С. КАРЛИН, В. СТАДДЕН

# ЧЕБЫШЕВСКИЕ СИСТЕМЫ И ИХ ПРИМЕНЕНИЕ В АНАЛИЗЕ И СТАТИСТИКЕ

Перевод с английского
под редакцией С. М. ЕРМАКОВА

ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
Москва 1976

517.8
К23
УДК 519.24

# TCHEBYCHEFF SYSTEMS: WITH APPLICATIONS IN ANALYSIS AND STATISTICS

SAMUEL KARLIN
WILLIAM J. STUDDEN

INTERSCIENCE PUBLISHERS
A. Division of John Wiley & Sons
New York • London • Sydney

*С. Карлин, В. Стадден*

Чебышевские системы и их применение
в анализе и статистике

М., 1976 г., 568 стр. с илл.

Редакторы *В. А. Абрамов, Е. Ю. Ходан*
Технический редактор *В. Н. Кондакова*
Корректоры *О. А. Сигал, Л. С. Сомова*

Сдано в набор 24.03. 1976 г. Подписано к печати 25 11 1976 г. Бумага $60\times90^{1}/_{16}$. Физ. печ. л. 35,5. Условн. печ л 35,5. Уч.-изд. л. 37,12. Тираж 6600 экз. Цена книги 2 р 91 к. Заказ 6-325.

Издательство «Наука»
Главная редакция физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

Киевская книжная типография научной книги Республиканского производственного объединения «Полиграфкнига» Госкомиздата УССР, Киев, Репина, 4.

$К\ \frac{20203-155}{053(02)-76}\ 43\text{-}76$

## ОГЛАВЛЕНИЕ

## ПРЕДИСЛОВИЕ РЕДАКТОРА

Предлагаемая советскому читателю книга американских математиков Самюэля Карлина и Вильяма Стаддена «Чебышевские системы и их применение в анализе и статистике» необычайно привлекательна богатством приложений излагаемой в ней фундаментальной теории. Приложения в области анализа, теории вероятностей, математической статистики и теории планирования эксперимента делают эту книгу весьма актуальной, несмотря на то что на русском языке имеется превосходная книга М. Г. Крейна и А. А. Нудельмана «Проблема моментов Маркова и экстремальные задачи», вышедшая в 1973 году, т. е. после выхода в свет книги С. Карлина и В. Стаддена. Обе эти книги в значительной своей части базируются на статье М. Г. Крейна «Идеи П. Л. Чебышева и А. А. Маркова в теории предельных величин интегралов и их дальнейшее развитие», опубликованной в УМН в 1951 году, и последняя в своей «теоретической» части заметно пересекается с первой. Читатель, интересующийся собственно экстремальными задачами, связанными с конечной проблемой моментов и их приложениями в теории аппроксимации, должен иметь в виду, что отдельные вопросы с большей полнотой освещены в книге М. Г. Крейна и А. А. Нудельмана. Вместе с тем нельзя не отметить, что даже в «области пересечения» изложение С. Карлина и В. Стаддена имеет свои несомненные достоинства и «теоретическая» часть их книги достаточно богата для того, чтобы она была интересна читателю, знакомому с книгой М. Г. Крейна и А. А. Нудельмана. Что же касается приложений, то американские авторы сумели сосредоточить свое внимание на областях, которые в последнее десятилетие получили широкое развитие. Это относится прежде всего к теории планирования эксперимента, которая благодаря важности ее приложений стала весьма актуальной в последние годы.

Десятая глава книги в основном посвящена теореме эквивалентности, играющей важную роль в теории планирования эксперимента, и с помощью этой теоремы в главе X строятся конкретные $D$-оптимальные планы регрессионных экспериментов. В §§ 9 и 10 главы XI содержится ряд результатов относительно аппроксимации функций обобщенными сплайн-многочленами. Как известно, эта область также бурно развивается в последние годы. Наконец, нельзя не отметить далеко не тривиальные приложения в области теории вероятностей, а также связь леммы Неймана — Пирсона с задачами чебышевского типа, излагаемыми в заключительной главе книги. Представляется особенно ценным, что столь разнородные приложения изложены с единых позиций.

Все это, а также то обстоятельство, что результаты русских и советских авторов достаточно полно представлены в книге С. Карлина и В. Стаддена, делает ее интересной для широкого круга читателей — математиков и специалистов, связанных с различными приложениями математики. Последовательное и ясное изложение делает книгу доступной в значительной своей части для студентов 3—4 курсов математико-механических специальностей, а также студентов, специализирующихся в области прикладной математики в университетах и технических вузах.

Перевод с английского выполнен В. А. Егоровым, В. Б. Меласом и Е. В. Седуновым.

*С. М. Ермаков*

## ПРЕДИСЛОВИЕ АВТОРОВ К АМЕРИКАНСКОМУ ИЗДАНИЮ

Чебышевская система (сокращенно $T$-система) состоит из множества непрерывных функций $\{u_i(t)\}_{i=0}^n$, определенных на вещественном интервале $[a, b]$ и характеризующихся тем свойством, что любая нетривиальная вещественная линейная комбинация $\sum_{i=0}^n a_i u_i(t)$ имеет, самое большее, $n$ различных нулей на $[a, b]$.

$T$-системы играют важную роль (иногда косвенно) во многих областях математического анализа, а именно: а) в теории аппроксимаций, где главными применениями являются интерполяционные методы, квадратурные формулы и методы обработки данных; b) в краевых задачах и задачах, связанных с осцилляционными свойствами нулей решений дифференциальных уравнений $n$-го порядка; с) в теории неравенств, особенно в статистических приложениях. Другими разделами, связанными с этими вопросами, являются: теория выпуклости, численный анализ, интегрирование, теория операторов и вариационные задачи, включающие классы многочленов.

Многие разделы геометрии классических моментных пространств, индуцированных специальной $T$-системой $\{1, t, t^2, \ldots, t^n\}$, и их ответвления могут быть обобщены на случай $T$-систем общего вида. Например, существенная часть теории аппроксимации, теории неравенств и осцилляционных свойств обычных многочленов может быть распространена на обобщенные многочлены, образованные $T$-системой. (Обобщенный многочлен определяется как выражение вида $u(t) = \sum_{i=0}^n a_i u_i(t)$, где $a_i$ — вещественные константы, а $\{u_i(t)\}_{i=0}^n$— заданная $T$-система.)

Первое систематическое изложение геометрической теории $T$-систем было опубликовано Крейном [1951]. Его работа оказала значительное влияние на многие разделы этой книги. Другой и, возможно, более фундаментальной основой для результатов, излагаемых здесь, явилась теория вполне положительных ядер. В 1959 г. соавтор этой книги, указанный первым, предпринял попытку разработать аналитическую теорию вполне положительных ядер с целью прояснить применения этой теории к стохастическим процессам диффузионного типа, к задачам теории принятия решений и к анализу движения механических систем в постановке Гантмахера и Крейна. В ходе этого исследования стало очевидно, что $T$-системы представляют собой особо важный феномен теории вполне положительных ядер. Ввиду многих уникальных свойств $T$-систем и их фундаментального значения в теории аппроксимации, а также вследствие геометрических формулировок и конструкций некоторых выпуклых тел в терминах $T$-систем, было признано целесообразным посвятить всю книгу исключительно этим системам. Имеется очень небольшое пересечение в содержании этой книги и ранее вышедшей двухтомной работы одного из авторов данной книги, фамилия которого указана первой, по теории вполне положительных ядер.

Мы не пытаемся здесь представить исчерпывающее изложение теории $T$-систем. В последние годы появился ряд замечательных книг по различным аспектам теории аппроксимации, например Мейнардус [1964] и Райс [1965]; любая из этих работ удовлетворила бы читателя, который интересуется общепринятым аналитическим изложением теории $T$-систем. Настоящая книга, напротив, представляет собой главным образом исследовательскую монографию. Около одной трети всех результатов являются новыми, и почти весь материал, включая обсуждения, излагается далее с единых позиций и часто в новой форме с акцентом на геометрический подход.

Мы искренне старались сопровождать изложение соответствующими замечаниями исторического характера и общими комментариями. Любая неточность или упущение в указаниях приоритетов на серьезные результаты являются неумышленными, и авторы будут глубоко сожалеть, если они встретятся. Все цитированные источники перечислены в библиографии в конце книги.

На наш взгляд, полезно сделать краткий обзор содержания книги.

В гл. I даются важные определения и приводится ряд конкретных примеров $T$-систем с целью продемонстрировать широкую область их приложения в математическом анализе. В гл. I устанавливаются также некоторые теоремы интерполяции и теоремы об основных осцилляционных свойствах обобщенных многочленов, образованных $T$-системами.

В гл. II изучается структура моментных пространств, а также дуальных к ним пространств, индуцированных $T$-системами $\{u_i(t)\}_{i=0}^n$, где $u_i(t)$ определена на конечном замкнутом интервале. Первая часть гл. III посвящается доказательству неравенства Маркова — Крейна; во второй части дополнительные разделы геометрической теории, изложенной в гл. II, интерпретируются как результаты решения некоторых экстремальных задач. В гл. IV приводятся некоторые конкретизации и усиления теории двух предшествующих глав для специальной $T$-системы $\{t^i\}_{i=0}^n$.

Моментные пространства, индуцированные $T$-системой, определенной на полубесконечном интервале, анализируются в гл. V. В гл. VI рассматриваются моментные пространства, связанные с $T$-системами периодических функций; используя соответствующие тождества, мы показываем, что $T$-система, определенная на $(-\infty, \infty)$, может трактоваться как $T$-система периодических функций. В гл. I—VI все рассматриваемые $T$-системы определены на интервале вещественной прямой; в гл. VII мы переходим к рассмотрению $T$-систем, определенных на дискретном подмножестве прямой, и анализируем некоторые примеры, включающие абсолютно монотонные функции.

В гл. VIII исследуются моментные пространства, индуцированные $T$-системами, на которые наложены ограничения; приведенное обсуждение устанавливает связь с классической теорией, касающейся множества значений векторной меры, и обобщает ее. Некоторые приложения теории $T$-систем к теории аппроксимации, а также некоторые обобщения и модификации известных неравенств Маркова — Бернштейна приводятся далее в гл. IX. Гл. X начинается с некоторых задач, имеющих отношение к интерполяционным многочленам; эти результаты затем используются при разработке теории оптимальных экспериментальных планов в регрессионном анализе.

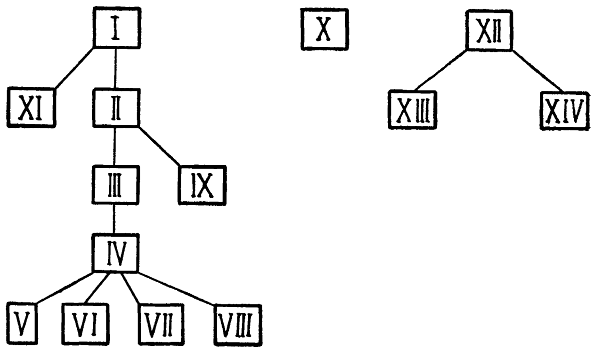
В гл. XI мы описываем некоторые естественные выпуклые конусы функций, определенные в терминах $T$-систем. Мы анализируем структуру этих конусов и дуальных к ним конусов, рассматривая различные классические неравенства (неравенства Фавара, Бервальда, Стеффенсена и другие), а также некоторые новые неравенства. В §§ 9 и 10 гл. XI излагаются некоторые новые результаты в теории аппроксимации функций обобщенными сплайн-многочленами.

В гл. XII—XIV приводится подробное рассмотрение различных типов чебышевских неравенств. В гл. XII мы получаем стандартные чебышевские неравенства, в которых ограничениями служат либо обычные моментные условия, либо условия типа условий гладкости (например, унимодальность, ограниченность), которые накладываются на плотность функции распределения. В гл. XIII рассматриваются несколько многомерных чебышевских неравенств. В гл. XIV обсуждаются различные классы нелинейных неравенств чебышевского типа.

Более полные комментарии приводятся во вводных параграфах к главам. Для удобства чтения приведена логическая зависимость глав.

*Самуэль Карлин, Вильям Стадден*

# ЛОГИЧЕСКАЯ ЗАВИСИМОСТЬ ГЛАВ




# СПИСОК ОСНОВНЫХ ОБОЗНАЧЕНИЙ

$R^r$—$r$-мерное евклидово пространство
$T$-система—система функций Чебышева
$CT$-система—полная система Чебышева
$WT$-система—слабая система Чебышева
$ET$-система—обобщенная система Чебышева
$U\begin{pmatrix}0,1,\ldots,n\\ t_0,t_1,\ldots,t_n\end{pmatrix}$—определитель $n+1$-го порядка с элементами $u_i(t_j)$, $i,j=0,1,\ldots,n$
$TP_k$—класс вполне положительных ядер порядка $k$
$STP_k$—класс обобщенных вполне положительных ядер порядка $k$
$D_j$—дифференциальный оператор
$Z(f)$—число различных нулей функции $f$
$\widetilde{Z}(f)$—число нулей функции $f$, когда неузловые нули учитываются дважды
$C^p$—класс $p$ раз непрерывно дифференцируемых функций
$S_\mu(f)$ — число существенных перемен знака функции $f$ относительно меры $\mu$
$Z^*(f)$—число нулей функции $f$ с учетом кратности
$\mathcal{M}_{n+1}$—пространство моментов
$M_{n+1}$—сечение $\mathcal{M}_{n+1}$
$C(G)$—выпуклый конус, натянутый на множество $G$
$C^+$—конус, дуальный к $C$
$I(G)$—индекс множества $G$
$DG$—граница множества $G$
$\mathrm{Int}\,G$—внутренняя часть множества $G$
$V(\mathbf{c})$—множество мер с фиксированным вектором моментов $\mathbf{c}$
$D(\mathbf{c})$—подкласс мер из $V(\mathbf{c})$, имеющих конечный спектр
$K(x,y)$—ядро интегрального оператора
$\rho(t)$—максимальная масса преставления
$\mathcal{P}_{n+1}$ — множество всех неотрицательных обобщенных многочленов порядка $n$
$u(t)$, $v(t)$—обобщенные многочлены
$l_\nu(x)$ — фундаментальный интерполяционный многочлен Лагранжа
$M(\xi)$—информационная матрица
$d(x,\xi)$—поверхность отклика
$\Omega_\Gamma(x)$—характеристическая функция множества $\Gamma$
$\mathrm{tr}\,\mathbf{A}$—след матрицы $\mathbf{A}$
$|\mathbf{A}|$—определитель матрицы $\mathbf{A}$
$T_n(x)$—многочлен Чебышева первого рода
$U_n(x)$—многочлен Чебышева второго рода
$L_n^{(\alpha)}(x)$—многочлен Лагерра
$H_n(x)$—многочлен Эрмита

Глава I

# СИСТЕМЫ ЧЕБЫШЕВА НА ЗАМКНУТОМ ИНТЕРВАЛЕ: ОПРЕДЕЛЕНИЯ, ПРИМЕРЫ И ПРЕДВАРИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

## § 1. Системы Чебышева

Определение 1.1. Пусть $u_0, u_1, \ldots, u_n$ обозначают непрерывные вещественные функции, определенные на замкнутом конечном интервале $[a, b]$. Эти функции будем называть *системой Чебышева* на $[a, b]$, сокращенно *T-системой* (без ссылки на интервал $[a, b]$, если это не приводит к неоднозначности), при условии, что определители $(n+1)$-го порядка [1])

$$U\begin{pmatrix} 0, & 1, & \ldots, & n \\ t_0, & t_1, & \ldots, & t_n \end{pmatrix} = \begin{vmatrix} u_0(t_0) & u_0(t_1) & \ldots & u_0(t_n) \\ u_1(t_0) & u_1(t_1) & \ldots & u_1(t_n) \\ \cdot & \cdot & \cdot & \cdot \\ u_n(t_0) & u_n(t_1) & \ldots & u_n(t_n) \end{vmatrix} \tag{1.1}$$

строго положительны всякий раз, когда $a \leqslant t_0 < t_1 < \ldots < t_n \leqslant b$. Функции $u_0, u_1, \ldots, u_n$ будем называть *полной системой Чебышева* (сокращенно *CT-системой*), если $\{u_0, u_1, \ldots, u_r\}$ является $T$-системой для любого $r = 0, 1, \ldots, n$.

Понятия, терминология и система обозначений в §§ 1 и 2 в большинстве своем общепринятые и классические (см., например, Ахиезер [1965] или Натансон [1949]).

Классическим примером $CT$-системы служит система функций $u_i(t) = t^i$, $i = 0, 1, \ldots, n$, определенная на любом конечном интервале $[a, b]$. В этом случае определитель в (1.1) сводится к известному определителю Вандермонда, равному

$$\prod_{0 \leqslant i < j \leqslant n} (t_j - t_i).$$

[1]) Если не оговорено противное, определители (1.1) имеют столбцы, расположенные так, что $t_i$ оказываются в возрастающем порядке.

Свойства чебышевской системы функций 1, $t$, $t^2$, . . . , $t^n$, их значение и применения обсуждаются во многих работах. Особенно рекомендуем книгу Полиа и Сеге [1925] (часть 5 гл. I), в которой содержатся и другие ссылки на классические работы в этой области.

В традиционном определении системы Чебышева требуется, чтобы определители в (1.1) сохраняли строго постоянный знак. Из непрерывности функций $u_0$, $u_1$, . . . , $u_n$ следует, что это требование эквивалентно условию, чтобы определители (1.1) никогда не обращались в нуль. Однако без потери общности можно нормировать систему умножением одной из функций, скажем $u_n$, на $+1$ или $-1$, и, следовательно, считать, что определители положительны. Формулировки большинства теорем в книге не зависят от нормировки функций, образующих чебышевскую систему, либо могут потребовать простых изменений. Если в случае $CT$-системы предполагается, что для каждого $r = 0,1$, . . . , $n$ определители сохраняют знак (возможно, зависящий от $r$), то можно очень просто изменить систему умножением каждой из функций соответственно на $+1$ или $-1$, чтобы сделать все определители положительными.

Системы функций Чебышева играют важную роль в различных областях математики, особенно в теории аппроксимации, в задачах интерполяции, в обобщенной проблеме моментов, в численном анализе, в исследовании осцилляционных свойств собственных функций задачи Штурма — Лиувилля, в обобщенной теории выпуклых тел, в теории неравенств и т. д.

Как отмечалось выше, прототипом системы Чебышева является набор степенных функций 1, $t$, $t^2$, . . . , $t^n$. Осцилляционные свойства обычных многочленов хорошо изучены, и их применимость в теории аппроксимации и смежных дисциплинах достаточно хорошо известна. Многие разделы из анализа систем Чебышева связаны с обобщениями осцилляционных и аппроксимативных свойств обычных многочленов.

Развитие теории и приложений $T$-систем, названных именем Чебышева, имеет давнюю историю. Среди математиков, которые внесли крупный вклад в эту теорию, были Бернштейн, Декарт, Хаар, Лагерр, Чебышев, Валле-Пуссен и другие. Существенно новые результаты и приложения были впоследствии получены Ахиезером, Гантмахером, Джексоном, Крейном, Натансоном, Полиа, Сеге и их учениками. За превосходными обзорами некоторых более ранних работ и соответствующими ссылками читатель может обратиться к Ахиезеру [1965], Джексону [1930], Натансону [1949] и Тиману [1963].

Многие разделы геометрической теории классической проблемы моментов и ее ответвлений в анализе (см. Карлин и Шепли [1953], М. Г. Крейн [1951], Рогозинский [1958], П. Розенблюм [1951]) могут быть распространены на $T$-системы. Эта возможность была обнаружена почти одновременно различными математиками, работавшими в этой области. Первое систематическое рассмотрение геометрической теории было предпринято Крейном [1951] (ср. Карлин и Шепли [1953]). Работа Крейна оказала влияние на некоторые разделы этой книги, что и будет отмечено в соответствующих местах. Геометрический подход является главным для нашей точки зрения, он лежит в основе доказательств, интерпретаций и технических приемов этой книги.

## § 2. Обобщенные и слабые системы Чебышева

Если система, состоящая из двух функций $u_0(t) \equiv 1$ и $u_1(t)$, образует $T$-систему на интервале $[a, b]$, то

$$\begin{vmatrix} 1 & 1 \\ u_1(t_1) & u_1(t_2) \end{vmatrix} > 0, \qquad a \leqslant t_1 < t_2 \leqslant b.$$

Функция $u_1$ является тогда строго возрастающей на интервале $[a, b]$. Условие монотонности функции $u_1$ может быть выражено в различных формах. Условие, что $u_1$ должна быть строго возрастающей, может быть ослаблено требованием только того, чтобы $u_1$ была неубывающей, т. е. $u_1(t_2) \geqslant u_1(t_1)$ для $t_2 > t_1$, или оно может быть усилено требованием того, чтобы $u_1$ была дифференцируемой и чтобы $u_1'(t) > 0$. Свойство неубывания функции $u_1$ может быть записано в виде следующего требования к определителю:

$$\begin{vmatrix} 1 & 1 \\ u_1(t_1) & u_1(t_2) \end{vmatrix} \geqslant 0, \qquad a \leqslant t_1 < t_2 \leqslant b, \tag{2.1}$$

в то время как условие $u_1'(t) > 0$ может быть представлено в форме

$$\begin{vmatrix} 1 & u_0'(t) \\ u_1(t) & u_1'(t) \end{vmatrix} = u_1'(t) > 0. \tag{2.2}$$

Два условия, записанные в форме (2.1) и (2.2), расширяют естественным образом систему функций $u_0, u_1, \ldots, u_n$.

Определение 2.1. Система $\{u_i\}_0^n$ непрерывных на $[a, b]$ функций называется *слабой чебышевской системой* на $[a, b]$, или *$WT$-системой,* если функции $u_0, u_1, \ldots, u_n$ линейно независимы на $[a, b]$ и определители

$$U\begin{pmatrix} 0, & 1, & \ldots, & n \\ t_0, & t_1, & \ldots, & t_n \end{pmatrix}$$

в (1.1) неотрицательны при $a \leqslant t_0 < t_1 < \ldots < t_n \leqslant b$.

Это определение является аналогом (2.1). Для того чтобы сформулировать аналог (2.2) для общих $T$-систем, необходимо обеспечить разумную оценку определителей (1.1), которая бы допускала равенства среди значений $t_i$.

Прежде чем приступить к этой задаче, важно принять во внимание, что $T$-система может рассматриваться как специальный случай функции ядра $K(s, t)$ двух вещественных переменных, которая обладает определенными свойствами знакорегулярности. Рассмотрим вещественную функцию $K(s, t)$, определенную для $(s, t) \in S \times T$, где $S$ и $T$ — заданные подмножества вещественной прямой. В случае, если $S$ состоит из конечного набора, который для удобства обозначим как $\{0, 1, \ldots, n\}$, ядро $K(s, t)$ можно представить как последовательность функций $K(i, t) = u_i(t)$, $i = 0, 1, \ldots, n$.

Для ядра общего вида определители (1.1) записываются как

$$K\begin{pmatrix} s_0, & s_1, \ldots, s_n \\ t_0, & t_1, \ldots, t_n \end{pmatrix} = \det \| K(s_i, t_j) \|_{i,j=0}^{n} \tag{2.3}$$

для $s_0 < s_1 < \ldots < s_n$, $t_0 < t_1 < \ldots < t_n$ и $(s_i, t_j) \in S \times T$, $i, j = 0, 1, \ldots, n$. Верхние значения $s_0, s_1, \ldots, s_n$ указывают строки в определителе, а нижние значения $t_0, t_1, \ldots, t_n$ — столбцы. Обозначения в (1.1) согласуются с обозначениями (2.3), если положить $s = \{0, 1, \ldots, n\}$, $T = [a, b]$ и $K(i, t) = u_i(t)$, $i = 0, 1, \ldots, n$.

Определение 2.2. Ядро $K(s, t)$ называется *вполне положительным порядка $k$*, сокращенно $(TP_k)$, если для каждого $r = 1, 2, \ldots, k$ имеем

$$K\begin{pmatrix} s_1, \ldots, s_r \\ t_1, \ldots, t_r \end{pmatrix} \geqslant 0 \tag{2.4}$$

всякий раз, когда $s_1 < s_2 < \ldots < s_r$, $t_1 < t_2 < \ldots < t_r$ и $(s_i, t_j) \in S \times T$, $i, j = 1, \ldots, r$. Если определители в (2.4) всегда положительны, то ядро $K(s, t)$ называется *строго вполне положительным порядка $k$ $(STP_k)$*.

Приведенные выше понятия вместе с другими условиями знакорегулярности обобщают определение $CT$-системы и встречаются во многих контекстах. Существует тщательно разработанная теория знакорегулярных ядер (см. Карлин [1967]). На всем протяжении этой книги внимание концентрируется исключительно на специальном случае, когда множество $T$ дискретно и конечно, что, таким образом, приводит нас к $T$-системам.

При формулировке аналога (2.2) для случая $T$-систем удобно для дальнейшего ввести понятие обобщенных систем для ядра общего вида. Если $T = [a, b]$ и если для каждого $s \in S$ функция $K(s, \cdot) \in C^{p-1}[a, b]$, $p \geqslant 1$, т. е. $K(s, t)$ имеет $p - 1$ непрерывных производных по $t$, то расширим понятие определителя, стоящего в правой части (2.4), чтобы допустить равенство, самое большее, среди $p$ из значений $t_i$. Для каждого набора равных $t_i$ заменим последовательные столбцы столбцами их последовательных производных. Точнее, если $s_0 < s_1 < \ldots < s_k$, $a \leqslant t_0 \leqslant t_1 \leqslant \ldots \leqslant t_k \leqslant b$ и

$t_{i-1} < t_i = t_{i+1} = \ldots = t_{i+q}$, $0 \leq q \leq p-1$, тогда

$$K^*\begin{pmatrix} s_0, & s_1, \ldots, s_k \\ t_0, & t_1, \ldots, t_k \end{pmatrix}$$

записывается как определитель (2.4), в котором $(i+1+j)$-й столбец, $0 \leq j \leq q$, заменяется вектор-столбцом [1])

$$\left(\frac{\partial^j}{\partial t_i^j} K(s_0, t_i), \quad \frac{\partial^j}{\partial t_i^j} K(s_1, t_i), \ldots, \frac{\partial^j}{\partial t_i^j} K(s_k, t_i)\right).$$

Например, если $K(i, t) = u_i(t)$ и $a \leq t_0 = t_1 = \ldots = t_q < t_{q+1} < \ldots$ $\ldots < t_{n-1} = t_n \leq b$, то

$$U^*\begin{pmatrix} 0, & 1, \ldots, n \\ t_0, & t_1, \ldots, t_n \end{pmatrix} =$$

$$= \begin{vmatrix} u_0(t_0) & u_0'(t_0) \ldots u_0^{(q)}(t_0) & u_0(t_{q+1}) \ldots u_0(t_{n-1}) & u_0'(t_{n-1}) \\ u_1(t_0) & u_1'(t_0) \ldots u_1^{(q)}(t_0) & u_1(t_{q+1}) \ldots u_1(t_{n-1}) & u_1'(t_{n-1}) \\ \cdot & \cdot & \cdot & \cdot \\ u_n(t_0) & u_n'(t_0) \ldots u_n^{(q)}(t_0) & u_n(t_{q+1}) \ldots u_n(t_{n-1}) & u_n'(t_{n-1}) \end{vmatrix}.$$

В случае $t_0 = t_1 = \ldots = t_n = t$, т. е. $q = n$, приведенный выше определитель сводится к определителю Вронского $W(u_0(t), u_1(t), \ldots$ $\ldots, u_n(t))$ функций $u_0, u_1, \ldots, u_n$. Аналогично, если функция $K$ дифференцируема по переменной $s$, можно допустить, чтобы встречались равные среди значений $s_i$, и внести соответствующие изменения в обозначения строк определителя. Наконец, если равенства встречаются как в том, так и в другом множестве, то соответствующие операции в строках и столбцах выполняются одновременно. Например, если $S$ и $T$ — два интервала, $s = s_0 =$ $= s_1 = \ldots = s_k$ и $t = t_0 = t_1 = \ldots = t_k$, то при условии, что функция $K(s, t)$ достаточно гладкая, имеем

$$K^*\begin{pmatrix} s, \ldots, s \\ t, \ldots, t \end{pmatrix} = \det \left\| \frac{\partial^{i+j}}{\partial s^i \partial t^j} K(s, t) \right\|_{i,j=0}^{k}.$$

Определение 2.3. Если функция $K(s, t)$ является достаточно гладкой по $t$, то говорят, что $K(s, t)$ — *обобщенное вполне положительное ядро порядка $k$ по отношению к $t$* ($ETP_k(t)$), если для каждого $r = 1, 2, \ldots, k$

$$K^*\begin{pmatrix} s_1, \ldots, s_r \\ t_1, \ldots, t_r \end{pmatrix} > 0$$

для всех наборов из $s_1 < s_2 < \ldots < s_r$, $t_1 \leq t_2 \leq \ldots \leq t_r$, $(s_i, t_j) \in$ $\in S \times T$, $i, j = 1, \ldots, r$. Аналогичное определение употребляется

---

[1]) Звездочка всегда будет означать, что оценка определителя, допускающая равенства среди значений $t$, берется в том виде, как здесь указано.

для $ETP_k(s)$, т. е. по отношению к переменной $s$. Когда равенства допускаются среди значений $s_i$ и (или) среди значений $t_i$, тогда говорят о $K(s, t)$ как о $ETP_k$ без указания на ту или другую переменную.

В случае, когда множество $S$ конечно, вышеприведенное определение сводится к определению 2.4.

Определение 2.4. Функции $u_0, u_1, \ldots, u_n$ будем называть *обобщенной чебышевской системой порядка* $p$ или *ET-системой порядка* $p$ при условии, что $u_i \in C^{p-1}[a, b]$, $i = 0, 1, \ldots, n$ и

$$U^*\begin{pmatrix} 0, & 1, \ldots, n \\ t_0, & t_1, \ldots, t_n \end{pmatrix} > 0 \tag{2.5}$$

для всех наборов $t_0 \leq t_1 \leq \ldots \leq t_n$ ($t_i \in [a, b]$), где равенство встречается в группах не более чем из $p$ последовательных значений $t_i$. Обобщенная система Чебышева порядка $n+1$ будет называться просто $ET$-системой.

Можно показать, что если $\{u_i\}_0^n$ является $T$-системой или $WT$-системой и функции $u_0, u_1, \ldots, u_n$ достаточно гладкие, то определители (2.5) удовлетворяют неравенствам

$$U^*\begin{pmatrix} 0, & 1, \ldots, n \\ t_0, & t_1, \ldots, t_n \end{pmatrix} \geqslant 0 \tag{2.6}$$

для $a \leq t_0 \leq t_1 \leq \ldots \leq t_n \leq b$.

Этот результат обобщает элементарный факт, который состоит в том, что если функция $u$ неубывающая и дифференцируемая, то $u'(t) \geqslant 0$.

Чтобы доказать (2.6), прежде всего заметим, что непосредственное применение теоремы о среднем значении к функции

$$g(t) = U\begin{pmatrix} 0, & 1, \ldots, n-1, & n \\ t_0, & t_1, \ldots, t_{n-1}, & t \end{pmatrix}, \qquad a \leq t_0 < t_1 \ldots < t_{n-1} < t \leq b,$$

показывает, что

$$U\begin{pmatrix} 0, & 1, \ldots, n \\ t_0, & t_1, \ldots, t_n \end{pmatrix} = (t_n - t_{n-1}) \begin{vmatrix} u_0(t_0) & \ldots & u_0(t_{n-1}) & u_0'(\xi_n) \\ u_1(t_0) & \ldots & u_1(t_{n-1}) & u_1'(\xi_n) \\ \cdot & \cdot & \cdot & \cdot \\ u_n(t_0) & \ldots & u_n(t_{n-1}) & u_n'(\xi_n) \end{vmatrix},$$

где $t_{n-1} < \xi_n < t_n$, так что

$$\operatorname{sign} U\begin{pmatrix} 0, & 1, \ldots, n \\ t_0, & t_1, \ldots, t_n \end{pmatrix} = \operatorname{sign} \begin{vmatrix} u_0(t_0) \ldots & u_0(t_{n-1}) & u_0'(\xi_n) \\ u_1(t_0) \ldots & u_1(t_{n-1}) & u_1'(\xi_n) \\ \cdot & \cdot & \cdot \\ u_n(t_0) \ldots & u_n(t_{n-1}) & u_n'(\xi_n) \end{vmatrix}.$$

Вторичное применение теоремы о среднем значении дает

$$\operatorname{sign} U\begin{pmatrix}0, & \ldots, & n\\ t_0, & \ldots, & t_n\end{pmatrix} = \operatorname{sign}\begin{vmatrix} u_0(t_0)\ldots u_0(t_{n-2}) & u_0'(\xi_{n-1}) & u_0'(\xi_n)\\ \cdot\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot\cdot\\ u_n(t_0)\ldots u_n(t_{n-2}) & u_n'(\xi_{n-1}) & u_n'(\xi_n)\end{vmatrix},$$

где $t_{n-2} < \xi_{n-1} < t_{n-1} < \xi_n < t_n$.

Повторение этой процедуры $[n(n+1)]/2$ раз образует последовательность $\eta_1, \eta_2, \eta_3, \ldots, \eta_n$ такую, что

$$0 \leqslant \operatorname{sign} U\begin{pmatrix}0, & 1, & \ldots, & n\\ t_0, & t_1, & \ldots, & t_n\end{pmatrix} = \operatorname{sign}\begin{vmatrix} u_0(t_0) & u_0'(\eta_1)\ldots & u_0^{(n)}(\eta_n)\\ u_1(t_0) & u_1'(\eta_1)\ldots & u_1^{(n)}(\eta_n)\\ \cdot\cdot\cdot & \cdot\cdot\cdot & \cdot\cdot\cdot\\ u_n(t_0) & u_n'(\eta_1)\ldots & u_n^{(n)}(\eta_n)\end{vmatrix},$$

где $t_0 < \eta_1 < \eta_2 < \ldots < \eta_n < t_n$. Если теперь устремить $t_n$ к $t_0$, то легко видеть, что

$$U^*\begin{pmatrix}0, & 1, & \ldots, & n\\ t_0, & t_0, & \ldots, & t_0\end{pmatrix} \geqslant 0.$$

Подобное рассуждение может быть применено к любому множеству из совпадающих значений $t_i$, откуда и следует (2.6).

Классический пример $u_i(t) = t^i$, $i = 0, 1, \ldots, n$, является $ET$-системой. Действительно, поскольку $\{t^i\}_0^n$ есть $T$-система, определители (2.5) по крайней мере все неотрицательны. Однако если для некоторого множества значений $t_i$, $i = 0, 1, \ldots, n$, удовлетворяющего условию $a \leq t_0 \leq t_1 \leq \ldots \leq t_n \leq b$, выполняется

$$U^*\begin{pmatrix}0, & 1, & \ldots, & n\\ t_0, & t_1, & \ldots, & t_n\end{pmatrix} = 0,$$

то можно образовать нетривиальный многочлен $\sum_{i=0}^{n} a_i t^i$, имеющий $n+1$ нулей с учетом кратности. Это противоречие означает, что $\{t^i\}_0^n$ является $ET$-системой, как и утверждалось.

В заключение этого параграфа приведем эквивалентную формулировку понятия $ET$-системы порядка $p$. Эти идеи не будут использоваться в оставшейся части книги, однако они обещают дальнейшие обобщения.

Для специальной системы $\{t^i\}_0^n$ обозначим определители (1.1) как

$$F\begin{pmatrix}0, & 1, & \ldots, & n\\ t_0, & t_1, & \ldots, & t_n\end{pmatrix},$$

и пусть $\Delta_{n+1} = \{\bar{t} = (t_0, t_1, \ldots, t_n) \mid a \leqslant t_0 < t_1 < \ldots < t_n \leqslant b\}$ и $\bar{\Delta}_{n+1} = \{\bar{t} \mid a \leqslant \leqslant t_0 \leqslant t_1 \leqslant \ldots \leqslant t_n \leqslant b\}$.

Грани симплекса $\Delta_{n+1}$ характеризуются равенствами, встречающимися среди значений $t_i$, или равенством с граничными точками $a$ и $b$. Определитель (1.1) может рассматриваться как функция, заданная на открытом симплексе $\Delta_{n+1}$.

Понятие $ET$-системы дает естественный метод продолжения этой функции до границы $\Delta_{n+1}$, т. е. до $\bar{\Delta}_{n+1}$.

Говорят, что система $\{u_i\}_0^n$ является $ET$-системой порядка $p$, если $u_i \in C^{p-1}[a, b]$, $i = 0, 1, \ldots, n$, и

$$\lim_{\bar{s}\to\bar{t}} \frac{U\begin{pmatrix}0, & 1, & \ldots, & n\\ s_0, & s_1, & \ldots, & s_n\end{pmatrix}}{F\begin{pmatrix}0, & 1, & \ldots, & n\\ s_0, & s_1, & \ldots, & s_n\end{pmatrix}} > 0, \tag{2.7}$$

где $\bar{s} \in \Delta_{n+1}$, $\bar{t} \in \bar{\Delta}_{n+1}$, и самое большее $p$ последовательных компонент в $\bar{t}$ равны. Это определение и определение 2.4 эквивалентны, так как предел левой части (2.7) в действительности равен

$$\frac{U^*\begin{pmatrix}0, & 1, & \ldots, & n\\ t_0, & t_1, & \ldots, & t_n\end{pmatrix}}{F^*\begin{pmatrix}0, & 1, & \ldots, & n\\ t_0, & t_1, & \ldots, & t_n\end{pmatrix}}, \tag{2.8}$$

а система $\{t^i\}_0^n$ является $ET$-системой, что делает знаменатель в (2.8) всегда положительным. Утверждение, что (2.8) совпадает с левой частью (2.7), доказывается путем неоднократно повторенного применения теоремы о среднем значении подобно выводу выражения (2.6) (см. также Полиа и Сеге [1925], т. II, стр. 54, задача 95). Мы опускаем формальное доказательство этого факта.

Обобщение определения $ET$-системы может теперь быть получено заменой знаменателя в (2.7). Например, если $\{u_i\}_0^n$ есть система непрерывных функций и $\{\varphi_i\}_0^n$ — $T$-система, рассмотрим

$$\lim_{\bar{s}\to\bar{t}} \frac{U\begin{pmatrix}0, & 1, & \ldots, & n\\ s_0, & s_1, & \ldots, & s_n\end{pmatrix}}{\Phi\begin{pmatrix}0, & 1, & \ldots, & n\\ s_0, & s_1, & \ldots, & s_n\end{pmatrix}},$$

где $\bar{s} \in \Delta_{n+1}$, $\bar{t} \in \bar{\Delta}_{n+1}$ и $\Phi(i, s) = \varphi_i(s)$, $i = 0, 1, \ldots, n$. Если этот предел существует и положителен для всех $\bar{t} \in \bar{\Delta}_{n+1}$, то система $\{u_i\}_0^n$ называется $ET$-системой по отношению к системе $\{\varphi_i\}_0^n$. (В дальнейшем по этим вопросам отсылаем читателя к Карлину [1967].)

## § 3. Примеры

В этом параграфе основное внимание уделяется ряду важных примеров $T$-систем, $ET$-систем и т. д. и описываются методы образования новых $T$-систем из типовых. Некоторые примеры, перечисленные ниже, известны и играют важную роль в различных областях математики, включая применения в теории аппроксимации и интерполяции.

Пример 1. Степенные функции. В первом параграфе указывалось, что функции $u_i(t) = t^i$, $i = 0, 1, \ldots, n$, образуют $CT$-систему, а фактически $ET$-систему. На самом деле в этом случае можно получить более сильный результат, а именно, если $\alpha_0, \alpha_1, \ldots$

$\ldots, \alpha_n$ есть любая строго возрастающая последовательность вещественных чисел, то система $\{t^{\alpha_i}\}_0^n$ является $T$-системой на любом замкнутом подынтервале интервала $(0, \infty)$.

То, что $\{t^{\alpha_i}\}_0^n$ является $T$-системой на $(0, \infty)$, следует (так как $t^\alpha = \exp(\alpha \ln t)$ для $t > 0$) из того факта, что

$$\det [\, \| \exp(x_i y_j) \|_{i,j=0}^n ] > 0, \tag{3.1}$$

если $-\infty < x_0 < x_1 < \ldots < x_n < \infty$ и $-\infty < y_0 < y_1 < \ldots$ $\ldots < y_n < \infty$.

Докажем неравенство (3.1), показав, что определители в (3.1) никогда не обращаются в нуль. Если знать, что определитель в (3.1) нигде в нуль не обращается, то легко установить индукционным доказательством по порядку определителя, что знак его положителен. Индукция выполняется при фиксированном $x_n > 0$ и последующем стремлении $y_n$ к $\infty$.

Необращение в нуль выражения (3.1) устанавливается, если показать, что любая функция вида

$$w_n(y) = \sum_{i=0}^n a_i e^{x_i y} \tag{3.2}$$

$\left(a_i \text{ вещественное, } i = 0, 1, \ldots, n, \text{ где } \sum_{i=0}^n a_i^2 > 0\right)$ допускает самое большее $n$ различных вещественных нулей. То, что экспоненциальный многочлен вида (3.2) имеет самое большее $n$ нулей с учетом кратности, доказывают по индукции, рассматривая функцию $\frac{d}{dy}\{w_{n+1}(y) \exp(-x_{n+1} y)\}$ и применяя теорему Ролля.

Это свойство представляет собой, конечно, специальный случай классического результата, согласно которому тождественно не равный нулю экспоненциальный многочлен $\sum_{i=0}^n q_i(y) e^{c_i y}$, где $q_i(y)$ есть обыкновенный алгебраический многочлен с вещественными коэффициентами степени $k_i$, допускает самое большее $\left[\sum_{i=0}^n (k_i + 1)\right] - 1$ нулей с учетом кратности (см. Полиа и Сеге [1925], т. II, стр. 48, задача 75).

**Пример 2. Элементарные операции, сохраняющие характерные свойства $T$-систем.** Существует много простых преобразований, которые изменяют данный набор функций $u_0, u_1, \ldots, u_n$, но сохраняют справедливость любых неравенств, которым удовлетворяют определители (1.1). Следующие две элементарные операции переводят $T$-систему в новую $T$-систему.

a) Если функция $r(t)$ положительна и непрерывна для $t \in [a, b]$, а $\{u_i\}_0^n$ является $T$-системой, то функции $v_i(t) = r(t) u_i(t)$, $i = 0, 1, \ldots, n$, образуют $T$-систему, так как

$$V\begin{pmatrix} 0, & 1, & \ldots, & n \\ t_0, & t_1, & \ldots, & t_n \end{pmatrix} = \left(\prod_{i=0}^{n} r(t_i)\right) U\begin{pmatrix} 0, & 1, & \ldots, & n \\ t_0, & t_1, & \ldots, & t_n \end{pmatrix}.$$

b) Если $r(t)$ есть строго возрастающая и непрерывная функция, определенная на $[c, d]$ со значениями в $[a, b]$, то система функций $w_i(t) = u_i(r(t))$, $i = 0, 1, \ldots, n$, является $T$-системой на интервале $[c, d]$, если только $\{u_i\}_0^n$ есть $T$-система на $[a, b]$.

Пример 3. $T$-системы, порожденные $TP$-ядрами. Большое многообразие различных типов $T$-систем может быть получено из ядер $K(s, t)$, удовлетворяющих некоторым условиям, накладываемым на определитель, таким, как в определении 2.2 или 2.3. Например, если ядро $K(s, t)$ принадлежит классу $STP_{n+1}$, где $T = [a, b]$, то система $\{K(s_i, t)\}_{i=0}^n$, где $s_0 < s_1 < \ldots < s_n$, образует $CT$-систему на $[a, b]$, если $K(s_i, t)$ — непрерывная функция от $t$ для каждого $s_i$, $i = 0, 1, \ldots, n$. Точно так же, если ядро $K(s, t)$ принадлежит классу $ETP_{n+1}(t)$, где $T = [a, b]$, то $\{K(s_i, t)\}_{i=0}^n$ есть $ET$-система. Следующие два примера конкретно иллюстрируют этот метод образования $T$-систем.

Пример 4. Ядро Коши. Ядро $K(s, t) = 1/(s+t)$ принадлежит классу $STP$ всех порядков при $s > 0$, $t > 0$. Определитель (1.1) в этом случае сводится к определителю Коши, значение которого вычисляется по формуле

$$K\begin{pmatrix} s_0, & \ldots, & s_n \\ t_0, & \ldots, & t_n \end{pmatrix} = \frac{\prod\limits_{j>i}(s_j - s_i)(t_j - t_i)}{\prod\limits_{i,j=0}^{n}(s_i + t_j)}.$$

Система $u_i(t) = 1/(s_i + t)$ для $i = 0, 1, \ldots, n$, где $0 < s_0 < s_1 < \ldots < s_n$, является, следовательно, $CT$-системой на любом замкнутом подынтервале интервала $(0, \infty)$.

Ядро Коши играет фундаментальную роль в установлении хорошо известной теоремы Мюнтца в задаче о наилучшей аппроксимации многочленами.

Пример 5. Ядро Гаусса. Ядро $K(s, t) = \exp[-(s-t)^2]$ принадлежит классу $ETP(t)$ всех порядков на $S = T = (-\infty, \infty)$. В действительности ядро $K(s, t)$ принадлежит классу $ETP$ по обеим переменным. Однако этот результат здесь не понадобится. Чтобы доказать, что ядро $K(s, t)$ принадлежит классу $ETP(t)$, прежде всего сошлемся на (3.1) и заметим, что ядро $e^{st}$ принадлежит классу $STP$ на $(-\infty, \infty)$. Теперь, записав $K(s, t)$ в виде $K(s, t) = \exp(-s^2) \times$

$\times \exp(2st)\exp(-t^2)$, можно утверждать, как в примере 2, что ядро $K(s, t)$ принадлежит классу $STP$. Далее покажем, что любая функция вида $v_n(t) = \sum_{i=0}^{n} a_i \exp[-(s_i - t)^2]$ имеет самое большее $n$ вещественных нулей с учетом кратности, если $s_0 < s_1 < s_2 < \ldots < s_n$, при условии, что $a_i$ все вещественны и $\sum_{i=0}^{n} a_i^2 > 0$. Чтобы доказать это, запишем

$$v_n(t) = e^{-t^2} \sum_{i=0}^{n} a_i e^{-s_i^2} e^{2s_i t} \tag{3.3}$$

и используем факт, указанный в примере 2, что любая функция вида $\sum_{i=0}^{n} b_i e^{2s_i t}$ имеет самое большее $n$ нулей с учетом кратности. Отсюда следует, что ядро $K(s, t)$ принадлежит классу $ETP$ и что

$$K^*\begin{pmatrix} s_0, & \ldots, & s_n \\ t_0, & \ldots, & t_n \end{pmatrix} > 0, \tag{3.4}$$

если

$$s_0 < s_1 < \ldots < s_n; \quad t_0 \leq t_1 \leq \ldots \leq t_n. \tag{3.5}$$

Пример 6. Собственные функции интегральных операторов, ядра которых принадлежат классу $ETP$. Более фундаментальный класс примеров $T$-систем и $CT$-систем получается следующим образом. Возьмем $K(s, t)$ из класса $C^\infty$ и $ETP$ всех порядков при $s$ и $t$, содержащихся в $[a, b]$, и определим линейный оператор $T$ с областью $D(T) = L_2[a, b]$ посредством равенства

$$T\varphi(s) = \int_a^b K(s, t)\,\varphi(t)\,dt.$$

Спектр оператора $T$ состоит из счетного набора простых положительных собственных чисел $\lambda_0 > \lambda_1 > \ldots$, стремящихся к нулю с ростом $n$. Если через $\varphi_0, \varphi_1, \ldots, \varphi_n, \ldots$ обозначить соответствующие собственные функции, которые определяются однозначно с точностью до постоянного множителя, то система $\{\varphi_i\}_0^n$ является $ET$-системой с точностью до знаков функций. То есть для каждого фиксированного $n$ определители

$$\Phi\begin{pmatrix} 0, & \ldots, & n \\ t_0, & \ldots, & t_n \end{pmatrix},$$

где $\Phi(i, t) = \varphi_i(t)$, имеют строго постоянный знак $\varepsilon_n$. Если умножить каждую из функций подходящим образом на $\pm 1$, то эта система преобразуется в истинную $ET$-систему.

Общий класс ядер, удовлетворяющих требуемым условиям, включает ядра вида

$$K(s, t) = \int_0^\infty \exp[u(s)\alpha(\eta)] \exp[v(t)\beta(\eta)]\, d\mu(\eta), \qquad (3.6)$$

$$a \leq s, \quad t \leq b,$$

где $u$, $v$ из класса $C^\infty$, $u'(s)\, v'(s) > 0$, $\alpha(\eta)$ и $\beta(\eta)$ — строго возрастающие функции, а $\mu$ — $\sigma$-конечная мера с бесконечным числом точек роста. Проверку того, что при этих условиях ядро (3.6) принадлежит классу $ETP$, и анализ осцилляционных свойств соответствующих собственных функций $\{\varphi_i\}$ можно найти у Карлина [1967] (см. также Карлин [1964 с]).

Пример 7. Собственные функции операторов Штурма — Лиувилля. Задачи на собственные значения операторов Штурма — Лиувилля также приводят к примерам $T$-систем и $TP$-ядер. Пусть оператор $L(\varphi)$ определяется как

$$L(\varphi) = -\frac{d}{dx}\left(p\frac{d\varphi}{dx}\right) + q\varphi \qquad (3.7)$$

(функция $p$ непрерывна и положительна на $[a, b]$). Рассмотрим задачу на собственное значение $L\varphi = \lambda\varphi$ при граничных условиях

$$\begin{aligned} &\varphi(x)\sin\alpha - p(x)\frac{d\varphi(x)}{dx}\cos\alpha\,\Big|_{x=a} = 0, \\ &\varphi(x)\sin\beta + p(x)\frac{d\varphi(x)}{dx}\cos\beta\,\Big|_{x=b} = 0. \end{aligned} \qquad (3.8)$$

Соответствующая функция Грина имеет вид

$$K(x, y) = \begin{cases} \psi(x)\chi(y), & a \leq x \leq y \leq b, \\ \psi(y)\chi(x), & a \leq y < x \leq b, \end{cases} \qquad (3.9)$$

где $\psi$ и $\chi$ удовлетворяют граничным условиям при $a$ и $b$ соответственно. Гантмахер и Крейн [1950] показали, что ядро (3.9) принадлежит классу $TP$ при условии, что $\psi(x)\chi(x) > 0$ для всех $a \leq x \leq b$ (последующие результаты приведены у Карлина [1967]).

Если итерация $K^{(r)}(x, y)$ ядра $K(x, y)$ принадлежит соответствующему классу непрерывности $C^n$ и $ETP_{n+1}$ на $(a, b) \times (a, b)$, то собственные функции $\varphi_0, \varphi_1, \ldots, \varphi_n$ дифференциального оператора (3.7), подчиненные граничным условиям (3.8), образуют $ET$-систему на любом замкнутом подынтервале интервала $(a, b)$.

Вышеприведенные результаты обобщаются на случай некоторых линейных дифференциальных операторов более высокого порядка. Для этого зададим последовательность дифференциальных операторов первого порядка $L_\nu u(x) = (r_\nu(x)\, u(x))'$, $\nu = 1, \ldots, 2k$, где $r_\nu(x)$ напрерывна и положительна на $[a, b]$.

Предположим, что $r_\nu(x) = r_{2k+1-\nu}(x)$, $\nu = 1, 2, \ldots, k$, и рассмотрим дифференциальный оператор

$$Mu = (-1)^k \left(\prod_{\nu=1}^{2k} L_\nu\right) u. \tag{3.10}$$

В случае $k = 2$ выражение (3.10) сводится к выражению, описывающему классический дифференциальный оператор источника колебаний.

У Карлина [1967] доказывается, что при определенных граничных условиях, при которых оператор (3.10) допускает только положительный дискретный спектр, соответствующая функция Грина $G(x, y)$ вполне положительна. Соответствующие собственные функции образуют систему Чебышева, как и в случае оператора Штурма — Лиувилля.

Пример 8. Композиционная формула. Для того чтобы облегчить построение (см. пример 2) новых $T$-систем из типовых, отвлечемся от обсуждения специфических примеров, с тем, чтобы представить на рассмотрение важную композиционную формулу. Эта формула послужит удобным инструментом для сглаживания $T$-систем до $ET$-систем.

Предположим, что $K$, $L$ и $M$ — функции двух переменных и что эти функции связаны равенством

$$M(s, t) = \int_{[c, d]} K(s, \eta)\, L(\eta, t)\, d\mu(\eta), \tag{3.11}$$

где $\mu$ — $\sigma$-конечная мера на интервале $[c, d]$. Предполагается, что интеграл в (3.11) существует для каждого $(s, t) \in S \times T$. Тогда можно выразить определители, включающие функцию $M$, через определители, включающие функции $K$ и $L$, посредством формулы

$$M\begin{pmatrix} s_1, \ldots, s_k \\ t_1, \ldots, t_k \end{pmatrix} =$$

$$= \int \cdots \int_{c \leqslant \eta_1 < \eta_2 < \ldots < \eta_k \leqslant d} K\begin{pmatrix} s_1, \ldots, s_k \\ \eta_1, \ldots, \eta_k \end{pmatrix} L\begin{pmatrix} \eta_1, \ldots, \eta_k \\ t_1, \ldots, t_k \end{pmatrix} d\mu(\eta_1) \ldots d\mu(\eta_k), \tag{3.12}$$

$$s_1 < s_2 < \ldots < s_k, \quad t_1 < t_2 < \ldots < t_k \text{ и } (s_i, t_j) \in S \times T,$$
$$i, j = 1, \ldots, k.$$

Если каждая из переменных $s$, $t$ и $\eta$ дискретна и конечна, то равенство (3.11) соответствует произведению матриц, а равенство (3.12) сводится к формуле Бине — Коши для вычисления миноров $M$ через миноры $K$ и $L$. Для доказательства формулы (3.12) отсылаем читателя к Полиа и Сеге [1925], т. I, стр. 48, задача 68.

В некоторых приложениях формулы (3.12) будет необходим ее усиленный вариант, допускающий равенства среди значений $t$. Если для каждого $s \in S$ и $\eta \in [c, d]$ функции $M(s, t)$ и $L(\eta, t)$ принадлежат классу непрерывных функций $C^{k-1}[a, b]$ по переменной $t$ и

$$\frac{\partial^i}{\partial t^i} M(s, t) = \int_{[c,d]} K(s, \eta)\left(\frac{\partial^i}{\partial t^i} L(\eta, t)\right) d\mu(\eta), \qquad i = 0, 1, \ldots, k-1,$$

то

$$M^*\begin{pmatrix} s_1, \ldots, s_k \\ t_1, \ldots, t_k \end{pmatrix} =$$

$$= \int \ldots \int_{c \leq \eta_1 < \ldots < \eta_k \leq d} K\begin{pmatrix} s_1, \ldots, s_k \\ \eta_1, \ldots, \eta_k \end{pmatrix} L^*\begin{pmatrix} \eta_1, \ldots, \eta_k \\ t_1, \ldots, t_k \end{pmatrix} d\mu(\eta_1) \ldots d\mu(\eta_k), \quad (3.13)$$

где $s_1 < s_2 < \ldots < s_k$ и $a \leq t_1 \leq \ldots \leq t_k \leq b$.

Формулы (3.12) и (3.13) в дальнейшем будем именовать *основными композиционными формулами*.

С помощью формулы (3.12) можно вывести некоторые неравенства относительно определителей, включающих функцию $M$, из аналогичных неравенств относительно определителей, включающих функции $K$ и $L$. Например, если функции $K$ и $L$ принадлежат классу $STP_k$, то функция $M$, очевидно, обладает тем же свойством, если мера $\mu$ имеет спектр, состоящий по крайней мере из $k$ точек. Если $u$, $v$, $\alpha$ и $\beta$ — строго возрастающие непрерывные функции на $(-\infty, \infty)$, то, согласно неравенству (3.1) и примеру 2, ядра $\exp[u(s)\alpha(\eta)]$ и $\exp[v(t)\beta(\eta)]$ принадлежат классу $STP$ всех порядков, поэтому функция

$$M(s, t) = \int_a^b \exp[u(s)\alpha(\eta)] \exp[v(t)\beta(\eta)]\, d\eta \qquad (3.14)$$

принадлежит классу $STP$ всех порядков.

Введение формул (3.12) и (3.13) преследует две цели. Пусть $S = \{0, 1, \ldots, n\}$ и $K(i, t) = u_i(t)$ для $i = 0, 1, \ldots, n$, и $t \in [a, b]$; определим функцию

$$v_i(t) = \int_a^b u_i(\eta) L(\eta, t)\, d\eta. \qquad (3.15)$$

Равенство (3.12) в этом случае сводится к следующему:

$$V\begin{pmatrix} 0, 1, \ldots, n \\ t_0, t_1, \ldots, t_n \end{pmatrix} = \int \ldots \int_{a \leq \eta_0 < \ldots < \eta_n \leq b} U\begin{pmatrix} 0, \ldots, n \\ \eta_0, \ldots, \eta_n \end{pmatrix} L\begin{pmatrix} \eta_0, \ldots, \eta_n \\ t_0, \ldots, t_n \end{pmatrix} d\eta_1 \ldots d\eta_n.$$

Если ядро $L(\eta, t)$ принадлежит классу $STP$ порядка $n+1$ (определение 2.2), а система $\{u_i\}_0^n$ является $T$-системой, то новая система $\{v_i\}_0^n$ является также $T$-системой. Этим способом новые $T$-системы могут быть легко образованы из типовых подстановкой множества различных ядер $L(\eta, t)$.

Формулы (3.12) и (3.13) будут использованы и для второй, возможно, более важной цели. Они снабдят нас стандартным процессом возмущения или сглаживания, который даст возможность преобразовывать $WT$-системы или $T$-системы в $ET$-системы, аппроксимируя исходные системы с требуемой точностью. Например, если возьмем

$$L_\sigma(\eta, t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\eta - t}{\sigma}\right)^2\right], \qquad \sigma > 0, \tag{3.16}$$

то (так как ядро (3.16) принадлежит классу $ETP(t)$ всех порядков (пример 3)) из (3.13) следует, что если $\{u_i\}_0^n$ является $T$-системой, то система $\{v_i(t, \sigma)\}_0^n$, полученная из (3.15), является $ET$-системой. Кроме того, легко проверить, что $\lim\limits_{\sigma \to 0} v_i(t, \sigma) = u_i(t)$ для всех $t$ в открытом интервале $(a, b)$. Процедура, которую мы выберем в конкретном случае, должна будет доказать требуемое свойство для модифицированной системы $\{v_i(t, \sigma)\}_0^n$, а затем использовать предельный переход при $\sigma \to 0$, чтобы вывести этот же результат для первоначальной системы.

Аппарат ядер Гаусса для аппроксимации $T$-систем $ET$-системами уже использовался Шенбергом [1953], Гантмахером и Крейном [1950] и Карлином [1967] для подобных целей.

Пример 9. Треугольное ядро. Ядро, определенное на $S \times T$, где $S = T = (-\infty, \infty)$, как

$$K(s, t) = \begin{cases} 1, & s \geqslant t, \\ 0, & s < t, \end{cases} \tag{3.17}$$

принадлежит классу $TP$ всех порядков. Покажем, что для $s_1 < s_2 < \ldots < s_k$ и $t_1 < t_2 < \ldots < t_k$ определители

$$K\begin{pmatrix} s_1, s_2, \ldots, s_k \\ t_1, t_2, \ldots, t_k \end{pmatrix} \tag{3.18}$$

равны нулю, за исключением случая $t_1 \leqslant s_1 < t_2 \leqslant s_2 < \ldots < t_k \leqslant s_k$, когда их значения равны единице.

Если $s_1 < t_1$, определитель (3.18) равен нулю, так как первая строка имеет все нулевые компоненты. Следовательно, для получения ненулевого значения требуется, чтобы $t_1 \leqslant s_1$. Если $t_1 < t_2 \leqslant s_1$, то первые два столбца являются идентичными, с компонентами, равными единице, и определители (3.18) равны нулю. Остается проверить возможность $t_1 \leqslant s_1 < t_2$. В этом случае первая строка имеет все нулевые компоненты, за исключением первого элемента, и тогда

$$K\begin{pmatrix} s_1, s_2, \ldots, s_k \\ t_1, t_2, \ldots, t_k \end{pmatrix} = K\begin{pmatrix} s_2, \ldots, s_k \\ t_2, \ldots, t_k \end{pmatrix}.$$

Продолжая этот процесс снижения порядка, видим, что определители (3.18) равны нулю, если только не выполняются неравенства $t_1 \leq s_1 < t_2 \leq s_2 < \ldots < t_k \leq s_k$, а в этом случае значение (3.18) равно единице.

Для ядра $K(s, t)$, определенного в (3.17), положим $u_i(t) = K(t, \xi_i)$, $i = 0, 1, \ldots, n$, где $\xi_0 < \xi_1 < \ldots < \xi_n$. Простая проверка показывает, что функции $u_0, u_1, \ldots, u_n$ линейно независимы. Система $\{u_i\}_0^n$ является, очевидно, $WT$-системой на любом интервале $[a, b]$, когда $a \leq \xi_0 < \xi_n \leq b$.

Пример 10. Усеченные степенные функции. Если определить для фиксированного неотрицательного целого $m$ функцию

$$x_+^m = \begin{cases} x^m, & x \geqslant 0, \\ 0, & x < 0, \end{cases} \tag{3.19}$$

то система функций $\{u_i\}_0^n$, где $u_i(t) = (t - \xi_i)_+^m$, $i = 0, 1, \ldots, n$ и $\xi_0 < \xi_1 < \ldots < \xi_n$ фиксированы, образует $WT$-систему на любом интервале $[a, b]$ при условии, что $a \leq \xi_0$ и $\xi_n \leq b$. Заметим, что пример 9 есть специальный случай данного, когда $m = 0$.

Чтобы доказать, что функции $(t - \xi_i)_+^m$, $i = 0, 1, \ldots, n$, образуют $WT$-систему, покажем, что функция

$$g_m(x) = \begin{cases} \dfrac{x^m e^{-\beta x}}{m!}, & x \geqslant 0, \\ 0, & x < 0 \end{cases}$$

($\beta \geqslant 0$), обладает тем свойством, что $g_m(x - y)$ принадлежит классу $TP$ всех порядков на $(-\infty, \infty)$. Функции с этим свойством известны как частотные функции Полиа (см. Шенберг [1951]). Выбор $\beta = 0$ немедленно приведет к желаемому результату.

В этом разделе мы впервые отмечаем, что если функции $f(x)$ и $g(x)$ таковы, что $f(x - y)$ и $g(x - y)$ принадлежат классу $TP$ на $(-\infty, \infty)$, то свертка

$$l(x) = \int_{-\infty}^{\infty} g(x - z) f(z)\, dz \tag{3.20}$$

обладает тем же свойством при условии, что интеграл существует. Ядро $l(x - y) = \int_{-\infty}^{\infty} g(x - w) f(w - y)\, dw$ представимо в форме (3.11); применяя основную композиционную формулу (3.12), можно показать, что $l(x - y)$ принадлежит классу $TP$.

Если теперь рассмотреть функцию

$$g_0(x) = \begin{cases} e^{-\beta x}, & x \geqslant 0, \\ 0, & x < 0, \end{cases}$$

то можно сделать вывод, что $g_0(x-y)$ принадлежит классу $TP$ на $(-\infty, \infty)$, так как она представима в виде $g_0(x-y) = e^{-\beta x} e^{\beta y} K(x, y)$, где $K(x, y)$ — ядро, определенное в (3.17).

Применяя операцию свертки, определяемую соотношением (3.20), к $g(x) = g_0(x)$ и $f(x) = g_{m-1}(x)$, найдем, что для $x \geqslant 0$

$$\int_0^\infty g_0(x-z)\, g_{m-1}(z)\, dz = \int_0^x e^{-\beta(x-z)} \frac{z^{m-1}}{(m-1)!} e^{-\beta z} dz =$$
$$= \frac{x^m}{m!} e^{-\beta x} = g_m(x).$$

Простое доказательство по индукции далее показывает, что $g_m(x-y)$ принадлежит классу $TP$ всех порядков.

Функции

$$h(t) = p_m(t) + \sum_{i=0}^{n} c_i (t - \xi_i)_+^m, \qquad (3.21)$$

где $p_m(t)$ — обычный многочлен степени самое большее $m$, называются сплайн-многочленами степени $m$ с узлами $\xi_0, \ldots, \xi_n$. Функции этого типа обладают многими из свойств обычных многочленов и встречаются в исследованиях, связанных с теорией аппроксимации, интерполяции и механических квадратур (см. § 9 гл. XI).

Пример 11. Сплайн-многочлены. Результат, доказанный в примере 10, согласно которому функция $(x-y)_+^m$ для $m \geqslant 0$ принадлежит классу $TP$ на $(-\infty, \infty)$, может быть использован для доказательства того, что система функций

$$t^m, t^{m-1}, \ldots, t, 1, (t-x_1)_+^m, (t-x_2)_+^m, \ldots, (t-x_r)_+^m, \qquad (3.22)$$

где $-1 < x_1 < x_2 < \ldots < x_r < 1$, образует $WT$-систему на интервале $[-1, 1]$.

Из того, что $L(x, y) = (x-y)_+^m$ принадлежит классу $TP$, следует, что если $-1 \leqq t_0 < t_1 < \ldots < t_{m+1} \leqq 1$, то система

$$u_i(s) = (t_i - s)_+^m, \qquad i = 0, 1, \ldots, m + r,$$

является $WT$-системой на $[-1, 1]$. Если выбрать значения $s_0 = s_1 = \ldots$ $\ldots = s_m = 1$ и $s_{m+i} = x_i$, $i = 1, \ldots, r$, то можно показать, как при установлении соотношения (2.6), что

$$0 \leq (-1)^{[m(m+1)]/2} \times U^* \begin{pmatrix} 0, 1, \ldots, m+r \\ s_0, s_1, \ldots, s_{m+r} \end{pmatrix} = \prod_{i=0}^{m-1} (m-i)^{m-i} \times$$

$$\times \begin{vmatrix} (t_0-1)^m & (t_0-1)^{m-1} \ldots 1 & (t_0-x_1)_+^m & \ldots (t_0-x_r)_+^m \\ (t_1-1)^m & (t_1-1)^{m-1} \ldots 1 & (t_1-x_1)_+^m & \ldots (t_1-x_r)_+^m \\ \cdot & \cdot & \cdot & \cdot \\ (t_{m+r}-1)^m & (t_{m+r}-1)^{m-1} \ldots 1 & (t_{m+r}-x_1)_+^m & \ldots (t_{m+r}-x_r)_+^m \end{vmatrix}.$$

Теперь из замены первых $m$ столбцов их соответствующими линейными комбинациями следует, что система функций (3.22) является $WT$-системой.

Чебышевский характер сплайн-кривых в неявном виде используется Шенбергом [1963], де Буром, Гревилем [1964 a, b], Уолшем, Албергом и Нильсоном [1963] и другими авторами в связи с некоторыми интерполяционными задачами (см. также §§ 9 и 10 гл. XI).

Пример 12. Общая конструкция $ET$-систем. Пусть $w_0, w_1, \ldots, w_n$ — положительные функции класса $C^n$ на $[a, b]$. Определим функции

$$\begin{aligned} u_0(x) &= w_0(x), \\ u_1(x) &= w_0(x) \int_a^x w_1(\xi_1)\, d\xi_1, \\ u_2(x) &= w_0(x) \int_a^x w_1(\xi_1) \int_a^{\xi_1} w_2(\xi_2)\, d\xi_2 d\xi_1, \qquad (3.23) \\ &\cdots\cdots\cdots\cdots \\ u_n(x) &= w_0(x) \int_a^x w_1(\xi_1) \int_a^{\xi_1} w_2(\xi_2) \ldots \int_a^{\xi_{n-1}} w_n(\xi_n)\, d\xi_n \ldots d\xi_1. \end{aligned}$$

Рассмотрение систем функций, образованных таким способом, будет проведено в § 5 гл. VIII и в гл. XI. Заметим, что степенные функции $1, t, \ldots, t^n$ могут быть образованы таким способом, если положить $a = 0$, $w_0(x) \equiv 1$ и $w_j(x) \equiv \mathrm{const} \equiv j$, $j = 1, 2, \ldots, n$.

Непосредственными выкладками можно получить формулу для определителя Вронского

$$W(u_0, u_1, \ldots, u_k)(x) = [w_0(x)]^{k+1} [w_1(x)]^k \ldots [w_k(x)]. \qquad (3.24)$$

Отсюда понятно, что определители Вронского $W(u_0, u_1, \ldots, u_k)$, $k = 0, 1, \ldots, n$, положительны на $[a, b]$ при $k = 0, 1, \ldots, n$. Оказывается, что положительность определителей Вронского достаточна, чтобы гарантировать, что система $\{u_i\}_0^n$ является $ET$-системой на $[a, b]$ (см. теорему 1.2 гл. XI).

Если ввести функциональные операторы

$$D_j f = \frac{d}{dx}\frac{f}{w_j}, \qquad j = 0, 1, \ldots, n,$$

то

$$w_{j+1} = D_j D_{j-1} \ldots D_0 u_{j+1}, \qquad j = 0, 1, \ldots, n-1,$$

и может быть показано с помощью формулы (3.24), что функции $w_j$ выражаются как некоторые отношения определителей Вронского (3.24) (см. замечание 1.3 гл. XI).

В главе XI представлено обширное исследование конуса функций $C(u_0, u_1, \ldots, u_n)$, состоящего из таких непрерывных функций $\varphi$, для которых система $\{u_0, u_1, \ldots, u_n, \varphi\}$ является $WT$-системой. Если $\varphi$ из класса $C^n$, то $\varphi \in C(u_0, u_1, \ldots, u_n)$ тогда и только тогда, когда

$$D_n D_{n-1} \ldots D_0 \varphi \geqslant 0. \tag{3.25}$$

Заметим, что $C(1, t)$ совпадает с классом выпуклых функций и (3.25) сводится к условию $\varphi''(x) \geqslant 0$.

Конусы $C(u_0, u_1, \ldots, u_n)$ естественным образом возникают при изучении теории моментов, дифференциальных уравнений и теории неравенств.

Появление $T$-систем в качестве решений линейных дифференциальных уравнений $n$-го порядка, как описано в этом примере, неявно подразумевается у Полиа [1922]. Другие применения этих идей можно найти у Карлина и Новикова [1963], Карлина и Циглера [1966] и Циглера [1966]. Относящиеся к этим вопросам результаты собраны в книге Беккенбаха и Беллмана [1961], гл. 4.

Пример 13. Периодические $T$-системы. В качестве последнего примера предлагаем систему $2m+1$ функций

$$1, \cos t, \sin t, \ldots, \cos mt, \sin mt, 0 \leq t < 2\pi. \tag{3.26}$$

Эти функции — периодические на $[0, 2\pi]$. Система (3.26) является $T$-системой на любом интервале $[a, b]$ длины, меньшей чем $2\pi$. Более общее понятие периодических $T$-систем будет рассмотрено в гл. VI.

## § 4. Эквивалентное определение $T$-систем для многочленов

Определение 4.1. Пусть система функций определена на некотором интервале $[a, b]$. Функцию вида $u = \sum_{i=0}^{n} a_i u_i$, где $a_i$ — вещественные числа, будем называть *u-многочленом* или, для краткости, просто многочленом. Говорят, что многочлен нетривиален, если $\sum_{i=0}^{n} a_i^2 > 0$.

Если $\{u_i\}_0^n$ — $T$-система, то нетривиальный многочлен, имеющий $n$ заданных нулей $t_0 < t_1 < \ldots < t_{n-1}$, конструируется с помощью явной формулы

$$u(t) = U\begin{pmatrix} 0, 1, \ldots, n-1, n \\ t_0, t_1, \ldots, t_{n-1}, t \end{pmatrix}.$$

Многочлен $u$, очевидно, меняет знак, как только $t$ проходит через каждое значение $t_i$, принадлежащее $(a, b)$. В случае, когда $\{u_i\}_0^n$ — — $ET$-система, можно рассмотреть $n$ точек $t_0 \leq t_1 \leq \ldots \leq t_{n-1}$ и образовать многочлен

$$u(t) = U^*\begin{pmatrix} 0, 1, \ldots, n-1, n \\ t_0, t_1, \ldots, t_{n-1}, t \end{pmatrix}.$$

Этот многочлен — нетривиальный, и множеством его нулей с учетом кратности как раз служит набор $\{t_0, t_1, \ldots, t_{n-1}\}$.

Если $\{u_i\}_0^n$ — $T$-система, то функции $u_0, \ldots, u_n$, очевидно, линейно независимы, так что любой многочлен $u = \sum_0^n a_i u_i$ однозначно определяется набором коэффициентов $a_0, \ldots, a_n$. Действительно, необращение в нуль определителя

$$U\begin{pmatrix} 0, \ldots, n \\ t_0, \ldots, t_n \end{pmatrix}$$

для $a \leq t_0 < \ldots < t_n \leq b$ означает, что многочлен определяется своими заданными значениями в любых $n+1$ различных точках. В частности, это означает, что любой нетривиальный многочлен $u$ имеет самое большее $n$ различных нулей.

В этом параграфе будет показано (теорема 4.1), что если система $\{u_i\}$ такова, что любой нетривиальный многочлен $\sum_0^n a_i u_i$ имеет самое большее $n$ различных нулей, то $\{u_i\}_0^n$ — $T$-система.

Две эквивалентные формулировки $T$-систем, первая — в форме требования выполнения некоторых неравенств для определителей, вторая — в форме требования на число нулей $u$-многочленов, имеют свои сравнительные преимущества. Обе формулировки предполагают другие определения более сильными или более слабыми, чем определение $T$-системы. Определение $T$-системы в терминах неравенств определителей обладает тем преимуществом, что оно дает возможность легко образовывать новые $T$-системы из типовых, или изменять $T$-системы и получать $ET$-системы, используя основную композиционную формулу (см. (3.13)).

В случае обычных многочленов существует целый ряд ограничений на число нулей, сформулированных в терминах вариаций в знаках коэффициентов. Простейшее ограничение состоит в том, что число различных нулей не может превышать степень многочлена. Обобщение этого свойства приводит к понятию $T$-системы. Если функции $u_i$ достаточно гладкие, то имеет смысл понятие кратности нулей, и становится оправданным понятие $ET$-системы.

Обычные многочлены также обладают свойством, которое известно как правило знаков Декарта (см. следствие 4.4), и легко установить общую формулировку этого понятия. Различные вспомогательные понятия также оказываются полезными, включая определение $CT$-систем, и т. д. Примеры такого рода встречаются во многих контекстах.

Определение 4.2. Число различных нулей непрерывной функции $f$ на $[a, b]$ обозначается через $Z(f)$.

Покажем теперь, что если $\{u_i\}_0^n$ является системой непрерывных функций на $[a, b]$ и $Z(u) \leq n$ для любого нетривиального многочлена $u$, то определители

$$U\begin{pmatrix}0, \ldots, n \\ t_0, \ldots, t_n\end{pmatrix}$$

сохраняют строго постоянный знак при любых наборах $a \leq t_0 < \ldots$ $\ldots < t_n \leq b$. То есть система $\{u_i\}_0^n$ является $T$-системой (см. теорему 4.1).

Действительно, можно заметить, что определители

$$U\begin{pmatrix}0, \ldots, n \\ t_0, \ldots, t_n\end{pmatrix},$$

где значения $t_i$ из множества $a \leq t_0 < t_1 < \ldots < t_n \leq b$, не могут принимать как отрицательные, так и положительные значения, не обращаясь в нуль для некоторого набора $\{t_i\}_0^n$. Это является непосредственным следствием того факта, что не обращающаяся в нуль непрерывная функция, определенная на связном множестве, необходимо сохраняет постоянным знак.

Резюмируя вышесказанное, сформулируем следующую теорему.

Теорема 4.1. *Если $\{u_i\}_0^n$ является $T$-системой, то $Z(u) \leq n$ для любого нетривиального многочлена $u$. Обратно, если система $\{u_i\}_0^n$ непрерывных функций на $[a, b]$ удовлетворяет условию $Z(u) \leq n$, где $u(t) \not\equiv 0$, то $\{u_i\}_0^n$ является $T$-системой с точностью до знака одной из функций*[1]).

---

[1]) Необходимость только доказывает, что определители нигде не обращаются в 0. Следовательно, согласно определению 1.1, все, что мы можем, собственно, утверждать, это то, что либо $\{u_0, u_1, \ldots, u_{n-1}, u_n\}$, либо $\{u_0, u_1, \ldots, u_{n-1}, -u_n\}$ является $T$-системой.

Следствие 4.1. *Если $\{u_i\}_0^n$ является СТ-системой, то для любого $k = 0, 1, \ldots, n$ справедливо неравенство $Z\left(\sum_{i=0}^{k} a_i u_i\right) \leq k$, если все $a_i$ вещественны и удовлетворяют условию $\sum_0^k a_i^2 > 0$.*

Выводы теоремы 4.1 могут быть усилены путем проведения различия между двумя типами нулей многочлена, *узловыми* и *неузловыми* нулями. Эти понятия соответствуют в случае дифференцируемых функций нулям четной или нечетной кратности соответственно. Заметим, что в теоремах 4.1 и 4.2 предполагается, что функции $u_0, u_1, \ldots, u_n$ просто непрерывны и не обязательно дифференцируемы.

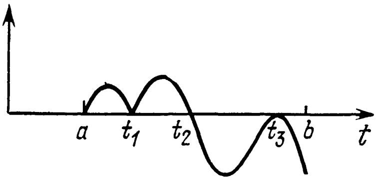


Рис. 1.

Определение 4.3. Для любой непрерывной функции $f$ на $[a, b]$ изолированный нуль $t_0 \in (a, b)$ функции $f$ будем называть *неузловым нулем*, если функция $f$ не изменяет знака в $t_0$. Все другие нули, включая нули в граничных точках $a$ и $b$, называются *узловыми нулями*. Число нулей функции $f$, когда узловые нули считаются один раз, а неузловые нули — дважды, обозначается через $\tilde{Z}(f)$.

На рис. 1 точки $t_1$ и $t_3$ — неузловые нули, в то время как $t_2$ и граничная точка $a$ — узловые нули.

Теорема 4.2. *Если $\{u_i\}_0^n$ — Т-система, то $\tilde{Z}(u) \leq n$ для любого нетривиального многочлена $u$. Обратно, если $\{u_i\}_0^n$ — система непрерывных функций на $[a, b]$ и $\tilde{Z}(u) \leq n$ для любого нетривиального многочлена $u$, то $\{u_i\}_0^n$ — Т-система, с точностью до знака функций* [1]).

Доказательство. Вторая часть теоремы содержится в теореме 4.1. Чтобы доказать первую часть, предположим противное, т. е. что $\tilde{Z}(u) \geqslant n + 1$ для некоторого нетривиального многочлена $u$. Можно также предположить, что $u$ имеет по крайней мере один неузловой нуль, в противном случае результат следует из теоремы 4.1. Обозначим различные нули $u$ через $t_1, \ldots, t_k$ и добавим к этим точкам точку $t_i + \varepsilon$ для каждого неузлового нуля $t_i$, а также точку $t_i - \varepsilon$ для первого неузлового нуля. Для достаточно малого $\varepsilon$ добавленные точки отличаются от $t_1, \ldots, t_k$ и содержатся в $[a, b]$. Следовательно, полученный набор содержит по крайней мере $n + 2$ точки. Расположим эти точки в возрастающем порядке и переобозначим первые $n + 2$ из этих точек как

[1]) См. сноску к теореме 4.1.

$s_0, s_1, \ldots, s_{n+1}$. Значения $u(s_i)$ должны тогда чередоваться в знаке, в том смысле, что $u(s_i) \geqslant 0$ для $i$ четного и $u(s_i) \leqslant 0$ для $i$ нечетного, или наоборот. В любом случае

$$\begin{vmatrix} u(s_0) & u(s_1) & \ldots & u(s_{n+1}) \\ u_0(s_0) & u_0(s_1) & \ldots & u_0(s_{n+1}) \\ \cdot & \cdot & \cdot & \cdot \\ u_n(s_0) & u_n(s_1) & \ldots & u_n(s_{n+1}) \end{vmatrix} = 0,$$

так как первая строка есть линейная комбинация последующих строк. Разлагая определитель по первой строке и используя чебышевское свойство системы $\{u_i\}_0^n$, получим $\sum_{i=0}^{n+1} u(s_i)\, a_i = 0$, где $a_i$ строго чередуются в знаке. Отсюда следует, что $u(s_j) = 0$ для $j = 0, 1, \ldots, n+1$. Таким образом, $Z(u) \geqslant n+1$, и это противоречит утверждению теоремы 4.1. Следовательно, наше первоначальное предположение ложно, и доказательство теоремы завершено.

Теорема 4.2, в частности, утверждает, что если $\{u_i\}_n^0$ есть $T$-система, то, считая дважды каждый нуль многочлена, который соответствует нулю нечетной кратности, общее число нулей никогда не может превысить $n$.

Определение 4.4. Если функция $f$, определенная на $[a, b]$, достаточно гладкая, то число нулей с учетом кратности обозначается через $Z^*(f)$.

Если для каждого многочлена $u$ мы можем рассчитать точную кратность его нулей, то имеет место следующая теорема.

Теорема 4.3. *Если $\{u_i\}_0^n$ — $ET$-система, то $Z^*(u) \leqslant n$ для любого нетривиального многочлена $u$. Обратно, если $u_i \in C^n[a, b]$, $i = 0, 1, \ldots, n$, и $Z^*(u) \leqslant n$ для любого нетривиального многочлена $u$, то $\{u_i\}_0^n$ — $ET$-система с точностью до знака одной из функций*[1].

Доказательство. Первая часть теоремы следует из того факта, что если $Z^*(u) \geqslant n+1$ для некоторого нетривиального многочлена $u$, то

$$U^*\begin{pmatrix} 0, 1, \ldots, n \\ t_0, t_1, \ldots, t_n \end{pmatrix} = 0,$$

где $t_0 \leqslant t_1 \leqslant \ldots \leqslant t_n$ являются любыми $n+1$ нулями. Это противоречит условию, что $\{u_i\}_0^n$ есть $ET$-система.

Обратное утверждение теоремы доказывается следующим образом. Если $Z^*(u) \leqslant n$ для всех $u \not\equiv 0$, то, в частности, $Z(u) \leqslant n$ для

[1] См. сноску к теореме 4.1.

любого такого $u$, поэтому по теореме 4.1 определители

$$U\begin{pmatrix}0, 1, \ldots, n\\ t_0, t_1, \ldots, t_n\end{pmatrix}$$

при $a \leq t_0 < t_1 < \ldots < t_n \leq b$ сохраняют строго постоянный знак. Для определенности положим, что эти определители положительны. Из неравенства (2.6) известно, что

$$U^*\begin{pmatrix}0, 1, \ldots, n\\ t_0, t_1, \ldots, t_n\end{pmatrix} \geqslant 0. \tag{4.1}$$

Теперь, если в (4.1) выполняется равенство для некоторого множества значений $t_i$, то, очевидно, можно построить нетривиальный многочлен $u$ с нулями в точках $t_i$ соответствующих кратностей, поэтому $Z^*(u) \geqslant n+1$ в противоречии с гипотезой. Доказательство обратного утверждения, следовательно, выполнено.

Заметим, не вдаваясь в подробности, что теорема промежуточного типа может быть установлена для случая, когда система $\{u_i\}_0^n$ является $ET$-системой порядка $p$.

В заключение этого параграфа докажем теорему, относящуюся к системам Декарта (определенным ниже), которая является аналогом теорем 4.1—4.3. Эта теорема не понадобится в дальнейшем, однако она имеет значение для указания связей некоторых неравенств с определителями и природой нулей многочленов.

Определение 4.5. Система $\{u_i\}_0^n$ непрерывных функций на $[a, b]$ называется *системой Декарта* или *$D$-системой*, если для каждого $k = 1, 2, \ldots, n+1$ определители

$$U\begin{pmatrix}i_0, i_1, \ldots, i_{k-1}\\ t_0, t_1, \ldots, t_{k-1}\end{pmatrix}.$$

сохраняют строго постоянный знак $\varepsilon_k = \pm 1$ всякий раз, когда $a \leqslant t_0 < t_1 \ldots \ldots < t_{k-1} \leqslant b$ и $i_0, i_1, \ldots, i_{k-1}$ — произвольные целые числа, удовлетворяющие условию $0 \leqslant i_0 < i_1 < \ldots < i_{k-1} \leqslant n$.

Заметим, что это определение сильнее, чем определение $CT$-системы. То есть если принять $(i_0, i_1, \ldots, i_{k-1}) = (0, 1, \ldots, k-1)$, то $D$-система является $CT$-системой при условии $\varepsilon_k \equiv 1$.

Теорема 4.4 (Гантмахер и Крейн [1950]). *Система $\{u_i\}_0^n$ есть $D$-система тогда и только тогда, когда*

$$Z\left(\sum_{i=0}^{n} a_i u_i\right) \leqslant S^-(a_0, a_1, \ldots, a_n)$$

*при условии $\sum_{i=0}^{n} a_i^2 > 0$ ($S^-(a_0, a_1, \ldots, a_n)$ обозначает число перемен знака в последовательности $a_0, a_1, \ldots, a_n$ после того, как нулевые члены отброшены).*

Доказательство. Пусть $\{u_i\}_0^n$ — $D$-система. Предположим, что для некоторого многочлена $u = \sum_{i=0}^{n} a_i u_i$, $\sum_{i=0}^{n} a_i^2 > 0$, имеем $S^-(a_0, a_1, \ldots, a_n) = p$, и $u$

обращается в нуль в $p+1$ различных точках $t_0, \ldots, t_p$. Так как $S^-(a_0, a_1, \ldots \ldots, a_n) = p$, то можно разбить последовательность $a_0, \ldots, a_n$ на $p+1$ групп $\{a_0, \ldots, a_{\nu_1}\}$, $\{a_{\nu_1+1}, \ldots, a_{\nu_2}\}, \ldots, \{a_{\nu_p+1}, \ldots, a_n\}$, где каждое $a_i$ в первой группе является, скажем, неотрицательным, каждое $a_i$ во второй группе является неположительным и т. д., и, следовательно, при этом существует по крайней мере один ненулевой элемент в каждой группе. Далее, так как $\sum_{i=0}^{n} a_i u_i(t_j) = 0$, $j=0, 1, \ldots, p$, имеем

$$\sum_{k=1}^{p+1} (-1)^{k+1} \sum_{i=\nu_{k-1}+1}^{\nu_k} |a_i| u_i(t_j) = 0, \qquad j = 0, 1, \ldots, p, \tag{4.2}$$

где $\nu_0 = -1$ и $\nu_{p+1} = n$ по соглашению. Предположим, что мы рассматриваем (4.2) как однородную систему линейных уравнений, для которой $(1, -1, \ldots \ldots, (-1)^{p+2})$ является решением. Тогда из известного результата линейной алгебры следует, что определитель системы

$$V = \begin{vmatrix} \sum_{i=0}^{\nu_1} |a_i| u_i(t_0) & \ldots & \sum_{i=\nu_p+1}^{n} |a_i| u_i(t_0) \\ \cdot & \cdot\;\cdot\;\cdot & \cdot \\ \sum_{i=0}^{\nu_1} |a_i| u_i(t_p) & \ldots & \sum_{i=\nu_p+1}^{n} |a_i| u_i(t_p) \end{vmatrix} =$$

$$= \sum_{l_0=0}^{\nu_1} \cdots \sum_{l_p=\nu_p+1}^{n} |a_{l_0} \ldots a_{l_p}| \, U\begin{pmatrix} l_0, \ldots, l_p \\ t_0, \ldots, t_p \end{pmatrix}$$

должен обращаться в нуль. Однако это невозможно, так как для некоторого набора индексов $l_0, \ldots, l_p$ произведение $|a_{l_0} \ldots a_{l_p}|$ положительно, и все определители

$$U\begin{pmatrix} l_0, \ldots, l_p \\ t_0, \ldots, t_p \end{pmatrix}$$

имеют строго постоянный знак для всех наборов $0 \leqslant l_0 < l_1 < \ldots l_p \leqslant n$. Это противоречие означает, что если $\{u_i\}_0^n$ является $D$-системой, то $Z\left(\sum_{i=0}^{n} a_i u_i\right) \leqslant S^-(a_0, \ldots, a_n)$ при условии $\sum_{i=0}^{n} a_i^2 > 0$.

Чтобы доказать обратное утверждение, прежде всего рассмотрим функции $u_{i_0}, u_{i_1}, \ldots, u_{i_{k-1}}$. По теореме 4.1 и из того, что $S^-(a_{i_0}, a_{i_1}, \ldots, a_{i_{k-1}}) \leqslant \leqslant k-1$, заключаем, что определители

$$U\begin{pmatrix} i_0, \ldots, i_{k-1} \\ t_0, \ldots, t_{k-1} \end{pmatrix}$$

имеют строго постоянный знак $\varepsilon(i_0, i_1, \ldots, i_{k-1})$, не зависящий от выбора $t_0$, $t_1, \ldots, t_{k-1}$, где $a \leqslant t_0 < t_1 < \ldots < t_{k-1} \leqslant b$. Теперь, если $a \leqslant t_0 < \ldots \ldots < t_{k-1} \leqslant b$, то многочлен

$$U\begin{pmatrix} i_0, \ldots, i_{k-1}, i_k \\ t_0, \ldots, t_{k-1}, t \end{pmatrix} = \sum_{j=0}^{k} (-1)^{k-j} U\begin{pmatrix} i_0, \ldots, i_{j-1}, i_{j+1}, \ldots, i_k \\ t_0 \quad \ldots, \quad t_{k-1} \end{pmatrix} u_{i_j}(t)$$

имеет $k$ различных нулей $t_0 < t_1 < \ldots < t_{k-1}$.

Следовательно, так как $Z\left(\sum_{j=0}^{k} a_j u_{i_j}\right) \leqslant S^-(a_0, \ldots, a_k)$, то $k+1$ коэффициентов вышеприведенного многочлена должны иметь $k$ строгих смен знака. Следовательно, знак определителя

$$U\begin{pmatrix} i_0, \ldots, i_{j-1}, i_{j+1}, \ldots, i_k \\ t_0, \quad \ldots, \quad t_{k-1} \end{pmatrix}$$

не зависит от $j = 0, 1, \ldots, k$. Тогда можно утверждать, что $\varepsilon(i_0, i_1, \ldots, i_{k-1}) = \varepsilon(1, \ldots, k) = \varepsilon_k$ зависит только от $k$. Этим и завершается доказательство.

Так как $\{t^{\alpha_i}\}_0^n$ является $D$-системой (см. пример 1), то получаем следующее

Следствие 4.4. *Число различных нулей на открытом интервале* $(0, \infty)$ *любого многочлена* $\sum_{i=0}^{n} a_i t^{\alpha_i}$, *где* $\sum_{i=0}^{n} a_i^2 > 0$ *и* $\alpha_0, \alpha_1, \ldots, \alpha_n$ — *возрастающая последовательность вещественных чисел, не превосходит* $S^-(a_0, \ldots, a_n)$.

Матричный вариант теории, ограничивающей число смен знака и фигурирующей в утверждении теоремы 4.4, впервые был предложен Шенбергом. Теорема 4.4 в том виде, как она сформулирована, принадлежит Гантмахеру и Крейну [1950]. Обобщения понятия систем Декарта до тщательно разработанной теории знакорегулярных ядер интенсивно развиваются в работах Гантмахера и Крейна [1950] и Карлина [1967].

## § 5. Многочлены с заданными нулями

Перейдем теперь к задаче построения неотрицательных многочленов, имеющих заданный набор нулей. В случае обычных многочленов построения тривиальны благодаря тому, что для функций $1, t, t^2, \ldots, t^n$ процесс факторизации нулей возможен. Эта особенность отсутствует в случае общих $T$-систем. Кроме того, дальнейшее осложнение состоит в том, что базисные функции не обязательно предполагаются дифференцируемыми.

Теоремы 5.1 и 5.2, установленные ниже, являются решающими для доказательств многих теорем в последующих главах. Эти результаты могли бы быть также установлены путем сглаживания системы $\{u_i\}_0^n$ с целью получения $ET$-системы и последующим переходом к пределу, как в § 3 (см. выражения (3.15) и (3.16)).

Пусть $T = \{t_1, \ldots, t_k\}$ — возрастающий набор различных точек на $[a, b]$. Если требуется построить неотрицательный многочлен, который обращается в нуль в каждой точке из $T$, то каждая внутренняя точка $t_i$ необходимо будет неузловым нулем, в то время как, если граничные точки $a$ или $b$ принадлежат $T$, эти точки будут узловыми нулями по определению. Следовательно, каждой точке $t_i$, принадлежащей $T$, приписывается вес $\omega(t_i)$, определенный как

$$\omega(t_i) = \begin{cases} 2, & t_i \in (a, b), \\ 1, & t_i = a \text{ или } b. \end{cases}$$

Если $\{u_i\}_0^n$ — $T$-система, то для того, чтобы существовал неотрицательный многочлен, который обращается в нуль на $T$, необходимо в соответствии с теоремой 4.2, чтобы $\sum_{i=1}^{k} \omega(t_i) \leq n$.

Теорема 5.1 (Крейн [1951]).

a) *Если* $\{u_i\}_0^n$ — *$T$-система и* $\sum_{i=1}^{n} \omega(t_i) \leq n$, *то существует нетривиальный неотрицательный многочлен* $u$, *обращающийся в нуль только в точках из* $T = \{t_1, \ldots, t_k\}$. *Единственное исключение состоит в том, что если* $n = 2m$ *и как раз одна из конечных точек* $a$ *или* $b$ *принадлежит* $T$, *то* $u(t)$ *может, кроме того, обращаться в нуль в другой конечной точке.*

b) *Если какое-либо одно из последующих условий выполняется, то всегда может быть построен многочлен, который обращается в нуль только при* $t_1, \ldots, t_k$:

1) $\{u_i\}_0^{n-1}$ *(т. е. набор функций* $u_0, u_1, u_2, \ldots, u_{n-1}$, *за исключением* $u_n$) *является $T$-системой.*

2) $\{u_i\}_0^n$ — *$T$-система на интервале* $[a', b']$, *содержащем* $[a, b]$, *где* $a' < a < b < b'$.

3) $\{u_i\}_0^n$ — *$ET$-система порядка 2 на* $[a, b]$.

Доказательство. Будем различать случаи четного и нечетного $n$.

I ($n = 2m + 1$). Рассмотрим первый случай, когда $a < t_1 < t_2 < \ldots$ $\ldots < t_k < b$. Присоединим к последовательности $\{t_i\}_1^k$ точки $a, t'_1, t'_2, \ldots, t'_{m-k}$, которые удовлетворяют условию $t_k < t'_1 < t'_2 < \ldots$ $\ldots < t'_{m-k} < b$. Теперь рассмотрим ряд $\{s_i(\varepsilon)\}_1^{2m+1}$, состоящий из точек

$$a_1, t_1, t_1 + \varepsilon, t_2, t_2 + \varepsilon, \ldots, t_k, t_k + \varepsilon, t'_1, t'_1 + \varepsilon, \ldots, t'_{m-k}, t'_{m-k} + \varepsilon \tag{5.1}$$

с $\varepsilon > 0$, выбранным достаточно малым, так, чтобы последовательность (5.1) была расположена в возрастающем порядке. Образуем многочлен

$$u_\varepsilon(t) = U\begin{pmatrix} 0, 1, \ldots, & 2m, & 2m+1 \\ s_1, s_2, \ldots, & s_{2m+1}, & t \end{pmatrix}. \tag{5.2}$$

Так как $\{u_i\}_0^n$ образует $T$-систему, то $u_\varepsilon(t)$ обращается в нуль точно на множестве $\{s_i\}$ и каждый нуль является узловым. Кроме того, для $t < s_{2m+1}$, очевидно, имеем $u_\varepsilon(t) > 0$. Из этого следует, в частности, что

$$\begin{aligned} &u_\varepsilon(t) > 0, \quad s_{2i-1} < t < s_{2i} \qquad (i = 1, 2, \ldots m), \\ &u_\varepsilon(t) > 0, \quad s_{2m+1} < t \leq b. \end{aligned} \tag{5.3}$$

Раскрывая определитель (5.2) по элементам последнего столбца, получим

$$u_\varepsilon(t) = \sum_{i=0}^{n} a_i(\varepsilon)\, u_i(t). \tag{5.4}$$

Многочлен $u_\varepsilon(t)$ может предполагаться нормированным (это достигается путем соответствующего умножения на положительную константу) в том смысле, что $\sum_{i=0}^{n} a_i^2(\varepsilon) = 1$. Допуская, что $\varepsilon \downarrow 0$, выберем предельный многочлен $\tilde{u}(t) = \sum_{i=0}^{n} a_i u_i(t) \quad \left(\sum a_i^2 = 1\right)$, который неотрицателен в соответствии с (5.3) и обращается в нуль на множестве

$$\{a, t_1, t_2, \ldots, t_k, t_1', t_2', \ldots, t_{m-k}'\}. \tag{5.5}$$

Так как точки из множества $\{t_i\}_1^k \cup \{t_i'\}_1^{m-k}$ определяют неузловые нули, то из теоремы 4.2 следует, что последовательность (5.5) заключает в себе полное множество нулей $\tilde{u}(t)$.

Подобным же образом можно построить неотрицательный многочлен $u^*(t)$, множество нулей которого точно совпадает с множеством $\{t_1, t_2, \ldots, t_k, t_1'', t_2'', \ldots t_{m-k}'', b\}$, где пересечение $\{t_j''\} \cap \{t_j'\}$ пусто по условию. Многочлен $u(t) = \tilde{u}(t) + u^*(t)$, очевидно, представляет собой неотрицательный многочлен, обращающийся в нуль точно на заданном множестве $\{t_i\}_1^k$.

Построение многочлена $u(t)$ требуемого типа для других возможных множеств $\{t_i\}$ осуществляется аналогично. Возможно, только случай $a = t_1 < t_2 < \ldots < t_{k-1} < t_k = b$ требует некоторого пояснения. Как и раньше, увеличиваем последовательность путем добавления $t_1', t_2', \ldots, t_{m-k+1}'$, удовлетворяющих условию $t_{k-1} < t_j' < t_k \quad (1 \leq j \leq m-k+1)$.

Рассмотрим теперь последовательность $\{s_i\}_1^{2m+1} = \{a = t_1, t_1 + \varepsilon, t_2, t_2+\varepsilon, \ldots, t_{k-1}, t_{k-1}+\varepsilon, t_1', t_1'+\varepsilon, \ldots, t_{m-k+1}', t_{m-k+1}' + \varepsilon, b\}$. Снова $\varepsilon > 0$ выбирается достаточно малым, и образуется многочлен (5.2). Процесс выбора, аналогичный проделанному выше, определяет нетривиальный неотрицательный многочлен $\tilde{u}(t)$, обращающийся в нуль в точках множества $\{a = t_1, t_2, \ldots, t_k = b, t'_1, t'_2, \ldots, t'_{m-k+1}\}$. Можно утверждать, что $\tilde{u}(t)$ не может обращаться в нуль в других точках, иначе нарушается заключение теоремы 4.2. Далее строим $u^*(t)$, заменив дополняющее множество $\{t'_j\}$ на $\{t''_j\}$. Тогда $u(t) = \tilde{u}(t) + u^*(t)$, очевидно, имеет желаемые свойства.

II ($n = 2m$). Доказательство выполняется способом, сходным с вышеприведенным доказательством п. I. Стоит предупредить о несколько более слабом заключении в случае, когда только одна из граничных точек $a$ или $b$ содержится в $T$. Не было решено, можно ли этот особый случай исключить из теоремы.

Результат, как и в п. I, достигается, если какое-либо одно из дополнительных условий 1), 2) или 3) выполняется. Формальные детали не представляют трудностей. Например, если выполняется условие 1), можно случай, когда $\sum_{i=0}^{n} \omega(t_i)$ нечетна, сразу же свести к анализу п. I при ограничении нашего рассмотрения системой $\{u_i\}_0^{n-1}$. Подобные аргументы пригодны и в том случае, когда либо условие 2) либо условие 3) выполняется.

Замечание 5.1. Почти во всех случаях, представляющих интерес, система $\{u_i\}_0^n$ является $CT$-системой, откуда, в частности, следует, что условие 1) пункта b) выполняется. Утверждения для $n$ нечетного или четного тогда идентичны.

Анализ, подобный предшествующему, приводит к следующей теореме.

Теорема 5.2. *Пусть $\{u_i\}_0^n$ — $T$-система. Рассмотрим множество $T = \{t_1, t_2, \ldots, t_k\}$, приписывая каждой точке вес $\omega(t_i) = 1$ или 2, подчиненный ограничению, что если $t_i$ является одной из граничных точек $a$ или $b$, то всегда $\omega(t_i) = 1$; и предположим, что $\sum_{i=1}^{k} \omega(t_i) \leq n$. Тогда существует многочлен $u(t)$ такой, что $u(t) \neq 0$ для $t \in (a, b) \setminus T$, и $t_i$ является узловым или неузловым нулем в зависимости от того, $\omega(t_i) = 1$ или $\omega(t_i) = 2$ соответственно.*

*Кроме того, при любом из условий* b) *теоремы* 5.1 *вышеприведенный многочлен обращается в нуль точно на множестве $T$.*

## § 6. Задача интерполяции

В этом параграфе обсуждаются некоторые задачи интерполяции. Решение этих задач будет дано в одном из доказательств теоремы Маркова — Крейна в гл. III, так что читатель может отложить чтение этого параграфа до гл. III. Так как используемые здесь построения подобны построениям в предыдущем параграфе, кажется целесообразным, чтобы эти два параграфа были представлены вместе.

Будем предполагать, что функции $u_0, u_1, \ldots, u_n$ и $\Omega$ удовлетворяют следующему условию.

Условие 6.1. Система $\{u_i\}_0^n$ — $CT$-система, $\Omega(t)>0$ для $t\in[a,b]$ и $\{u_0(t),\ldots,u_k(t),\Omega(t)\}$ — $T$-система для каждого $k=0,1,2,\ldots,n$.

Введем функции

$$w_i(t)=\frac{u_i(t)}{\Omega(t)},\qquad i=0,1,\ldots,n.$$

Пусть $\{u_0,u_1,\ldots,u_k,\Omega\}$ — $T$-система и $\Omega(t)>0$, тогда для $a\le t_1<t_2<\ldots<t_{k+2}\le b$ справедливо неравенство

$$\begin{vmatrix} w_0(t_1) & w_0(t_2) & \ldots & w_0(t_{k+2})\\ w_1(t_1) & w_1(t_2) & \ldots & w_1(t_{k+2})\\ \cdot & \cdot & \cdot & \cdot\\ w_k(t_1) & w_k(t_2) & \ldots & w_k(t_{k+2})\\ 1 & 1 & \ldots & 1\end{vmatrix}>0, \tag{6.1}$$

$$k=0,1,\ldots,n.$$

В дальнейшем будет необходима следующая лемма, представляющая и некоторый самостоятельный интерес.

Лемма 6.1. *Пусть условие* (6.1) *выполнено. Тогда*

$$(-1)^r\cdot\begin{vmatrix} w_0(t_1) & \ldots & w_0(t_{k-r+2}) & w_0(t_{k-r+3}) & \ldots & w_0(t_{k+2})\\ w_1(t_1) & \ldots & w_1(t_{k-r+2}) & w_1(t_{k-r+3}) & \ldots & w_1(t_{k+2})\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ w_k(t_1) & \ldots & w_k(t_{k-r+2}) & w_k(_{k-r+3}) & \ldots & w_k(t_{k+2})\\ 1 & & 1 & 0 & \ldots & 0\end{vmatrix}=$$

$$=(-1)^rM_{r,k}>0,\qquad 0\le r\le k+1\le n+1, \tag{6.2}$$

*где $r$ обозначает число нулей в последней строке. Заметим, что нулевые компоненты расположены друг за другом вслед за последовательным рядом из единиц.*

Доказательство. Доказательство проводится индукцией по $r$. Заметим, что (6.1) выражает справедливость (6.2) для всех $k$, когда $r=0$. Кроме того, имеем

$$(-1)^{k+1}M_{k+1,k}=(-1)^{2k+2}W\begin{pmatrix} w_0, w_1,\ldots, w_k\\ t_2, t_3,\ldots, t_{k+2}\end{pmatrix}>0.$$

Предположим теперь, что неравенство (6.2) выполняется для $r-1$ и всех $k$ и для $r$, когда $k=r-1,\ldots,l-1$, и рассмотрим (6.2) с $r$ нулевыми компонентами и $k=l$. Применяя тождество Сильвестра к $M_{r,l}(t_1,\ldots,t_{l+2})$, где вспомогательный блок образуется исключением последних двух строк первого и последнего столбцов, получаем

$$W\begin{pmatrix} w_0, \dots, w_{l-1} \\ t_2, \dots, t_{l+1} \end{pmatrix} M_{r,l}(t_1, \dots, t_{l+2}) =$$

$$= \begin{vmatrix} W\begin{pmatrix} w_0 \dots w_l \\ t_1 \dots t_{l+1} \end{pmatrix} & W\begin{pmatrix} w_0 \dots w_l \\ t_2 \dots t_{l+2} \end{pmatrix} \\ M_{r-1,l-1}(t_1, \dots, t_{l+1}) & M_{r,l-1}(t_2, \dots, t_{l+2}) \end{vmatrix}^{1)}. \qquad (6.3)$$

Умножая это соотношение на (— 1) и используя предположение индукции, видим, что правая часть, очевидно, положительна, так как каждый член в разложении положителен.

Очевидное преобразование определителя в (6.2) приводит к следствию.

Следствие 6.1. *При условиях леммы* 6.1

$$W_\nu(t_1, \dots, t_{k+2}) =$$

$$= \begin{vmatrix} w_0(t_1) - w_0(t_2), \dots, w_0(t_{\nu-1}) - w_0(t_\nu), w_0(t_{\nu+1}), w_0(t_{\nu+2}), \dots, w_0(t_{k+2}) \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ w_k(t_1) - w_k(t_2), \dots, w_k(t_{\nu-1}) - w_k(t_\nu), w_k(t_{\nu+1}), w_k(t_{\nu+2}), \dots, w_k(t_{k+2}) \end{vmatrix} > 0 \qquad (6.4)$$

$$(a \leq t_1 < t_2 < \dots < t_{k+2} \leq b, \quad 1 \leq \nu \leq k+2).$$

Перейдем теперь к обсуждению важной интерполяционной задачи, решение которой будет существенным инструментом в выводе теоремы Маркова — Крейна (теорема 2.1 гл. III).

Предположим, что условие 6.1 имеет место. Пусть $n+1$ точек, заданных на $[a, b]$, удовлетворяют условию

$$a \leq t_1 < t_2 < \dots < t_h < \tau_1 < \tau_2 < \dots < \tau_k \leq b \qquad (6.5)$$
$$(h + k = n + 1; \quad 1 \leq h).$$

---

[1]) Частный случай тождества Сильвестра, который здесь используется, состоит в том, что если $A(i, j)$ определено для $i, j = 1, 2, \dots, n$, то для $1 \leqslant i < j \leqslant n$ и $1 \leqslant k < l \leqslant n$ имеем

$$A\begin{pmatrix} 1, 2, \dots, n \\ 1, 2, \dots, n \end{pmatrix} A\begin{pmatrix} 1, \dots, i-1, i+1, \dots, j-1, j+1, \dots, n \\ 1, \dots, k-1, k+1, \dots, l-1, l+1, \dots, n \end{pmatrix} =$$

$$= \begin{vmatrix} A\begin{pmatrix} 1, \dots, j-1, j+1, \dots, n \\ 1, \dots, l-1, l+1, \dots, n \end{pmatrix} & A\begin{pmatrix} 1, \dots, j-1, j+1, \dots, n \\ 1, \dots, k-1, k+1, \dots, n \end{pmatrix} \\ A\begin{pmatrix} 1, \dots, i-1, i+1, \dots, n \\ 1, \dots, l-1, l+1, \dots, n \end{pmatrix} & A\begin{pmatrix} 1, \dots, i-1, i+1, \dots, n \\ 1, \dots, k-1, k+1, \dots, n \end{pmatrix} \end{vmatrix}.$$

Второй определитель в левой части равенства упоминается как вспомогательный блок.

Пусть $\{\alpha_p\}$ определяются как решения системы уравнений

$$\sum_{p=0}^{n} \alpha_p u_p(t_i) = \Omega(t_i), \qquad i = 1, \ldots, h,$$
$$\sum_{p=0}^{n} \alpha_p u_p(\tau_j) = 0, \qquad j = 1, \ldots, k. \tag{6.6}$$

Матрица системы (6.6) не вырождена (так как $\{u_j\}_0^n$ — $T$-система), и, следовательно, вектор коэффициентов $\{\alpha_p\}$ определяется однозначно.

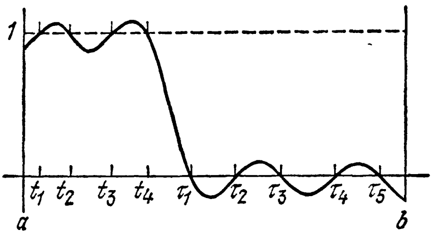


Рис 2.

Деля на $\Omega$, можем переписать (6.6) в виде

$$\sum_{p=0}^{n} \alpha_p w_p(t_i) = 1, \qquad i = 1, \ldots, h,$$
$$\sum_{p=0}^{n} \alpha_p w_p(\tau_j) = 0, \qquad j = 1, \ldots, k. \tag{6.7}$$

Записывая решение для последнего коэффициента $\alpha_n$ и обращаясь к лемме 6.1, заключаем, что

$$(-1)^k \alpha_n > 0. \tag{6.8}$$

Пусть $W(t) = \sum\limits_{p=0}^{n} \alpha_p w_p(t$ . Общий вид многочлена $w(t)$ в случае $h = 4$ и $k = 5$ показан на рис. 2.

Для удобства дальнейшего изложения разделим рассуждения на ряд этапов.

1) Мы утверждаем, что $W(t) = 1$ на сегменте $[a, \tau_1]$ только в точках $\{t_j\}_1^h$. Действительно, предположим противное, т. е. что существует точка $t_0 \neq t_i$ $(i = 1,\ldots, h)$, $a \leq t_0 < \tau_1$, для которой $W(t_0) = 1$. (Ясно, что $W(\tau_1) = 0$.) Предполагаем, что $k \geqslant 1$, так как утверждение для $k=0$ очевидно. Пусть $\{s_l\}$ представляет собой набор точек $\{t_1, \ldots, t_h, t_0, \tau_1, \ldots, \tau_{k-1}\}$, переставленных в возрастающем порядке. Тогда

$$\sum_{p=0}^{n} \alpha_p w_p(s_l) = 1, \qquad l = 1, \ldots, h+1,$$
$$\sum_{p=0}^{n} \alpha_p w_p(s_l) = 0, \qquad l = h+2, \ldots, n+1. \tag{6.9}$$

Находя $\alpha_n$ из (6.9) и снова обращаясь к лемме 6.1, заключаем, что

$$(-1)^{k-1} \alpha_n > 0. \tag{6.10}$$

Очевидно, условия (6.8) и (6.10) несовместимы, что и доказывает утверждение 1).

Аналогичным образом можно доказать

2) $W(t)$ обращается в нуль на сегменте $[t_h, b]$ только в точках $\{\tau_1, \tau_2, \ldots, \tau_k\}$.

Следующее свойство является фундаментальным.

3) Точки $\{t_i\}_{i=1}^h$ — узловые нули $W(t)-1$ и $\{\tau_i\}_{i=1}^k$ — узловые нули $W(t)-1$ и $\{\tau_i\}_{i=1}^k$ — узловые нули $W(t)$.

Доказательство. Рассмотрим множество точек $t_{i_1}, \ldots, t_{i_r}$, в которых $W(t)-1$ изменяет знак, и множество точек $\tau_{i_1}, \ldots, \tau_{i_s}$, в которых изменяет знак $W(t)$. Не считая точек $t_1$ и $\tau_k$, набор $\{t_{i_1}, \ldots, t_{i_r}, 2^{-1}(t_h+\tau_1), \tau_{i_1}, \ldots, \tau_{i_s}\}$ содержит не более $n$ точек. Теперь применим теорему 5.2 и определим $u(t)$ как многочлен, который имеет узловой нуль в каждой из этих точек, изменяет знак с положительного на отрицательный при $t = 2^{-1}(t_h+\tau_1)$ и не обращается в нуль в других точках открытого интервала $(a, b)$. Из этого следует, что $W_\varepsilon(t) = W(t) + \varepsilon(u(t)/\Omega(t))$ при достаточно малом $\varepsilon > 0$ является таким, что $W_\varepsilon(t) - 1$ обращается в нуль по крайней мере однажды в окрестности каждой точки $t_i$ и по крайней мере дважды вблизи $t_i$, если $t_i$ — неузловой нуль $W(t)-1$. Аналогично функция $W_\varepsilon(t)$ обращается в нуль по крайней мере однажды вблизи каждой точки $\tau_i$ и дважды, если $\tau_i$ — неузловой нуль $W(t)$.

Если $W(t)-1$ имеет неузловой нуль в одной из точек $t_i$, то, используя (6.9), можно прийти к противоречию, как в вышеприведенном утверждении 1). Это доказывает, что $\tau_i$, $i = 1, \ldots, k$, — узловые нули $W(t)$.

Предшествующий анализ доказывает следующую теорему.

Теорема 6.1 (Крейн [1951]). *Пусть условие 6.1 выполнено и $n+1$ точек заданы, как в (6.5). Тогда существует единственный многочлен $u(t) = \sum_{j=0}^{n} a_j u_j(t)$ такой, что $u(t) - \Omega(t)$ имеет точно $h$ нулей $\{t_i\}_{i=1}^h$ на $[a, \tau_1]$, каждый из которых является узловым, и $u(t)$ имеет точно $k$ нулей $\{\tau_j\}_{j=1}^k$ на $[t_h, b]$, которые также узловые. (Если $t_1 = a$ или $\tau_k = b$, то эти точки будут узловыми нулями по определению.)*

Пусть теперь $T\{\xi_1, \ldots, \xi_k\}$ — множество различных точек из $[a, b]$, расположенных в возрастающем порядке, и припишем вес $\omega(\xi_i)$ каждой из точек, как в теореме 5.1, т.е. положим

$$\omega(\xi_i) = \begin{cases} 2, & \xi_i \in (a, b), \\ 1, & \xi_i = a \text{ или } b. \end{cases}$$

Тогда, используя результаты теоремы 6.1, можно доказать следующую теорему.

Теорема 6.2 (Крейн [1951]). *Пусть условие (6.1) выполняется и $T = \{\xi_1, \ldots, \xi_k\}$ — множество точек, удовлетворяющих условию $\sum_{i=1}^{k} \omega(\xi_i) \leq n + 2$; пусть еще $\xi_{i_0} \in T \cap (a, b)$.*

a) *Если $\xi_{i_0} > \xi_1$, то существует многочлен $u^*(t)$, обладающий следующими свойствами:*

1) $\Omega(t) \geqslant u^*(t)$, $t \in [a, \xi_{i_0}]$, *причем равенство достигается в точках $\xi_1, \xi_2, \ldots, \xi_{i_0-1}$;* 2) $u^*(t) \leq 0$, $t \in [\xi_{i_0}, b]$, *причем равенство достигается в точках $\xi_{i_0}, \ldots, \xi_k$.*

b) *Далее, для любой точки $\xi_{i_0}$, принадлежащей $T$, существует многочлен $u^{**}(t)$ со свойствами:*

1′) $u^{**}(t) \geqslant \Omega(t), t \in [a, \xi_{i_0}]$, *причем равенство достигается в точках $\xi_1, \xi_2, \ldots, \xi_{i_0}$;* 2′) $u^{**}(t) \geqslant 0$, $t \in [\xi_{i_0}, b]$, *причем равенство достигается в точках $\xi_{i_0+1}, \ldots, \xi_k$.*

Доказательство. a) Эти результаты получаются предельным переходом, аналогичным доказательству теоремы 5.1. За счет добавления дополнительных точек, если это необходимо, можно считать, что $\sum_{i=1}^{k} \omega(\xi_i) \geqslant n + 1$. Мы дублируем точки из $T$, за исключением $\xi_{i_0}$. Таким образом, рассматриваются два множества точек

$$\xi_1, \xi_1 + \varepsilon, \xi_2, \xi_2 + \varepsilon, \ldots, \xi_{i_0-1}, \xi_{i_0-1} + \varepsilon \tag{6.11}$$

и

$$\xi_{i_0+1}, \xi_{i_0+1} + \varepsilon, \ldots, \xi_k, \xi_k + \varepsilon \tag{6.12}$$

с условием, что если $\xi_1 = a$, то $\xi_1 + \varepsilon$ исключается, и аналогично, если $\xi_k = b$. Общее число точек, включая точку $\xi_{i_0}$, равно либо $n$, либо $n + 1$. Если это число равно $n$, то добавляем точку $a + \varepsilon$, так что в обоих случаях имеем точно $n + 1$ точек. По теореме 6.1 существует многочлен $u^*_\varepsilon(t)$, удовлетворяющий условию

$$\begin{aligned} u^*_\varepsilon(\xi') &= \Omega(\xi'), \\ u^*_\varepsilon(\xi_{i_0}) &= 0, \\ u^*_\varepsilon(\xi'') &= 0, \end{aligned} \tag{6.13}$$

где $\xi'$ принадлежит множеству (6.11), а $\xi''$ — множеству (6.12). Легко доказывается, что коэффициенты $\{\alpha_p(\varepsilon)\}$ из соотношения

$$u_\varepsilon(t) = \sum_{p=0}^{n} \alpha_p(\varepsilon)\, u_p(t)$$

равномерно ограничены. Это является следствием того факта, что $u_\varepsilon(t)$ по теореме 6.1 заключен на интервале $(t_h + \varepsilon, \tau_1)$ между 0 и $\Omega(t)$. Следовательно, мы можем ограничить многочлены $u_\varepsilon(t)$ в $n+1$ заданных точках. Теперь для решений с коэффициентами $\{\alpha_i(\varepsilon)\}_0^n$ следует, что они равномерно ограничены по $\varepsilon$. Переходя к предельному многочлену $u^*(t)$ (при $\varepsilon \to 0$) и вспоминая осцилляционные свойства $u_\varepsilon(t)$, которые описаны в теореме 6.1, заключаем, что $u^*(t)$ обладает свойствами, описанными в п. а) теоремы.

Доказательство второй половины теоремы проводится аналогично, единственное изменение в (6.13) состоит в условии $u_\varepsilon^{**}(\xi_{i_0}) = \Omega(\xi_{i_0})$.

Следствие 6.2. *Если* $\sum_{i=1}^{k} \omega(\xi_i) \geqslant n+1$, *то* $u^*, \Omega - u^*, u^{**}, \Omega - u^{**}$ *могут обращаться в нуль на открытом интервале* $(a, b)$ *только в точках* $\xi_i$, *которые задаются в формулировке теоремы.*

Доказательство. Проверим это утверждение только для многочлена $u^*$. Если $\Omega - u^*$ или $u^*$ обращаются в нуль на открытом интервале $(a, b)$ в некоторой точке, отличной от точек $\xi_1, \ldots, \xi_{i_0-1}$ и $\xi_{i_0}, \ldots, \xi_{i_p}$ соответственно, то многочлен $u = u^* + v(\varepsilon) \geqslant \geqslant v(t) > 0$, $t \in [a, b]$ таков, что $u$ и $\Omega - u$ имеют суммарное число нулей, равное $n+2$. Однако это невозможно, так как, если мы определяем многочлен $u$ из (6.6), как имеющий $n+1$ из этих нулей, то справедливы утверждения 1) и 2) на стр. 44, 45, и дополнительный нуль у многочленов $u$ или $\Omega - u$ не может существовать.

В заключение подчеркнем, что интерполяционные результаты (теоремы 6.1 и 6.2) принадлежат Крейну [1951], но доказательства, приводимые здесь, отличаются от доказательств Крейна. Лемма 6.1 представляет и определенный самостоятельный интерес.

Глава II

# МОМЕНТНЫЕ ПРОСТРАНСТВА, ИНДУЦИРОВАННЫЕ $T$-СИСТЕМАМИ, И ДВОЙСТВЕННЫЕ К НИМ

В этой главе мы исследуем моментные пространства $\widetilde{\mathcal{M}}_{n+1}$ (см. уравнение (1.1) ниже), индуцированные чебышевской системой $\{u_i\}_0^n$. Это пространство можно охарактеризовать как наименьший выпуклый конус, содержащий кривую, порожденную точками $(u_0(t), u_1(t), \ldots, u_n(t))$, когда $t$ изменяется внутри интервала $[a, b]$.

По теореме Каратеодори каждая точка в $\widetilde{\mathcal{M}}_{n+1}$ может быть представлена как линейная комбинация с положительными коэффициентами по крайней мере из $n+2$ точек этой кривой. В случае $T$-систем мы покажем, что каждая точка внутри $\widetilde{\mathcal{M}}_{n+1}$ имеет одно параметрическое семейство выпуклых представлений, которые включают приблизительно $(n+2)/2$ точек кривой.

Будут также подробно изучены перемежаемость, свойства монотонности и другие характерные черты этих представлений.

Мы рассмотрим также двойственное пространство $\widetilde{\mathcal{P}}_{n+1}$, состоящее из всех неотрицательных многочленов. При изучении $\widetilde{\mathcal{P}}_{n+1}$ будут доказаны (§ 10) некоторые теоремы, которые обобщают классический результат, состоящий в том, что обычный неотрицательный многочлен может быть представлен как сумма квадратов двух многочленов.

Геометрический подход к теории моментов восходит к Каратеодори [1907]. Он в особенности исследовал природу области коэффициентов для множества аналитических функций, определенных в единичном круге с положительной вещественной частью. Приемы, связанные с выпуклыми множествами и выпуклыми конусами, играли существенную роль в этих исследованиях. Важную роль играли также хорошо известные для случая $n$-мерного евклидова пространства теоремы относительно числа слагаемых в представлении точки в выпуклой оболочке ее через множество крайних точек.

Геометрическое изучение обобщенных моментных пространств, индуцированных $T$-системами, было систематически развито в классической работе Крейна [1951]. Независимо и одновременно Карлин и Шепли [1953] разработали геометрию приведенных моментных пространств, т. е. моментных пространств, порожденных специальной $T$-системой $\{1, t, t^2, \ldots, t^n\}$. Точка зрения и методы Карлина и Шепли, с одной стороны, и Крейна, с другой, сходны и являются первоначальным источником для большей части материала этой главы.

Некоторые обобщения упомянутого геометрического подхода на случай общих систем (не обязательно $T$-систем) и на случай бесконечномерных пространств были даны Рогозинским [1958]. Другой исключительно плодотворный вариант этой теории с большим числом приложений к классическому анализу можно найти у Розенблюма [1951].

Классическое в аналитическом духе изложение обычных моментных пространств, неравенств для моментов, доказательства существования решения проблемы моментов и т. д. дано у Шохата и Тамаркина [1943] (см. также приведенные там ссылки) и у Ахиезера и Крейна [1962]. Эти две последние работы очень полно изучают задачу об определенности или неопределенности проблемы моментов. Последняя оказывается тесно связанной с исследованиями граничных задач для дифференциальных операторов второго порядка. Случай моментов является по существу дискретным вариантом этой теории. Превосходное изложение этого аспекта теории моментов содержится в книге Ахиезера [1965]. Здесь методы гильбертова пространства используются красиво и удачно.

## § 1. Определение и общая структура $\mathcal{M}_{n+1}$

Пусть $\{u_i\}_0^n$ — чебышевская система на интервале $[a, b]$. Моментное пространство $\mathcal{M}_{n+1}$ по отношению к $\{u_i\}_0^n$ образуется следующим образом:

$$\mathcal{M}_{n+1} = \{\mathbf{c} = (c_0, c_1, \ldots, c_n) \in R^{n+1} | c_i = \\ = \int_a^b u_i(t)\, d\sigma(t), \quad i = 0, 1, \ldots, n\}, \qquad (1.1)$$

где $\sigma(t)$ принадлежит множеству всех неубывающих, непрерывных справа функций ограниченной вариации.

Заметим, что нормировка $\int_a^b d\sigma(t) = 1$ в определении не предполагается выполненной, так что $\mathcal{M}_{n+1}$ является конусом в евклидовом $n+1$-мерном пространстве $R^{n+1}$, т. е. если $\mathbf{c} \in \mathcal{M}_{n+1}$, то $\lambda\mathbf{c} \in \mathcal{M}_{n+1}$ для всех $\lambda \geqslant 0$.

Конус $\mathcal{M}_{n+1}$ не содержится в $n$-мерной гиперплоскости. Действительно, если $t_0, t_1, \ldots, t_n$ является $n+1$ различным значением на $[a, b]$, то точки $c^i = (u_0(t_i), \ldots, u_n(t_i)$, $i = 0, \ldots, n)$ все принадлежат к $\mathcal{M}_{n+1}$ и по чебышевскому свойству эти точки являются линейно независимыми.

Теорема 1.1. *Моментное пространство $\mathcal{M}_{n+1}$ является замкнутым выпуклым конусом.*

Доказательство. Что $\mathcal{M}_{n+1}$ выпукло и является конусом, очевидно. Покажем, что $\mathcal{M}_{n+1}$ замкнуто. Пусть $u(t) = \sum_{i=0}^{n} a_i u_i(t)$ обозначает строго положительный многочлен, существование которого гарантируется теоремой 5.1 гл. I. Далее предположим $\lim_{r\to\infty} \mathbf{c}(r) = \mathbf{c}$, где $\mathbf{c}^{(r)} \in \mathcal{M}_{n+1}$. Тогда для некоторой константы $M$

$$M \geqslant \sum_{i=0}^{n} c_i^{(r)} a_i = \int_a^b u(t)\, d\sigma^{(r)}(t) \geqslant (\min_{0\leqslant r\leqslant b} u(\tau)) \int_a^b d\sigma^{(r)}(t), \tag{1.2}$$

что показывает, что $\sigma^{(r)}(t)$ имеют равномерно ограниченную вариацию. Применяя принцип выбора Хелли и связанные с ним теоремы непрерывности, мы получим представление

$$\lim_{r\to\infty} c_i^{(r)} = c_i = \int_a^b u_i(t)\, d\sigma(t),$$

где $\sigma(t)$ — возрастающая функция ограниченной вариации. Таким образом, $\mathbf{c}$ принадлежит к $\mathcal{M}_{n+1}$. Мы сейчас рассмотрим другую характеризацию $\mathcal{M}_{n+1}$. Пусть $C_{i+1}$ есть кривая в $R^{n+1}$, что может быть представлено в параметрической форме следующим образом:

$$C_{t+1} = \{\gamma_t = (u_0(t), u_1(t), \dots, u_n(t)) \mid a \leqslant t \leqslant b\}. \tag{1.3}$$

Пусть $\mathcal{C}(C_{n+1})$ обозначает наименьший выпуклый конус, содержащий $C_{n+1}$. Мы требуем, чтобы $\mathcal{C}(C_{n+1})$ был замкнут и чтобы каждое $\gamma$, принадлежащее $\mathcal{C}(C_{n+1})$, могло быть представлено в форме

$$\gamma_i = \sum_{j=1}^{n+2} \lambda_j u_i(t_j), \qquad i = 0, 1, \dots, n \tag{1.4}$$

$$(\lambda_j \geqslant 0,\ a \leqslant t_j \leqslant b,\ j = 1, \dots, n+2).$$

Ясно, что множество $\Gamma$ всех векторов $\gamma$ вида (1.4) содержится в $\mathcal{C}(C_{n+1})$ и вследствие классической теоремы Каратеодори[1]) всякий вектор в конусе $C_{n+1}$ представим в виде (1.4). Следовательно, $\Gamma = \mathcal{C}(C_{n+1})$. Что $\Gamma$ замкнуто, доказывается аналогично в теореме 1.1.

В следующей теореме мы покажем, что $\mathcal{M}_{n+1} = \Gamma = \mathcal{C}(C_{n+1})$. Рис. 3 показывает «разрез» $\mathcal{M}_{n+1}$, где $c_0 = 1$, в случае $u_i(t) = t^i$,

---

[1]) Упомянутая теорема Каратеодори утверждает, что каждая точка, принадлежащая выпуклой оболочке данного множества $A$ в евклидовом $n+1$-мерном пространстве, может быть представлена как линейная комбинация $n+2$ точек из $A$ с суммой коэффициентов, равной 1.

$i = 0, 1, 2,$ и $[a, b] = [0, 1]$. Жирной линией изображена кривая $(t, t^2)$ или $(1, t, t^2)$, где $t$ пробегает интервал $[0, 1]$, и очерченная поверхность является сечением $\mathcal{M}_{n+1}$, получающимся для $c_0 = 1$. Соответствующее сечение для кривой $(1, t, t^2, t^3)$ при $t$, изменяющемся в $[0, 1]$, изображено на рис. 4.

Теорема 1.2. *$\mathcal{M}_{n+1}$ является внутренней конической оболочкой $\mathscr{C}(C_{n+1})$ из $C_{n+1}$.*

Доказательство. Ясно, что $\mathscr{C}(C_{n+1}) \subset \mathcal{M}_{n+1}$. Предположим, что $c^0 \in \mathcal{M}_{n+1}$, но $c^0 \notin \mathscr{C}(C_{n+1})$. Рассмотрим гиперплоскость, строго

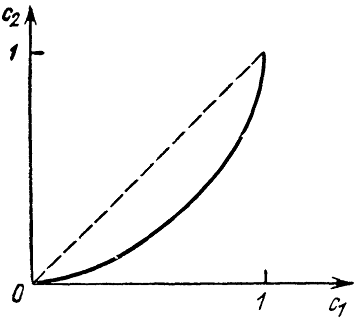



Рис. 3.

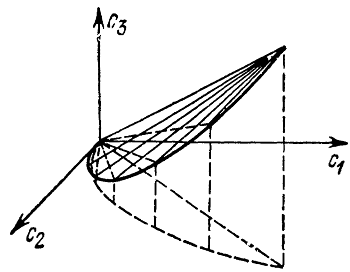



Рис. 4.

отделяющую $c^0$ от $\mathscr{C}(C_{n+1})$. Она существует, так как $\mathscr{C}(C_{n+1})$ является замкнутым выпуклым конусом. Иначе говоря, существуют действительные константы $a_i$, $i = 0, 1, \ldots, n$, не все равные нулю, и действительное $d$, для которых

$$\sum_{i=0}^{n} a_i c_i^0 + d < 0, \quad \sum_{i=0}^{n} a_i \gamma_i + d \geqslant 0, \qquad \gamma \in \mathscr{C}(C_{n+1}). \tag{1.5}$$

В частности,

$$\sum_{i=0}^{n} a_i \lambda u_i(t) + d \geqslant 0, \qquad a \leqslant t \leqslant b, \quad \lambda \geqslant 0. \tag{1.6}$$

Ясно, что $d \geqslant 0$, иначе (1.6) было бы несправедливо для $\lambda = 0$. Первое соотношение в (1.5) может быть записано как

$$\int_a^b \left( \sum_{i=0}^{n} a_i u_i(t) \right) d\sigma^0(t) + d < 0.$$

С другой стороны, интегрируя (1.6) относительно $\sigma^0(t)$ и выбирая $\lambda = \int_a^b d\sigma^0(t) > 0$, мы приходим к противоречию. Следовательно, $\mathcal{M}_{n+1} = \mathscr{C}(C_{n+1})$.

Определение 1.1. *Сечением* $\mathcal{M}_{n+1}$ будем называть всякое подмножество $S$ множества $\mathcal{M}_{n+1}$ со свойствами: $S$ находится в линейном многообразии, и, если $\mathbf{c}^0 \in \mathcal{M}_{n+1}$ и $\mathbf{c}^0 \neq 0$, то существует единственное $\lambda$ ($\lambda > 0$) такое, что $\lambda\mathbf{c}^0 \in S$. Сечения $\mathcal{M}_{n+1}$ получаются с помощью следующих построений. Зададим многочлен $u(t) = \sum_{i=0}^{n} a_i u_i(t)$, строго положительный на $[a, b]$. Тогда множество всех $\mathbf{c} \in \mathcal{M}_{n+1}$, удовлетворяющих равенству $\sum_{i=0}^{n} a_i c_i = \alpha > 0$, определяет сечение. Мы можем показать, что каждое сечение $S$ может быть получено таким образом. Предположим, что $S$ лежит на гиперплоскости $\sum_{i=0}^{n} a_i x_i = \alpha$, где $\alpha \geqslant 0$ и $\sum_{i=0}^{n} a_i^2 > 0$. В этом случае $\alpha > 0$, так как из $\alpha = 0$ следует, что $\sum_{i=0}^{n} a_i u_i(t) = 0$ для всех $t \in [a, b]$, что противоречит предположению $\sum_{i=0}^{n} a_i^2 > 0$. Многочлен $u = \sum_{i=0}^{n} a_i u_i$ может тогда использоваться для определения сечения $S$. Каждое сечение выпукло и ограничено.

## § 2. Граница и крайние лучи $\mathcal{M}_{n+1}$

Из теоремы 1.2 и уравнения (1.4) мы знаем, что каждое $\mathbf{c} \in \mathcal{M}_{n+1}$ допускает представление в форме (1.4), содержащей не более $n+2$ точек кривой $C_{n+1}$. Вообще говоря, существует много таких представлений. Мы покажем в § 3, что для $T$-систем каждое $\mathbf{c}$ из внутренности $\mathcal{M}_{n+1}$ допускает представление в форме (1.4), которое включает приблизительно $(n+2)/2$ точек кривой $C_{n+1}$. Однако перед анализом такого представления внутренних точек мы исследуем представление граничных точек $\mathcal{M}_{n+1}$. Следующее определение вводит важное понятие.

Определение 2.1. *Индексом* $I(\mathbf{c})$ точек $\mathbf{c}$ в $\mathcal{M}_{n+1}$ называется минимальное число точек из $C_{n+1}$, которые могут представить $\mathbf{c}$ как выпуклую комбинацию, причем при подсчете каждая из точек $(u_0(b), \ldots, u_n(b))$ и $(u_0(a), \ldots, u_n(a))$ считается за половину, а $(u_0(t), \ldots, u_n(t))$ при $a < t < b$ — за единицу. (Смысл вышеприведенной процедуры подсчета станет ясен позднее).

С помощью этого определения мы можем охарактеризовать очень просто границу.

Теорема 2.1. *Вектор* $\mathbf{c}^0 \in \mathcal{M}_{n+1}$ $(\mathbf{c}^0 \neq 0)$ *является граничной точкой* $\mathcal{M}_{n+1}$ *тогда и только тогда, когда* $I(\mathbf{c}^0) < (n+1)/2$. *Бо-*

*лее того, каждая граничная точка* $\mathbf{c}^0$ *допускает единственное представление*

$$c_i^0 = \sum_{j=1}^{p} \lambda_j u_i(t_j), \qquad i = 0, 1, \ldots, n, \tag{2.1}$$

*где* $p \leqslant \frac{n+2}{2}$ *и* $\lambda_j > 0$, $j = 1, 2, \ldots, p$.

Доказательство. Предположим, что $\mathbf{c}^0 \in \mathscr{D}\mathscr{M}_{n+1}$ ($\mathbf{c}^0 \neq 0$). Тогда существует опорная гиперплоскость к $\mathscr{M}_{n+1}$ в $\mathbf{c}^0$. Она может быть задана вектором $\{a_i\}_0^n$ и $d$, причем $\sum_{i=0}^{n} a_i^2 > 0$. Для этих точек справедливо неравенство

$$\sum_{i=0}^{n} a_i c_i + d \geqslant 0, \tag{2.2}$$

где $\mathbf{c} \in \mathscr{M}_{n+1}$ и $\sum_{i=0}^{n} a_i c_i^0 + d = 0$.

Покажем, что $d = 0$. Это геометрически ясно, так как $\mathscr{M}_{n+1}$ — конус и любая гиперплоскость, которая является касательной и опорной для конуса, должна проходить через начало координат. Формальное доказательство проводится следующим образом. Ясно, что $d \geqslant 0$, так как $\mathbf{c} = (0, \ldots, 0) \in \mathscr{M}_{n+1}$. Если $d > 0$, тогда $\sum_{i=0}^{n} a_i c_i^0 < 0$, так как выполняется уравнение $\sum_{i=0}^{n} a_i c_i^0 + d = 0$. Однако, так так $\lambda \mathbf{c}^0 \in \mathscr{M}_{n+1}$ для всех $\lambda \geqslant 0$, мы должны иметь $\lambda \sum_{i=0}^{n} a_i c_i^0 + d < 0$ для достаточно больших $\lambda$, что противоречит (2.2). Следовательно,

$$\begin{cases} \sum_{i=0}^{n} a_i c_i \geqslant 0 \quad \text{для всех} \quad \mathbf{c} \in \mathscr{M}_{n+1}, \\ \sum_{i=0}^{n} a_i c_i^0 = 0. \end{cases} \tag{2.3}$$

Теперь определим $n^0(t) = \sum_{i=0}^{n} a_i u_i(t)$. Сравнивая с (2.3), мы видим, что $u^0(t) \geqslant 0$ для $a \leqslant t \leqslant b$ и $\int u^0(t)\, d\sigma^0(t) = 0$, где $c_i^0 = \int u_i(t)\, d\sigma^0(t)$. Следовательно, $u^0(t) = 0$ в каждой точке роста $\sigma^0$. Это означает, что спектр $\sigma^0$ содержится вне носителя $u^0(t)$.

Таким образом, так как $u^0(t) \geqslant 0$, то имеем $2I(\mathbf{c}^0) = \tilde{Z}(u^0)$ (где обозначение $\tilde{Z}(u^0)$ согласуется с обозначением определения 4.3 гл. I). Однако по теореме 4.2 гл. I $\tilde{Z}(u^0) \leqslant n$, так что $I(\mathbf{c}^0) < (n+1)/2$. Далее рассмотрим представление $\mathbf{c}^0$ вида (2.1), где $T = \{t_j\}^m$ — корни $u^0(t)$, так что $T$ содержит $n+1$ различных точек. Так как $\det \| u_i(t) \| > 0$, мы выводим, что коэффициенты $\lambda_j$ однозначно определяются по $\mathbf{c}^0$. Обратно, пусть $\mathbf{c}^0$ обозначает вектор из $\mathcal{M}_{n+1}$, для которого $I(\mathbf{c}^0) < (n+1)/2$. Таким образом, $\mathbf{c}^0$ допускает представление вида (2.1). Согласно теореме 5.1 гл. I мы можем построить неотрицательный нетривиальный полином $u^0(t)$, образующийся в нуле на $T = \{t_j\}$, где $t_j$ соответствуют точкам $C_{n+1}$, входящим в представление (2.1). Ясно, что коэффициенты полинома $u^0(t)$ определяют опорную гиперплоскость к $\mathcal{M}_{n+1}$ в $\mathbf{c}^0$ и, следовательно, $\mathbf{c}^0$ лежит на границе с $\mathcal{M}_{n+1}$. Доказательство теоремы 2.1 завершено.

Следствие 2.1. *Для $n \geqslant 2$ граничные лучи $\mathcal{M}_{n+1}$ представимы в форме $\lambda\mathbf{c}$, где $\lambda \geqslant 0$ и $\mathbf{c}$ лежит на $C_{n+1}$.*

Доказательство. Согласно теореме 1.2 мы имеем $\mathcal{M}_{n+1} = \mathcal{C}(C_{n+1})$, так что достаточно провести опорную гиперплоскость к $\mathcal{M}_{n+1}$, которая касается только вдоль луча $(\lambda u_0(t), \dots, \lambda u_n(t))$, $\lambda > 0$, $t$ фиксировано. Если $t \in (a, b)$, то мы применяем теорему 5.1 гл. I, которая утверждает существование неотрицательного многочлена $u$ с единственным нулем в $t$. Заметим, что нам необходимо, чтобы $n \geqslant 2$, так как $\tilde{Z}(u) = 2 \leqslant n$. Коэффициенты этого многочлена дают нужную опорную гиперплоскость[1]). Если $t = a$, то существует неотрицательный многочлен, обращающийся в нуль в граничной точке $a$ и, возможно, в другой граничной точке $t = b$. Предположим теперь противное, что $(u_0(a), \dots, u_n(a))$ не принадлежит граничному лучу. Тогда

$$(u_0(a), \dots, u_n(a)) = \alpha\mathbf{c}^{(1)} + (1-\alpha)\mathbf{c}^{(2)}, \tag{2.4}$$

где точки $\mathbf{c}^{(1)}$, $\mathbf{c}^{(2)}$ и $(u_0(a), \dots, u_n(a))$ лежат на различных лучах. Если $\sigma_1$ и $\sigma_2$ — меры, соответствующие $\mathbf{c}^{(1)}$ и $\mathbf{c}^{(2)}$, тогда их объединенный спектр должен содержать точки открытого интервала $(a, b)$. При этом условии мы приходим к противоречию, так как, согласно (2.4),

$$0 = u(a) = \alpha \int_a^b u\, d\sigma_1 + (1-\alpha) \int_a^b u\, d\sigma_2,$$

в то время как многочлен $u$ не обращается в нуль на $(a, b)$.

---

[1]) Мы будем часто говорить об опорном многочлене, подразумевая, что коэффициенты многочлена определяют опорную гиперплоскость в $\mathcal{M}_{n+1}$.

## § 3. Внутренность $\mathcal{M}_{n+1}$

Определение 3.1. Значения $t_j$, входящие в представление

$$c_i^0 = \sum_{j=1}^{p} \lambda_j u_i(t_j), \qquad i = 0, 1, \dots, n, \quad \lambda_j > 0, \quad 1 \leqslant j \leqslant p, \tag{3.1}$$

называются *корнями представления*. Мы также иногда будем говорить, что $\mathbf{c}^0$ содержит $t_1, \dots, t_p$. В записи (3.1) мы будем всегда подразумевать, если не оговорено противное, что $\{t_j\}$ расположены в возрастающем порядке. Возрастающую функцию $\sigma$, которая изменяется только в точках $\{t_j\}$ с относительными весами $\lambda_j$, назовем ассоциированной мерой представления. Индекс множества $T = \{t_1, \dots, t_p\}$ определяется как число, получающееся при подсчете внутренних точек с весом 1 и граничных точек $a$ и $b$ с весом 1/2. *Индексом* меры $\sigma$, порождающей $\mathbf{c}^0$, мы будем называть индекс ее корней в представлении (3.1). (Заметим, что имеется некоторое совпадение между определением 3.1 и определением 2.1.)

Относительно представления внутренних точек $\mathcal{M}_{n+1}$ мы имеем следующий важный результат.

Теорема 3.1. *Пусть* $\mathbf{c}^0$ — *внутренняя точка* $\mathcal{M}_{n+1}$. *Для каждого* $t^*$, $a \leqslant t^* \leqslant b$, *существует представление*

$$c_i^0 = \sum_{j=1}^{p} \lambda_j u_i(t_j), \qquad i = 0, 1, \dots, n, \quad \lambda_j > 0, \quad j = 1, 2, \dots, p,$$

*с индексом* $(n+1)/2$ *или* $(n+2)/2$, *которое содержит точку* $t^*$.

Доказательство. Рассмотрим сечение $S_{n+1}$ множества $\mathcal{M}_{n+1}$ такое, что $\mathbf{c}^0$ — внутренняя точка в $S_{n+1}$ (внутренность рассматривается относительно линейного расширения $S_{n+1}$). Проведем прямую $L$ из $\mathbf{c}^*(\lambda u_0(t^*), \dots, \lambda u_n(t^*))$ в $S_{n+1}$ через $\mathbf{c}^0$, где значение $\lambda$ определяется таким образом, чтобы $\mathbf{c}^* \in S_{n+1}$. Пусть $\tilde{\mathbf{c}} \neq \mathbf{c}^*$ будет второй точкой $\mathcal{M}_{n+1}$, где $L$ пересекает границу $\mathcal{M}_{n+1}$ в $S_{n+1}$. Ясно, что

$$\mathbf{c}^0 = \alpha\tilde{\mathbf{c}} + (1 - \alpha)\,\mathbf{c}^* \tag{3.2}$$

для некоторого $\alpha$ $(0 < \alpha < 1)$. Более того, $I(\tilde{\mathbf{c}})$ равно $(n-1)/2$ или $n/2$, ибо в противном случае $I(\mathbf{c}^0) < (n+1)/2$, что приводит к тому, что $\mathbf{c}^0$ принадлежит границе $\mathcal{M}_{n+1}$, в противоречии с предположением. Требуемое представление получается из (3.2). Представление $\mathbf{c}^0$, даваемое теоремой 3.1, с индексным значением $(n+1)/2$ или $(n+2)/2$, является единственным для любого заданного $t^*$, $a < t^* < b$.

Доказательство этого факта откладывается до обсуждения результатов теоремы 3.2.

Определение 3.2. Пусть $\mathbf{c}^0$ — внутренняя точка $\mathcal{M}_{n+1}$. Представление для $\mathbf{c}^0$ с индексом $I(\mathbf{c}^0) = (n+1)/2$ называется *основным*, а всякое представление с индексом $I(\mathbf{c}^0) \leqslant (n+2)/2$ — *канониче-*

*ским*. Каноническое или главное представление будет называться *верхним*, если оно содержит конечную точку $b$, и *нижним* в противном случае.

Теорема 3.1 утверждает, что для всякого $\mathbf{c}^0 \in \operatorname{Int}(\mathcal{M}_{n+1})$ существует каноническое представление, содержащее произвольную заданную точку $t^*$ с положительным весом. Мы покажем, что только два из этих представлений являются основными, одно из них — верхнее, другое — нижнее.

Замечание 3.1. Если мы прибавим к $T$-системе $\{u_i\}_0^n$ другую функцию $u_{n+1}(t)$ такую, что $\{u_i\}_0^{n+1}$ образует снова $T$-систему, то каноническое представление с индексом $(n+2)/2$ внутренней точки $(c_0^0, c_1^0, \ldots, c_n^0)$ будет соответствовать основному представлению точки $(c_0^0, c_1^0, \ldots, c_n^0, c_{n+1})$ при всяком $c_{n+1}$. Верхнее и нижнее основные представления $(c_0^0, c_1^0, \ldots, c_n^0)$ с индексом $(n+1)/2$ соответствуют однозначному представлению верхнего и нижнего граничного значений величины $c_{n+1}$ (см. § 6).

Случай I ($n = 2m$). Имеется по крайней мере два основных представления; опишем их, классифицируя по корням.

Первое основное представление получается при применении построения теоремы 3.1 с $t^* = a$. Так как всякая точка $\mathcal{D}\mathcal{M}_{n+1}$ имеет индекс, меньший $(n+1)/2 = m + 1/2$, то мы должны иметь $I(\tilde{\mathbf{c}}) = m$ в (3.2). Также представление $\tilde{\mathbf{c}}$ не может содержать $b$, так как иначе $I(\tilde{\mathbf{c}})$ не целое. Итак, для $t^* = a$ представление (3.2) имеет индекс $m + 1/2$, и, следовательно, является главным. Второе главное представление получается аналогично при $t^* = b$. В итоге для $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}$ существуют ровно два главных представления с корнями

$$\begin{gathered} a = t_1^* < t_2^* < \ldots < t_{m+1}^* < b, \\ s_1^* < s_2^* < \ldots < s_{m+1}^* = b. \end{gathered} \tag{3.3}$$

В соответствии с определением 3.2 главное представление, содержащее $b$, называется верхним, а другое главное представление — нижним. Эти названия будут обоснованы в § 6 (см. также замечание 3.1).

Случай II ($n = 2m + 1$). Придавая $t^*$ значение $a$, мы получаем основное представление, которое автоматически также содержит $b$, так как здесь мы имеем $I(\tilde{\mathbf{c}}) = m + 1/2$. По определению это представление является верхним основным представлением и его корнями являются

$$a = s_1^* < s_2^* < \ldots < s_{m+1}^* < s_{m+2}^* = b. \tag{3.4}$$

Другое основное представление строится иначе. В этом случае мы рассмотрим часть кривой $C_{n+1}$, именно $C_{n+1}(d) = \{\mathbf{v}_t = (u_1(t), \ldots, u_n(t)) \mid d < t < b\}$, и обозначим $\mathcal{M}_{n+1}(d)$, $d > a$, выпуклый конус, натянутый на кривую $C_{n+1}(d)$. Мы утверждаем, что существует

$d' > a$, для которого $c^0 \in \mathcal{D}\mathcal{M}_{n+1}(d')$. Если это не так, то мы можем разбить $(a, b)$ на два непустых открытых множества $A_1 \cup A_2$, где $A_1 = \{d \mid c^0 \in \operatorname{Int} \mathcal{M}_{n+1}(d)\}$ и $A_2 = \{d \mid c^0 \in \mathcal{M}_{n+1}(d)\}$. Таким образом, получается противоречие, так как открытый интервал $(a, b)$ — связное множество. Следовательно, $c^0 \in \mathcal{D}\mathcal{M}_{n+1}(d')$ для некоторого $d' \in (a, b)$. Но из $c^0 \in \mathcal{D}\mathcal{M}_{n+1}(d')$ и $c^0 \in \operatorname{Int} \mathcal{M}_{n+1}$ по теореме 2.1 вытекает существование представления с индексом $m + 1/2$ относительно $[d', b]$, которое не содержит $b$. Таким образом, $d'$ входит в представление. Относительно интервала $[a, b]$ это представление имеет индекс $m + 1$ и корни

$$t_1^* < t_2^* < \ldots < t_{m+1}^* \; (a < t_1^*, t_{m+1}^* < b). \tag{3.5}$$

Нижнее главное представление таким образом определено.

Мы будем обозначать меры, индуцированные верхним и нижним представлением, через $\bar{\sigma}$ и $\underline{\sigma}$ соответственно. Следующие две теоремы показывают, что имеется только два главных представления, и выделяют связи между корнями этих представлений. Более точно, мы анализируем связи между корнями двух независимых канонических представлений. Нам будет нужна следующая лемма.

Лемма 3.1. *Пусть* $c^0 \in \operatorname{Int} \mathcal{M}_{n+1}$, *и пусть* $\sigma^*$ *и* $\sigma$ — *две различные меры, удовлетворяющие условию* $\int u_i u\sigma^* = \int u_i d\sigma = c_i^0$, *где* $\sigma$ — *каноническая мера, т. е.* $\sigma$ *имеет индекс* $(n+1)/2$ *или* $(n+2)/2$. *Тогда для каждой пары точек* $t_j$, $t_{j+1}$, *в которых сосредоточены массы меры* $\sigma$, *существует точка роста меры* $\sigma^*$, *лежащая в открытом интервале* $(t_j, t_{j+1})$. *Если* $\sigma$ *имеет индекс* $(n+1)/2$, *то это остается справедливым при условии* $t_j = a$ *или* $t_{j+1} = b$.

Доказательство. Пусть $t_j$, $t_{j+1}$ — два последовательных корня $\sigma$. Если $\sigma$ имеет индекс $(n+2)/2$, то будем считать, что корни расположены на открытом интервале $(a, b)$, а если $\sigma$ имеет индекс $(n+1)/2$, то тогда будем допускать, что один из них лежит в конечной точке. Если $\sigma$ не растет внутри $(t_j, t_{j+1})$, то мы построим, согласно теореме 5.2 гл. I, полином

$$u(t) = \begin{cases} \geqslant 0, & t \notin [t_j, t_{j+1}], \\ < 0, & t \in [t_j, t_{j+1}], \end{cases}$$

обращающийся в нуль на $(a, b)$ только на корнях $\sigma$. Заметим, что $t_j$ и $t_{j+1}$ являются узловыми корнями. Ясно, что

$$0 = \int u(t)\,(d\sigma^* - d\sigma) = \int u(t)\, d\sigma^* =$$
$$= \int\limits_{[a, t_j]} u(t)\, d\sigma^* + \int\limits_{[t_{j+1}, b]} u(t)\, d\sigma^* \geqslant 0. \tag{3.6}$$

Соотношение (3.6) возможно только в случае, когда $\sigma^*$ изменяется только в корнях $\sigma$. Но $\sigma$ имеет менее чем $n + 1$ корней $(n \geqslant 1)$

и, следовательно, $\sigma$ и $\sigma^*$ совпадают, что противоречит условиям леммы. Отсюда вытекает, что $\sigma^*$ изменяется на интервале $(t_j, t_{j+1})$, что и нужно было показать.

Следствие 3.1. *Для каждого* $c^0 \in \operatorname{Int} \mathcal{M}_{n+1}$ *существует ровно два основных представления. Между каждыми двумя корнями одного представления лежит ровно один корень другого.*

Доказательство. Мы уже выделили два основных представления $\underline{\sigma}$ и $\bar{\sigma}$ для каждого $c^0 \in \operatorname{Int} \mathcal{M}_{n+1}$. Корни этих представлений должны строго перемежаться (по лемме 3.1).

Предположим, что существует другое главное представление, отличное от двух уже построенных. Если $n$ нечетно, скажем $n = 2m + 1$, и это представление имеет $m + 1$ внутренних корней $a < t'_1 < \ldots < t'_{m+1} < b$, то эти корни должны лежать между корнями (3.5). Не теряя общности, мы можем предположить, что $t'_1 < t^*_1$. Мы построим нетривиальный полином $u$ с неузловыми нулями в $t'_i$, $i = 2, \ldots, m + 1$, и узловым нулем в $t'_1$ и такой, что $u(t) \geqslant 0$ для $t \geqslant t'_1$. Тогда $0 = \int u\, d\underline{\sigma} - \int u\, d\sigma = \int u\, d\underline{\sigma} > 0$, что абсурдно. Во всех других случаях представление $\sigma$ должно содержать один или оба конца $a$ и $b$. Так как $a$ и $b$ являются корнями $\bar{\sigma}$ и $\underline{\sigma}$, то, следовательно, существует два главных представления, имеющих по крайней мере одну из граничных точек. Мы сразу же придем к противоречию, используя лемму 3.1, если сравним корни этих двух представлений.

Теорема 3.2. *Пусть* $c^0 \in \operatorname{Int} \mathcal{M}_{n+1}$. *Рассмотрим два различных представления* $\sigma'$ *и* $\sigma''$ *точек* $c^0$ *с индексом не большим, чем* $(n + 2)/2$, *корнями которых соответственно являются* $\{t'_j\}_1^p$ *и* $\{t''_j\}_1^q$, *т. е.*

$$c_i^0 = \sum_{j=1}^{p} \lambda'_j u_i(t'_j) = \sum_{j=1}^{q} \lambda''_j u_i(t''_j), \qquad i = 0, 1, \ldots, n. \tag{3.7}$$

*Тогда*

$$\left[\frac{n+2}{2}\right] \leqslant p, \quad q \leqslant \left[\frac{n+2}{2}\right] + 1$$

(*[r], как обычно, обозначает целую часть r*). *Корни* $\{t'_j\}^p$ *и* $\{t''_j\}^p$ *в* $(a, b)$ *строго перемежаются и могут быть одной или обеими граничными точками* $a$ *и* $b$. *Более того, если* $t'_1 = t''_1 = a$, *то* $\lambda'_1 \neq \lambda''_1$ *и* $\lambda' > \lambda''$ *тогда и только тогда, когда* $t'_2 > t''_2$. *Аналогично, если* $t'_p = t''_q = b$, *то* $\lambda'_p \neq \lambda''_q$ *и* $\lambda'_p > \lambda''_q$ *тогда и только тогда, когда* $t''_{q-1} > t'_{p-1}$.

Доказательство. Соотношение (3.8) легко устанавливается проверкой различных ситуаций, которые могут возникнуть, когда $n$ четно и нечетно.

Чтобы показать, что внутренние корни $\{t'_j\}^p$ и $\{t''_j\}^q$ строго перемежаются, мы разберем возможные случаи сначала для $n$ нечетного и затем для $n$ четного. Во всех случаях число внутренних точек среди $\{t_j'\}^p$ и $\{t_j''\}^q$ отличается, самое большее, на единицу, так что корни должны строго перемежаться (по лемме 3.1). Заметим, что случай, когда оба представления главные, был полностью разобран в следствии 3.1. Мы теперь допустим, что индекс $\sigma'$ равен $(n+2)/2$ и что $t_1' = t_1'' = a$. Если $\sigma''$ имеет индекс $(n+1)/2$, то тогда (по лемме 3.1) $t'' > t'$, так что $\sigma$ должна иметь точку роста $b\,(t_1'', t_2'')$. Если $\sigma''$ также имеет индекс $(n+2)/2$, то мы можем предположить без потери общности, что $t_2'' > t_2'$. Мы можем построить на основании теоремы 5.2 гл. I нетривиальный многочлен, имеющий узловые нули в $t_2'$, а также в конечной точке $b$, если $t_p' = b$, и неузловые нули в других $t_j'$, $j = 3, \ldots, p$. В обоих случаях $n = 2m$ или $n = 2m+1$ этот многочлен имеет точно $n$ нулей, если мы считаем неузловые нули дважды и узловые один раз. По теореме 4.2 гл. I многочлен обращается в нуль точно в точках $t_j$, $j = 2, \ldots, p$, и мы можем задать $u(a) = -1$. Мы тогда имеем

$$0 = \int u\,d\sigma'' - \int u\,d\sigma' = \int\limits_{[t_2', b]} u\,d\sigma'' + \lambda_1' - \lambda_1''.$$

Требуемое неравенство $\lambda_1'' > \lambda_1'$ следует из того, что

$$\int\limits_{[t_2', b]} u\,d\sigma'' > 0.$$

С помощью аналогичных соображений доказывается $t_q'' = t_p' = b$.

Следствие 3.2. *Пусть* $c^0 \in \operatorname{Int} \mathcal{M}_{n+1}$, *тогда для всякого* $t^*$, $a < t^* < b$, *существует единственное каноническое представление* $c^0$, *содержащее* $t^*$. *При особом выборе* $t^*$ *каноническое представление оказывается основным.*

## § 4. Экстремальная характеризация канонической меры

Каноническое представление может быть охарактеризовано с точки зрения экстремальных задач. Пусть $c^0 \in \operatorname{Int} \mathcal{M}_{n+1}$. Мы определим функционал $\mathcal{E}(u) = \sum_{i=0}^{n} c_i^0 a_i$ для всякого полинома $u(t) = \sum_{i=0}^{n} a_i u_i(t)$. Рассмотрим следующую задачу. Определим полином, доставляющий минимум

$$\min_{u \in \mathcal{P}_{n+1}} \frac{\mathcal{E}(u)}{u(t_0)} = \rho(t_0) \tag{4.1}$$

($t_0$ фиксированно, $a \leqslant t_0 \leqslant b$), где $\mathcal{P}_{n+1}$ обозначает множество всех неотрицательных многочленов и $\mathscr{E}(u)/u(t_0)$ считается равным $+\infty$, если $u(t_0)=0$. (Мы пишем минимум в (4.1), так как ниже будет показано, что нижняя грань достигается.)

Ясно, что если $c_i^0=\int u_i(t)\,d\sigma(t)$, $i=0,\dots,n$, то тогда для $u\in\mathscr{P}_{n+1}$ имеем

$$\mathscr{E}(u)=\int u(t)\,d\sigma(t)\geqslant\sigma(\{t_0\})\,u(t_0), \tag{4.2}$$

где $\sigma(\{t_0\})$ обозначает меру множества, состоящего из точки. Это показывает, что

$$\inf_{u\in\mathscr{P}_{n+1}}\frac{\mathscr{E}(u)}{u(t_0)}\geqslant\sup_{\sigma\in V(\mathbf{c}^0)}\sigma(\{t_0\}), \tag{4.3}$$

где верхняя грань берется по множеству всех мер, представляющих $\mathbf{c}^0$, т. е.

$$V(\mathbf{c}^0)=\{\sigma\mid c_i^0=\int u_i\,d\sigma,\quad i=0,1,\dots,n\}. \tag{4.4}$$

Пусть $\sigma_{t_0}$ есть мера единственного канонического (или главного) представления $\mathbf{c}^0$, которое содержит $t_0$ для $t_0\in(a,b)$, и главного представления, содержащего $t_0$, когда $t_0=a$ или $t=b$.

Разбирая различные случаи для $n$ четного и нечетного, мы видим, что можно, согласно теореме 5.1 гл. I, построить многочлен $u^*\in\mathscr{P}_{n+1}$, который обращается в нуль на всех корнях $\sigma_{t_0}$, кроме точки $t_0$ и такой, что $u^*(t_0)>0$. Для этого специального многочлена имеем

$$\frac{\mathscr{E}(u^*)}{u^*(t_0)}=\sigma(\{t_0\})$$

и, сравнивая с (4.3), мы видим, что

$$\frac{\mathscr{E}(u^*)}{u^*(t_0)}\geqslant\inf_{u\in\mathscr{P}_{n+1}}\frac{\mathscr{E}(u)}{u(t_0)}\geqslant\sup_{\sigma\in V(\mathbf{c}^0)}\sigma(\{t_0\})\geqslant\sigma_{t_0}(\{t_0\}).$$

Следовательно,

$$\min_{u\in\mathscr{P}_{n+1}}\frac{\mathscr{E}(u)}{u(t_0)}=\sigma_{t_0}(\{t_0\}),$$

$$\rho(t_0)=\sigma_{t_0}(\{t_0\})=\max_{\sigma\in V(\mathbf{c}^0)}\sigma_0(\{t_0\}).$$

Более того, из (4.2) ясно, что нижняя грань в (4.3) достигается, только если $u$ обращается в нуль во всех корнях $\sigma_{t_0}$, исключая $t_0$. Если $t_0$ не внутренний корень одного из главных представлений, то многочлен $u^*$ имеет $n$ нулей, если считать неузловые нули дважды и узловые нули один раз, т. е. $\widetilde{Z}(u^*)=n$. Если $\{u_i\}_0^n$ является $ET$-системой порядка 2, то многочлен $u^*$, доставляющий минимум в (4.1), является однозначно определенным с точностью до постоянного множителя. Если $t_0$ — внутренний корень основного представления, то $\widetilde{Z}(u^*) = n-1$. В этом случае, если мы ограничим рассмотрение множеством $\mathcal{P}_n$ всех неотрицательных многочленов, построенных по $\{u_i\}_0^{n-1}$, то $u^*$ является однозначно определенным, если $\{u_i\}_0^{n-1}$ является $ET$-системой порядка 2.

Итоги предшествующего анализа мы подводим в следующей теореме.

Теорема 4.1 (Крейн [1951]).

(a) *Пусть* $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}$, *тогда среди всех* $\sigma \in V(\mathbf{c}^0)$ *единственная мера, обладающая максимальной массой* $\rho(t_0)$ *в точке* $t_0$, *является канонической мерой* $\sigma_{t_0}$, *имеющей скачок в* $t_0$, *при условии, что* $t_0 \in (a, b)$ (см. *следствие* 3.2), *и основным представлением со скачком в* $t_0$, *когда* $t_0 = a$ *или* $b$. *Более того,* $\rho(t_0)$ *может быть вычислена по формуле* (4.1).

(b) *Пусть* $\{u_i\}_0^n$ *будет* $ET$-*системой порядка* 2 *и предположим, что* $t_0$ *не является внутренним корнем главного представления* $\mathbf{c}^0$. *Тогда многочлен* $u^* \in \mathcal{P}_{n+1}$, *нормированный так, что* $u^*(t_0) = 1$, *доставляющий минимум в* (4.1), *однозначно определяется условием, что* $u^*$ *обращается в нуль в корнях* $\sigma_{t_0}$, *исключая* $t_0$.

*Если* $t_0$ *является внутренним корнем главного представления и* $\{u_i\}_0^{n-1}$ *является* $ET$-*системой порядка* 2, *то* $u^*$ (*удовлетворяющий* $u^*(t) = 1$ *и доставляющий минимум в* (4.1)) *однозначно определяется в* $\mathcal{P}_n$ — *множестве всех неотрицательных многочленов, построенных по* $\{u_i\}_0^{n-1}$.

## § 5. Свойства непрерывности канонического представления

Пусть $\mathbf{c}^0 = (c_0^0, c_1^0, c_2^0, \ldots, c_n^0) \in \operatorname{Int} \mathcal{M}_{n+1}$. Ниже обсуждаются результаты для случая $n = 2m$ и формулируются без доказательства соответствующие утверждения для случая $n = 2m+1$.

Рассмотрим два главных представления $\mathbf{c}^0$, описываемые их корнями:

$$a = t_1^* < t_2^* < \ldots < t_{m+1}^* \quad \text{(нижнее)}, \tag{5.1}$$

$$s_1^* < s_2^* < \ldots < s_{m+1}^* = b \quad \text{(верхнее)}, \tag{5.2}$$

где

$$t_i^* < s_i^* < t_{i+1}^* < s_{i+1}^* \qquad (i = 1, 2, \ldots, m). \tag{5.3}$$

Приведем утверждение, обратное тому, что корни в главных представлениях $\mathbf{c}^0$ перемежаются, как в (5.3).

Теорема 5.1 ($n = 2m$). *Пусть два множества точек вида* (5.1) *и* (5.2) *удовлетворяют* (5.3). *Тогда существует единственный луч* $\lambda \mathbf{c}^0$, $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}$, *такой, что два главных представления* $\mathbf{c}^0$ *имеют корнями* (5.1) *и* (5.2). *Аналогичное утверждение справедливо для* $n = 2m+1$.

Доказательство. Пусть $\{r_k\}_{k=1}^{n+1}=\{t_i^*\}_{i=1}^{m+1}\bigcup\{s_i^*\}_{i=1}^{m+1}$ и эти точки расположены в возрастающем порядке. Рассмотрим систему из $n+1$ уравнения с $n+2$ переменными $\{a_k\}_{k=1}^{n+2}$

$$\sum_{k=1}^{n+2} a_k u_i(r_k)=0 \qquad i=0,\dots,n. \tag{5.4}$$

Матрица $U=\|u_i(r_k)\|$ имеет ранг $n+1$, и компоненты единственного решения (5.4), с точностью до постоянного множителя, пропорциональны алгебраическим дополнениям, полученным вычеркиванием одного из столбцов $u$. Так как $\{u_i\}$ составляет $T$-систему, мы заключаем, что $a_k$ строго различны по знаку. Зафиксировав нормирующий множитель, мы можем записать (5.4)' в форме

$$\sum_{j=1}^{m+1}\lambda_j^* u_i(t_j^*)=\sum_{j=1}^{m+1}\alpha_j^* u_i(s_j), \qquad i=0,1,\dots,n, \tag{5.5}$$

где $\lambda_j^*>0$, $\alpha_j^*>0$ $(j=1,2,\dots,m+1)$. Компоненты вектора $\mathbf{c}^0$ тогда равны величинам из (5.5). Ясно, в силу теоремы 2.1, что $\mathbf{c}^0$ — внутренняя точка в $\mathcal{M}_{n+1}$.

Замечание 5.1. Для фиксированного сечения $S$ множества $\mathcal{M}_{n+1}$ доказанная выше теорема утверждает, что существует взаимно однозначное соответствие между парами строго перемежающихся корней (5.1) и (5.2) и $\operatorname{Int} S$. Тот же самый результат справедлив для $n$ нечетного, если мы возьмем корни (3.4) и (3.5). Легко показать, что это соответствие непрерывно в обе стороны. Если $u^0(t)\equiv 1$ и мы зададим сечение $S=M^n$ условием, что $c^0=1$, то, как будет показано в гл. III (см. теорему 2.1 гл. III), веса $\{\lambda_j^*\}_1^{m+1}$ и $\{a_j^*\}_1^{m+1}$, соответствующие (5.1) и (5.2) для данной моментной точки $\mathbf{c}\in\operatorname{Int} M^n$, удовлетворяют условию

$$0<\sum_{j=1}^{k}\lambda_j^*<\sum_{j=1}^{k}a_j^*<\sum_{j=1}^{k+1}\lambda_j^*, \qquad k=1,\dots,m$$
$$\left(\sum_{j=1}^{m+1}\lambda_j^*=\sum_{j=1}^{m+1}a_j^*=1\right).$$

Можно также показать, что написанные выше неравенства определяют единственную моментную точку $\mathbf{c}\in\operatorname{Int} M^n$ со свойствами, что $\{\lambda_j^*\}_1^{m+1}$ и $\{a_j^*\}_1^{m+1}$ являются весами нижнего и верхнего главного представления $\mathbf{c}$ (см. Карлин и Шумейкер [1966]). В дальнейшем мы начнем изучать свойства непрерывности канонической меры. Для этой цели введем интервалы

$$J_i=(t_i^*, s_i^*), \qquad i=1,2,\dots,m+1,$$
$$K_i=(s_i^*, t_{i+1}^*), \qquad i=1,2,\dots,m.$$

Рассмотрим $\xi \in J_1$. Мы хотим более полно описать природу корней $\{t_i(\xi)\}_{i=1}^{m+1}$ канонической меры $\sigma_\xi$.

Теорема 3.2 о взаимном расположении, примененная к множествам (5.1), (5.2) и $\{t_i(\xi)\}$, означает, что $t_i(\xi) \in J_i$, $i = 1, \ldots$ $\ldots, m+1$ $(t_1(\xi) = \xi)$. Более того, из критерия единственности вытекает, что когда $\xi$ пробегает $J_1$ непрерывно и монотонно слева направо, то $t_i(\xi)$ изменяется на $J_i$ непрерывно и монотонно слева направо. Конечно, корни канонической меры являются непрерывными функциями любого из них при непрерывном его изменении.

Далее рассмотрим $\xi$ в $K_1$. Снова, так как корни перемежают множества (5.1) и (5.2), и, имея в виду, что каждое $\xi$ принадлежит к некоторой канонической мере $\sigma_\xi$, мы находим, что

$$t_0(\xi) = a, \quad t_1(\xi) = \xi, \quad t_i(\xi) \in K_i, \qquad i = 2, \ldots, m$$

и

$$t_{m+1}(\xi) = b.$$

Снова, на основе единственности канонической меры и теоремы 3.2, выводим, что, когда $\xi$ пробегает $K_1$ непрерывно слева направо, $t_i(\xi)$ пробегает $K_i$ непрерывно слева направо. Вес $\lambda_0$ в $t_0(\xi) = a$ представления возрастает монотонно от нуля к максимальной массе $\rho(a)$, в то время как вес $\lambda_{m+1}$ в $t_{m+1}(\xi) = b$ уменьшается непрерывно от $\rho(b)$ до нуля.

Непрерывность $t_i(\xi)$ как функции $\xi$ также ясна из геометрических построений теоремы 3.1. Что веса изменяются непрерывно, становится очевидным, когда мы представим их как решение неособой системы линейных уравнений, у которой коэффициенты матрицы изменяются непрерывно и она остается неособой. Монотонность вытекает из единственности.

В случае $n = 2m + 1$ мы сформулируем результаты без доказательства.

Два основных представления имеют вид

$$t_1^* < t_2^* < \ldots < t_{m+1}^* \qquad \text{(нижнее)}, \tag{5.6}$$

$$a = s_1^* < s_2^* < \ldots < s_{m+1}^* < s_{m+2}^* = b \qquad \text{(верхнее)}, \tag{5.7}$$

и $s_i^* < t_i^* < s_{i+1}^*$, $i = 1, 2, \ldots, m+1$.

Рассмотрим интервалы

$$K_i = (s_i^*, t_i^*), \qquad i = 1, 2, \ldots, m+1,$$

$$J_i = (t_i^*, s_{i+1}^*), \qquad i = 1, 2, \ldots, m+1.$$

Для каждого $\xi \in K_i$ каноническая мера $\sigma_\xi$ имеет корни $t_i(\xi) \in K_i$, $i = 1, 2, \ldots, m+1$ и $t_{m+2}(\xi) = b$. Корни пробегают непрерывно и монотонно эти интервалы. Более того, веса $\lambda_{m+2}$ в $t_{m+2}(\xi) = b$ в представлении $c^0$ изменяются непрерывно от $\rho(b)$ до нуля. Аналогично для $\xi \in J_1$ каноническая мера $\sigma_\xi$ имеет корни $t_0(\xi) = a$.

$t_1(\xi) = \xi$ и $t_i(\xi) \in J_i$ $(i = 2, \dots, m+2)$, которые изменяются монотонно и непрерывно на этих интервалах. Более того, вес $\lambda_0$ в $t_0(\xi) = a$ увеличивается непрерывно от нуля до $\rho(a)$. Заметим, что $K$-интервалы всегда связаны каноническим представлением с точкой $b$.

## § 6. Иной подход к каноническим мерам и одномерным сечениям

Все изложение в §§ 2—5 следует Крейну [1951] и Карлину и Шепли [1953] с некоторым изменением порядка и отдельными усовершенствованиями. Общее определение важного понятия канонических мер дается следуя Крейну [1951]. Разумеется, канонические меры, связанные с моментной точкой, родственны в классическом случае понятию «квазиортогональных многочленов» (см. гл. IV), которые играют важную роль при изучении проблемы моментов. Характеризация максимальной массы в § 4 следует Крейну. То же справедливо для обсуждения свойств непрерывности канонических мер. То, что непрерывность свойственна мерам, следует ожидать ввиду единственности характеризации канонических мер. Теорема 5.1, по-видимому, является новой.

В этом параграфе дается геометрическая интерпретация канонических мер, причем изложение близко к изложению Карлина и Шепли [1953].

Определение 6.1. Каноническое представление назовем *верхним*, если его корни принадлежат $\{K_i\}$. Интервалы $K_i$ и $J_i$ будем называть *верхними и нижними каноническими интервалами* соответственно. Это совпадает с нашей предшествующей терминологией (см. определение 3.2).

Мы укажем другой метод конструирования канонических мер, соответствующий случаю, когда $T$-система $\{u_i\}_0^n$ может быть расширена до $T$-системы из $n+2$ функций

$$u_0(t), u_1(t), \dots, u_n(t), u_{n+1}(t) = \Omega(t). \tag{6.1}$$

Моментное пространство, введенное (6.1), обозначим через $\mathcal{M}_{n+1}$.

Пусть $\mathbf{c}^0 = (c_0^0, c_1^0, \dots, c_n^0)$ принадлежит $\operatorname{Int} \mathcal{M}_{n+1}$.

Определение 6.2. Интервал $R(\mathbf{c}^0)$ есть множество величин

$$\int_a^b \Omega(t)\, d\sigma(t) \qquad (\sigma \in V(\mathbf{c}^0)), \tag{6.2}$$

образованное в результате изменения $\sigma$ в множестве всех мер, представляющих $\mathbf{c}^0$. Этот интервал будем называть *одномерным сечением.*

Нетрудно доказать, что $R(\mathbf{c}^0)$ ограничено. Мы представим интервал в форме

$$R(\mathbf{c}^0) = \{\gamma\colon\ \underline{\gamma} \leqslant \gamma \leqslant \overline{\gamma}\}. \tag{6.3}$$

Очевидно, $\bar{\mathbf{c}} = \{c_0^0, c_1^0, \ldots, c_n^0, \bar{\gamma}\}$ и $\underline{\mathbf{c}} = \{c_0^0, c_1^0, \ldots, c_n^0, \underline{\gamma}\}$ являются граничными точками и, следовательно, каждая допускает единственное представление с индексом, меньшим чем $(n+1)/2$ (согласно теореме 2.1). Индекс каждой точки есть $(n+1)/2$, ибо в противном случае предположение, что $\mathbf{c}^0$ находится в $\operatorname{Int} \mathcal{M}_{n+1}$, привело бы к противоречию. Так как $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}$, то она имеет в точности два различных представления с индексом $(n+1)/2$ так, что интервал (6.3) не может выродиться в точку. Поэтому $\bar{\mathbf{c}}$ отлично от $\underline{\mathbf{c}}$. Представление $\bar{\mathbf{c}}$ и $\underline{\mathbf{c}}$ приводит, очевидно, к двум различным главным представлениям, связанным с $\bar{\mathbf{c}}^0$. Мы докажем, что $\bar{\mathbf{c}}$ соответствует верхнему ($\underline{\mathbf{c}}$ — нижнему), т. е. представление $\bar{\mathbf{c}}$ включает точку $b$. В этом случае пусть $\{s_j^*, \alpha_j^*\}_{j=1}^p$, и $\{t_j^*, \lambda_j^*\}_{j=1}^q$ означает корни и веса, входящие в представление $\bar{\mathbf{c}}$ и $\underline{\mathbf{c}}$ соответственно. Так как представления $\bar{\mathbf{c}}$ и $\underline{\mathbf{c}}$ являются главными для $\bar{\mathbf{c}}^0$, то мы знаем (следствие 3.1), что множества $\{t_j^*\}$ и $\{s_j^*\}$ разделены так, что их объединение содержит в точности $n+2$ точки. Более того, $b$ входит в одно из этих представлений.

Теперь мы имеем

$$\begin{gathered}0 = c_i^0 - c_i^0 = \sum_{j=1}^{p} \alpha_j^* u_i(s_j^*) - \sum_{j=1}^{q} \lambda_j^* u_i(t_j^*),\\ 0 < \bar{\gamma} - \underline{\gamma} = \sum_{j=1}^{p} \alpha_j^* \Omega(s_j^*) - \sum_{j=1}^{q} \lambda_j^* \Omega(t_j^*),\\ i = 0, 1, \ldots, n.\end{gathered} \tag{6.4}$$

Мы можем записать (6.4) в форме

$$\begin{gathered}0 = \sum_{k=1}^{n+2} a_k u_i(r_k), \qquad i = 0, 1, \ldots, n,\\ 0 < \bar{\gamma} - \underline{\gamma} = \sum_{k=1}^{n+2} a_k \Omega(r_k),\end{gathered} \tag{6.5}$$

где $\{r_k\}$ обозначает объединение $\{t_j^*\} \cup \{s_j^*\}$ точек, расположенных в порядке возрастания, и $a_k$ являются соответствующими коэффициентами. Решение (6.5) дает

$$a_{n+2} = (\bar{\gamma} - \underline{\gamma}) \frac{U\begin{pmatrix}0, & 1, & \ldots, & n\\ r_1, & r_2, & \ldots, & r_{n+1}\end{pmatrix}}{U\begin{pmatrix}0, & 1, & \ldots, & n+1\\ r_1, & r_2, & \ldots, & r_{n+2}\end{pmatrix}} > 0.$$

Это неравенство имеет смысл, только если $a_{n+2} = \alpha_p^*$, и поэтому $r_{n+2} = s_p^* = b$, как и ожидалось.

Название главных представлений верхним и нижним основано на геометрическом факте, что верхнее связано с $\bar{\mathbf{c}}$, а нижнее — с $\underline{\mathbf{c}}$, когда на самом деле $\bar{\gamma} > \underline{\gamma}$.

Каноническое представление $\mathbf{c}^0$ связано с (6.3) следующим образом. Для каждого $\gamma$ ($\underline{\gamma} < \gamma < \bar{\gamma}$) мы рассмотрим

$$\mathbf{c}(\gamma) = (c_0^0, c_1^0, \dots, c_n^0, \gamma) \in \operatorname{Int} \mathcal{M}_{n+2}. \tag{6.6}$$

При этом в $\mathcal{M}_{n+2}$ существует два главных представления $\mathbf{c}(\gamma)$, т. е. представления с индексом $(n+2)/2$. Они, разумеется,— различные канонические представления $\mathbf{c}^0$, одно из которых является верхним, а другое нижним (см. определение 6.1).

Мы легко вывели, что когда $\gamma$ пересекает $R(\mathbf{c}^0)$, соответствующая пара главных представлений, относящихся к $\mathcal{M}_{n+2}$, покрывает набор всех канонических представлений $\mathbf{c}^0$ непрерывным образом.

Проведенный выше анализ позволяет параметризовать интервал $R(\mathbf{c}^0)$ следующим образом.

Теорема 6.1 (Крейн [1951], Карлин и Шепли [1953]). *Пусть* $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}$, *и пусть* $I$ *является замыканием одного из верхних или нижних канонических интервалов. Тогда каноническое представление* $\sigma_t$ (*главное, если* $t = a$ *или* $b$) *индуцирует непрерывное взаимно однозначное отображение* $\gamma(t)$ *интервала* $I$ *на интервал* $R(\mathbf{c}^0)$, *определяемое формулой*

$$\gamma(t) = \int_a^b u_{n+1}(x)\, d\sigma_t(x).$$

*Функция* $\gamma(t)$ *возрастает на* $I = \bar{J}_i$ *и убывает на* $I = \bar{K}_i$.

## § 7. Двумерные сечения

В этом параграфе мы полагаем, что $\{u_i\}_0^{n+1}$ является такой системой функций на интервале $[a, b]$, что $\{u_i\}_0^k$ есть $T$-система для $k = n, n+1, n+2$. Для фиксированной моментной точки $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}$ мы рассмотрим множество

$$\Gamma(\mathbf{c}^0) = \left\{ (c_{n+1}, c_{n+2}) \mid c_i = \int_a^b u_i d\sigma, \quad i = n+1,\ n+2, \quad \sigma \in V(\mathbf{c}^0) \right\}.$$

Множество $\Gamma(\mathbf{c}^0)$ будем называть двумерным сечением. Аналогично тому, как это сделано в теореме 6.1, рассмотрим, в какой связи находится множество $\Gamma(\mathbf{c}^0)$ с теми мерами $\sigma \in V(\mathbf{c}^0)$, которые имеют индекс $(n+3)/2$.

Пусть $t_1^0$ и $t_2^0$ $(t_1^0 < t_2^0)$ являются двумя последовательными корнями верхнего главного представления $\bar{\sigma}$ точки $\mathbf{c}^0$ и для каждого $t' \in (t_1^0, t_2^0)$ $\rho(t')$ является массой в $t''$ канонической меры $\sigma_t$ точки $\mathbf{c}^0$. Определим двумерное множество

$$B = \{(t', \lambda) \mid t' \in (t_1^0, t_2^0), \quad 0 < \lambda < \rho(t')\}.$$

Теорема 7.1. *Для каждого* $(t', \lambda) \in B$ *существует единственное представление* $\sigma(t; t', \lambda)$ *для* $\mathbf{c}^0$, *имеющее индекс* $(n+3)/2$ *и включающее конечную точку* $b$ *и точку* $t'$ *с весом* $\lambda$. *Отображение* $\varphi$: $B \to \Gamma(\mathbf{c}^0)$, *определяемое соотношением*

$$\varphi(t', \lambda) = (c_{n+1}, c_{n+2}),$$

*где* $c_k = \int_a^b u_k(t)\, d\sigma(t; t', \lambda)$, $k = n+1,\ n+2$, *взаимно однозначно, непрерывно и покрывает внутренность области* $\Gamma(\mathbf{c}^0)$.

Замечание 7.1. Можно ожидать, что возможно следующее обобщение теоремы 7.1 на сечения более высокого порядка. Предположим, что функции $u_0, u_1, \ldots, u_n, u_{n+1}, \ldots, u_{n+m}$ таковы, что каждая система $\{u_i\}_0^k$ для $k = n,\ n+1, \ldots, n+m$ является $T$-системой. Для фиксированной $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}$ положим

$$\Gamma_m(\mathbf{c}^0) = \left\{(c_{n+1}, \ldots, c_{n+m}) \mid c_i = \int_a^b u_i\, d\sigma,\right.$$
$$\left. i = n+1, \ldots, n+m, \quad \sigma \in V(\mathbf{c}^0)\right\}.$$

Теперь естественно задать вопрос, может ли множество $\Gamma_m(\mathbf{c}^0)$ быть параметризовано при использовании мер $\sigma \in V(\mathbf{c}^0)$ индекса $(m+n+1)/2$. Это справедливо, но оказывается, что соответствующее множество $B$ очень трудно описать в обозримой форме.

Доказательство теоремы 7.1. Для каждого $\lambda \in (0, \rho(t'))$ точка

$$(c_0^*, \ldots, c_n^*) = (c_0^0, \ldots, c_n^0) - \lambda(u_0(t'), \ldots, u_n(t')) \tag{7.1}$$

лежит во внутренности $\mathcal{M}_{n+1}$ и, следовательно, принадлежит верхнему главному представлению, включающему точку $b$. Это представление не может включать $t'$, так как в противном случае при переносе члена $\lambda(u_0(t'), \ldots, u_n(t'))$ в левую часть (7.1) получающееся представление $(c_0^0, \ldots, c_n^0)$ имеет индекс $(n+1)/2$, включает точки $b$ и $t'$ и отличается от $\bar{\sigma}$. Это невозможно в силу следствия 3.2.

Перенося второй член, получаем представление $(c_0^0, \ldots, c_n^0)$ индекса $(n+3)/2$, включающее конечную точку $b$ и точку $t'$ с весом $\lambda \in (0, \rho(t'))$. Мы обозначим соответствующую меру через $\sigma(t; t', \lambda)$.

Существует только одно такое представление, ибо если бы было два, мы могли бы получить два верхних главных представления $(c_0^*, \dots, c_n^*)$ из (7.1).

Для окончания доказательства мы ограничимся случаем $n=2m+1$. Случай четного $n$ может быть рассмотрен аналогично. Полагая, что $a = s_1^* < s_2^* < \dots < s_{m+1}^* < s_{m+2}^* = b$ являются корнями $\bar{\sigma}$, как это полагалось в (5.7), видим, что любое представление $\mathbf{c}^0$ индекса $(n+3)/2$, содержащее конечную точку $b$, необходимо имеет корни $a = s_1 \leq s_2 < \dots < s_{m+2} < s_{m+3} = b$, которые перемежаются с корнями $\bar{\sigma}$, т. е. имеем

$$a < s_2 < s_2^* < s_3 < \dots < s_{m+1}^* < s_{m+2} < b. \tag{7.2}$$

Это следует из леммы 3.1.

Отображение $\varphi$, следовательно, является «отображением на», ибо любая точка $(c_0^0, \dots, c_{n+2}^0) \in \operatorname{Int} \mathcal{M}_{n+3}$ имеет верхнее главное представление $\sigma'$ индекса $(n+3)/2$, содержащее единственный корень $t'$ из интервала $(t_1^0, t_2^0)$. Соответствующий вес $\lambda$ должен быть меньше чем $\rho(t)$ ввиду характеризации $\rho(t')$ как максимальной массы, принадлежащей каноническому представлению $\sigma_{t'}$ (см. § 4). Так как представление индекса $(n+3)/2$ с весом $\lambda$ в $t'$ и весом в $b$ единственно, как было показано, представление $\sigma'$ должно совпадать с $\sigma(t; t', \lambda)$.

Отображение $\varphi$ — взаимно однозначное; это следует из того, что при $\varphi(t', \lambda) = \varphi(t'', \lambda) = (c_{n+1}, c_{n+2})$ с необходимостью $\sigma(t; t', \lambda) = \sigma(t; t'', \lambda)$, так как $(c_0^0, \dots, c_n^0, c_{n+1}, c_{n+2})$ имеет единственное верхнее главное представление. Вывод этот, однако, несостоятелен, так как любое представление точки $(c_0, \dots, c_n)$ индекса $(n+3)/2$, имеющее вес в точке $b$, имеет единственный корень в интервале $(t_1^0, t_2^0)$.

Непрерывность $\varphi$ доказывается следующим образом. Если $(t^{(n)}, \lambda^{(n)}) \to (t', \lambda)$, то требуется показать, что $\sigma(t; t^{(n)}, \lambda^{(n)})$ сходится к $\sigma(t; t', \lambda)$. В этом случае мы выбираем подпоследовательность $\sigma(t; t^{(n)}, \lambda^{(n)})$, сходящуюся к некоторой мере $\sigma$. Из свойства перемежаемости (7.2) следует, что мера $\sigma$ имеет вес $\lambda$ в $t'$ и имеет индекс $(n+3)/2$. В противном случае точка $c^* = (c_0^*, \dots, c_n^*)$ в (7.1) должна была бы иметь индекс $n/2$, что невозможно согласно теореме 2.1, так как $c^*$ является внутренней точкой $\mathcal{M}_{n+1}$. Следовательно, из свойства единственности, доказанного ранее, мы заключаем, что $\sigma(t) = \sigma(t; t', \lambda)$. Таким образом, каждая подпоследовательность $\sigma(t; t^{(n)}, \lambda^{(n)})$ сходится к $\sigma(t; t', \lambda)$, что и требовалось доказать.

Отметим теперь некоторые дальнейшие факты, вытекающие из приведенного выше анализа. Большинство из утверждений получается непосредственно и не потребует детальных доказательств.

Отображение $\varphi$ непрерывно на границе $\Gamma(\mathbf{c}^0)$ в том смысле, что, когда $\lambda \downarrow 0$, веса и корни $\sigma(t; t', \lambda)$ стремятся к весам и корням представления $\bar{\sigma}$ точки $c^0$; и в аналогичном смысле, когда $\lambda \uparrow \rho(t')$ мера $\sigma(t; t', \lambda)$ стремится к канонической мере $\sigma_{t'}$.

Область $\Gamma(c^0)$ замкнута, ограничена и выпукла. Более того, когда $t'$ пробегает интервал $[t_1^0, t_2^0]$, канонические меры $\sigma_{t'}$ отображают границу $\Gamma(c^0)$ начиная с точки, индуцированной $\sigma'$, и продолжая в направлении часовой стрелки. Для фиксированной $t' \in (t_1^0, t_2^0)$, когда вес $\lambda$ изменяется от $\rho(t')$ до нуля, точка $(c_{n+1}, c_{n+2})$ движется вдоль дуги во внутренности $\Gamma(c^0)$, которая соединяет соответствующую граничную точку и точку, индуцированную $\bar{\sigma}$.

Приведенные выше доводы, в которых использовалось главное верхнее представление, могут быть повторены с использованием нижнего представления. Здесь рассматривается $t'$, изменяющееся в интервале между двумя последовательными корнями в нижнем главном представлении $\underline{\sigma}$ точки $c^0$. В конструкции $\sigma(t; t', \lambda)$ используется нижнее главное представление точки $(c_0^*, \dots, c_n^*)$, определенное формулой (7.1).

## § 8. Монотонность корней некоторых главных представлений

Пусть $\{u_i\}_0^n$ является $T$-системой на интервале $[a, b]$, $w(t, \tau)$ — весовая функция, которая зависит от параметра $\tau$ и определяет $c(\tau) = (c_0(\tau), \dots, c_n(\tau))$ по формулам $c_i(\tau) = \int_a^b u_i(\tau)\, w(t, \tau)\, dt$. Всюду в этом параграфе будем предполагать, что $w(t, \tau) > 0$ для $t \in (a, b)$, параметр $\tau$ принимает значения из интервала $(\tau_1, \tau_2)$ и $c_i(\tau)$, $i = 1, \dots, n$, конечны и непрерывны для $\tau \in (\tau_1, \tau_2)$. Следующие теоремы обобщают результат Маркова [1886] относительно монотонного изменения нулей многочленов, ортогональных с весом $w(t, \tau)$.

Теорема 8.1. *Если $\{u_i\}_0^n$ является $T$-системой на $[a, b]$ и $w(t, \tau)$ есть $TP_2$* (см. определение 2.2 гл. I) *для $t \in (a, b)$ и $\tau \in (\tau_1, \tau_2)$, то внутренние корни верхнего и нижнего главных представлений $c(\tau)$ являются неубывающими непрерывными функциями $\tau$ на $(\tau_1, \tau_2)$.*

Доказательство. Проведем доказательство только для нижнего главного представления в случае нечетного $n = 2m - 1$, так как остальные случаи рассматриваются в основном аналогично.

Пусть

$$\Xi = \{(t_1, \dots, t_m, \lambda_1, \dots, \lambda_m) \mid a < t_1 < \dots < t_m < b,\ \lambda_i > 0,$$
$$i = 1, 2, \dots, m\}.$$

Рассмотрим отображение $f$ из $\Xi$ в $\mathcal{M}_{n+1}$, определяемое следующим образом: $c_k = \sum_{i=1}^m u_k(t_i)\lambda_i$, $k = 0, 1, \dots, n$. Из предшествующих рассмотрений нам известно, что $f$ отображает $\Xi$ взаимно однозначно и непрерывно на внутренность $\mathcal{M}_{n+1}$. Обратная функция $f^{-1}$, отображающая $\operatorname{Int}\mathcal{M}_{n+1}$ на $\Xi$ является также непрерывной. Доказательство состоит в следующем. Пусть $\lim\limits_{k\to\infty} \mathbf{c}^{(k)}=\mathbf{c}$, где $\mathbf{c}$ и $\mathbf{c}^{(k)}$ для $k=1, 2,\ldots$ содержатся в $\operatorname{Int}\mathcal{M}_{n+1}$. Тогда, если $f^{-1}(\mathbf{c}^{(k)}) = (t_1^{(k)}, \ldots, t_m^{(k)}, \lambda_1^{(k)}, \ldots \ldots, \lambda_m^{(k)})$, то сходимость и существование строго положительного полинома предполагает (как в (1.2)), что $\lambda_i^{(k)}$ ограничены по $k$. Непрерывность $f^{-1}$ теперь легко следует из того факта, что $\mathbf{c}$ принадлежит $\operatorname{Int}\mathcal{M}_{n+1}$ и обладает единственным нижним главным представлением.

Непрерывность $f^{-1}$ доказывает, в частности, так как $c_j(\tau)$ непрерывны по $\tau$, для $\tau \in (\tau_1, \tau_2)$ и $\mathbf{c}(\tau) \in \operatorname{Int}\mathcal{M}_{n+1}$, что корни $t_i(\tau)$, $i = 1, \ldots, m$, в нижнем главном представлении $\mathbf{c}(\tau)$ являются непрерывными функциями $\tau$. Остается доказать, что $t_i(\tau)$ — неубывающие функции.

Предположим противное, что $t_{j_0}(\tau') > t_{j_0}(\tau'')$ для некоторого $j$ и $\tau_1 < \tau' < \tau < \tau_2$. В силу непрерывности $t_i(\tau)$ мы можем считать без потери общности, что $t_i(\tau'') > t_{j_0}(\tau')$ для $i > j_0$, и $t_i(\tau') < t_{j_0}(\tau'')$ для $i < j_0$. Ссылаясь на теорему 5.2 гл. I, определим $u = \sum\limits_{i=0}^{n} a_i u_i$ как многочлен, имеющий корни второй кратности в $t_0 = 2^{-1}(t_{j_0}(\tau') + t_{j_0}(\tau''))$, первой кратности в $t_i(\tau')$ для $i < j_0$ и $t_i(\tau'')$ для $i > j_0$. Отметим также, что многочлен $u$ не обращается в нуль больше нигде на открытом интервале $(a, b)$ и что он изменяется от отрицательных значений к положительным в точке $t_0$.

Отсюда следует, что функция

$$h(\tau) = \int u(t)\, w(t, \tau)\, dt = \sum_{i=1}^{m} u(t_i(\tau))\, \lambda_i(\tau) \tag{8.1}$$

является отрицательной в $\tau''$ и положительной в $\tau'$. Определим далее

$$\varphi_1(t) = \begin{cases} 0, & t < t_0, \\ u(t), & t \geqslant t_0, \end{cases}$$

$$\varphi_2 = \varphi_1(t) - u(t),$$

$$h_i(\tau) = \int \varphi_i(t)\, w(t, \tau)\, dt, \qquad i = 1, 2.$$

Тогда

$$h_1(\tau'')\, h_2(\tau') - h_2(\tau'')\, h_1(\tau') =$$

$$= \int\limits_{s=t_0}^{b} \int\limits_{t=a}^{t_0} \varphi_1(s)\, \varphi_2(t) \begin{vmatrix} w(t, \tau') & w(t, \tau'') \\ w(s, \tau') & w(s, \tau'') \end{vmatrix} dt\, ds. \tag{8.2}$$

Так как $h(\tau'') < 0$, $h(\tau') > 0$, $h(\tau) = h_1(\tau) - h_2(\tau)$ и $h_i(\tau) \geqslant 0$, $i = 1, 2$, то заключаем, что $0 \leqslant h_1(\tau'') < h_2(\tau'')$ и $0 \leqslant h_2(\tau') < h_1(\tau')$.

Отсюда следует, что левая часть (8.2) строго отрицательна. Однако правая часть очевидным образом неотрицательна, ибо $\varphi_1(t) \geqslant 0$ на $[t_0, b]$, $\varphi_2(t) \geqslant 0$ на $[a, t_0]$ и $w$ есть $TP_2$, что означает положительность определителя в подынтегральном выражении. Мы пришли к противоречию, которое завершает доказательство.

Более жесткие предположения относительно весовой функции $w(t, \tau)$ приводят к следующим результатам.

Теорема 8.2. *Если* $\{u_i\}_0^n$ *является* $T$*-системой на* $[a, b]$ *и* $w(t, \tau)$ *есть* $STP_2$ (см. определение 2.2 гл. I), *то внутренние корни верхнего и нижнего главных представлений* $c(\tau)$ *являются строго возрастающими функциями* $\tau \in (\tau_1, \tau_2)$

Доказательство. Мы снова ограничимся случаем нижнего представления для $n = 2m - 1$ и заметим, что утверждение о непрерывности $t_i(\tau)$, $i = 1, \ldots, m$, остается справедливым без изменения доказательства. Далее, если для некоторого $\tau' < \tau''$ и для некоторого $j$ мы имеем $t_j(\tau'') < t_j(\tau')$ и $u$ определяется так же, как и выше, то можно заключить, что функция $h(\tau)$ из уравнения (8.1) такова, что $h(\tau') \geqslant 0$ и $h(\tau'') \leq 0$. Это сразу же приводит к заключению, что левая часть (8.2) неположительна. С другой стороны, $STP_2$-свойство $w(t, \tau)$ предполагает, что правая часть (8.2) строго положительна, что приводит к противоречию, как и при доказательстве теоремы 8.1. Поэтому мы должны иметь $t_j(\tau') < t_j(\tau'')$ для всех $j$ при $\tau' < \tau''$.

Если наложить более строгое требование, что $\{u_i\}_0^n$ является $ET$-системой порядка 2 и что

$$\frac{w_\tau(t, \tau)}{w(t, \tau)} \qquad \left(w_\tau(t, \tau) = \frac{\partial}{\partial\tau} w(t, \tau)\right)$$

— возрастающие функции $t$, то мы достигнем следующего усиления результатов теоремы 8.2.

Теорема 8.3. *Если*

(i) $\{u_i\}_0^n$ *является* $ET$*-системой порядка* 2 (см. определение 2.4 гл. I),

(ii) $w(t, \tau)$ *непрерывно дифференцируема по* $\tau$,

(iii) *производная* $\frac{\partial}{\partial\tau} c_i(\tau)$ *задается равенством*

$$\frac{\partial}{\partial\tau} c_i(\tau) = \int_a^b u_i(t)\, w_\tau(t, \tau)\, dt,$$

*и непрерывна по* $\tau$,

(iv) *для каждого* $\tau \in (\tau_1, \tau_2)$, $w_\tau(t, \tau)/w(t, \tau)$ *является неубывающей функцией* $t$ *и не есть тождественная константа,*

*то внутренние корни верхнего и нижнего главных представлений имеют непрерывные производные, которые строго положительны для* $\tau \in (\tau_1, \tau_2)$.

Доказательство. Рассмотрим только нижние главные представления для нечетного случая $n = 2m - 1$. Пусть $f$ является функцией, определенной при доказательстве теоремы 8.1, которая отображает $\Xi$ в $\operatorname{Int} \mathcal{M}_{n+1}$. Мы утверждаем, так как $u \in C^1$, что частные производные первого порядка $f$ непрерывны. Действительно, якобиан преобразования $f$ легко вычисляется и есть

$$(-1)^{[m(m+1)]/2} \lambda_1 \dots \lambda_m U^* \begin{pmatrix} 0, 1, \dots, 2m-2, 2m-1 \\ t_1, t_2, \dots, t_m, \quad t_m \end{pmatrix}.$$

Это выражение отлично от нуля вследствие $ET$-свойства системы $\{u_i\}_0^n$. Используя теорему об обратном преобразовании (см. Апостол [1957], стр. 144), видим, что $f^{-1} = g$ непрерывна и имеет непрерывные частные производные.

Так как $(\partial/\partial\tau) c_i(\tau)$ непрерывна по $\tau$, то $(\partial/\partial\tau) t_i(c(\tau))$ и $(\partial/\partial\tau) \lambda_i(\tau)$ существуют и непрерывны для $\tau \in (\tau_1, \tau_2)$.

Пусть $u = \sum_{i=0}^{n} a_i u_i$ — произвольный многочлен. Имеем

$$\int u(t) w(t, \tau) \, dt = \sum_{j=1}^{m} u(t_j(\tau)) \lambda_j(\tau),$$

$$\int u(t) w_\tau(t, \tau) \, dt = \sum_{j=1}^{m} u'(t_j(\tau)) \lambda_j(\tau) + \sum_{j=1}^{m} u(t_j(\tau)) \lambda'_j(\tau).$$

Выбирая теперь в качестве $u$ нетривиальный многочлен, имеющий простые нули в $t_{j_0}(\tau)$, двойные нули в $t_i(\tau)$ для $i \neq j_0$ и неотрицательный для $t \leqslant t_{j_0}(\tau)$, получаем

$$\int u(t) w_\tau(t, \tau) \, dt = u'(t_{j_0}(\tau)) \, t'_{j_0}(\tau) \lambda_{j_0}(\tau).$$

Левую часть этого равенства можно переписать в виде выражения

$$\int u(t) \left( w_\tau(t, \tau) - \frac{w_\tau(t_{j_0}(\tau), \tau)}{w(t_{j_0}(\tau), \tau)} w(t, \tau) \right) dt,$$

которое является строго отрицательным при наших предположениях относительно $w_\tau(t, \tau)/w(t, \tau)$. Поэтому $u'(t_{j_0}(\tau)) \, t'_{j_0}(\tau) \lambda_{j_0}(\tau) < 0$ и, так как $u'(t_{j_0}(\tau))$ должно быть меньше нуля вследствие $ET$-свойства $\{u_i\}_0^n$ и свойств $u$, то $t'_{j_0}(t) > 0$.

Следствие 8.3. *Если $\{u_i\}_0^n$ является $ET$-системой порядка 2 и $w(t, \tau'')/w(t, \tau')$ не убывает и отлично от константы, какими бы ни были $\tau' < \tau$, то корни верхнего и нижнего главных представлений* $c(\tau)$ *строго возрастают по* $\tau$.

Доказательство. При любых $\tau' < \tau''$ для функции $w(t, \tau) = (1 - \tau) w(t, \tau') + \tau w(t, \tau'')$ выполняется равенство

$$\frac{w_\tau(t, \tau)}{w(t, \tau)} = \tau^{-1} - \tau^{-1} \left[ (1 - \tau) + \tau \frac{w(t, \tau'')}{w(t, \tau')} \right]^{-1}$$

так, что $w(t, \tau)$ для $\tau \in (0, 1)$ удовлетворяет условиям теоремы 8.3. Следствие доказано.

В частном случае ортогональных многочленов Марков [1886] (см. также Сеге [1959]) использовал результаты теоремы 8.3, чтобы установить монотонное изменение корней многочленов Якоби $P_n^{(\alpha,\beta)}(t)$ при изменяющихся параметрах $\alpha$ и $\beta$.

## § 9. Двойственный конус $\mathcal{P}_{n+1}$

Для данного выпуклого конуса $\mathcal{C}$, содержащегося в $n+1$-мерном евклидовом пространстве, двойственный конус $\mathcal{C}^+$ определяется следующим образом:

$$\mathcal{C}^+ = \{\mathbf{a} \in R^{n+1} \mid (\mathbf{a}, \mathbf{c}) > 0 \quad \text{для всех} \quad \mathbf{c} \in \mathcal{C}\},$$

где $(\mathbf{a}, \mathbf{c}) = \sum_{i=0}^{n} a_i c_i$. Непосредственными рассуждениями легко проверить, что $\mathcal{C}^+$ является замкнутым и выпуклым. Для полноты мы включили здесь следующую элементарную характеризацию двойственного выпуклого конуса.

Лемма 9.1. *Если $\mathcal{C}$ — выпуклый конус, то $\overline{\mathcal{C}} = \mathcal{C}^{++}$, где $\overline{\mathcal{C}}$ означает замыкание $\mathcal{C}$.*

Доказательство. Из определения следует, что если $\mathbf{c} \subset \mathcal{C}$, то $(\mathbf{a}, \mathbf{c}) \geqslant 0$ для всех $\mathbf{a} \in \mathcal{C}^+$ и $\mathbf{c} \in \mathcal{C}^{++}$. Следовательно, $\mathcal{C} \subset \mathcal{C}^{++}$. Замыкание $\overline{\mathcal{C}}$ является замкнутым выпуклым конусом и, так как $\mathcal{C}^{++}$ замкнуто, то $\overline{\mathcal{C}} \subset \mathcal{C}^{++}$. Теперь, если $\mathbf{y} \notin \overline{\mathcal{C}}$, то существует вектор $\mathbf{a} \in R^{n+1}$ такой, что $(\mathbf{y}, \mathbf{a}) < (\mathbf{c}, \mathbf{a})$ для всех $\mathbf{c} \in \overline{\mathcal{C}}$. Так как $\overline{\mathcal{C}}$ является конусом, то заключаем, что $(\mathbf{y}, \mathbf{a}) < 0 \leqslant (\mathbf{c}, \mathbf{a})$ для всех $\mathbf{c} \in \overline{\mathcal{C}}$. Следовательно, $\mathbf{a} \in \mathcal{C}^+$ и $\mathbf{y} \notin \mathcal{C}^{++}$. Доказательство завершено.

Двойственная теория для выпуклых конусов встречается в чрезвычайно многих областях. Для моментного пространства мы увидим ниже, что двойственный конус может быть отождествлен с пространством неотрицательных многочленов. Это геометрическое толкование позволяет дать интуитивное решение для различных экстремальных задач, связанных с многочленами (см., в частности, гл. IX, X).

Определение 9.1. Если $\{u_i\}_0^n$ является $T$-системой, то обозначим множество всех неотрицательных многочленов $u$ через $\mathcal{P}_{n+1}$.

Так как каждый многочлен единственным образом определяется вектором своих коэффициентов, мы можем отождествить $\mathcal{P}_{n+1}$ с множеством векторов $\mathbf{a} \in R^{n+1}$, для которых

$$\sum_{i=0}^{n} a_i u_i(t) \geqslant 0 \quad \text{при всех} \quad t \in [a, \beta].$$

Если $\mathbf{a} \in \mathcal{M}_{n+1}$, то $\sum_{i=0}^{n} a_i c_i \geqslant 0$ для всех $\mathbf{c} \in \mathcal{M}_{n+1}$. В частности, $u(t) = \sum_{i=0}^{n} a_i u_i(t) \geqslant 0$, $u \in \mathcal{P}_{n+1}$. Обратно, каждый многочлен $u \in \mathcal{P}_{n+1}$ определяет некую точку в $\mathcal{M}_{n+1}$. Следовательно, имеем $\mathcal{M}_{n+1}^{+} = \mathcal{P}_{n+1}$. Так как множество $\mathcal{M}_{n+1}$ замкнуто, то использование леммы 9.1 дает следующую теорему.

Теорема 9.1. *Если $\{u_i\}_0^n$ является $T$-системой, то $\mathcal{M}_{n+1}^{+} = \mathcal{P}_{n+1}$ и $\mathcal{P}_{n+1}^{+} = \mathcal{M}_{n+1}$.*

Следующая лемма понадобится в гл. IV.

Лемма 9.2. *Если $\{u_i\}_0^n$ является $T$-системой на $[a, b]$, то*

(i) $\mathbf{c}^0 = (c_0^0, \ldots, c_n^0) \in \mathcal{M}_{n+1}$ *тогда и только тогда, когда $\sum_{i=0}^{n} a_i c_i^0 \geqslant 0$ для любого многочлена $\sum_{i=0}^{n} a_i u_i \in \mathcal{P}_{n+1}$;*

(ii) $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}$, *если и только если $\sum_{i=0}^{n} a_i u_i \in \mathcal{P}_{n+1}$ для каждого нетривиального многочлена.*

Доказательство. Пункт (i) леммы является простой перефразировкой равенства $\mathcal{M}_{n+1} = \mathcal{P}_{n+1}^{+}$. Доказательство пункта (ii) проводится следующим образом. Если

$$\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1} \quad \text{и} \quad u(t) = \sum_{i=0}^{n} a_i u_i(t) \geqslant 0,$$

где $\sum_{i=0}^{n} a_i^2 > 0$, то $\sum_{i=0}^{n} a_i u_i(t_0) > 0$ для некоторого $t_0$ и можно проинтегрировать $u(t)$ на промежутке $[a, b]$ по отношению к канонической мере $\mathbf{c}^0$, имеющей положительную массу в $t_0$, чтобы получить $\sum_{i=0}^{n} a_i c_i^0 > 0$. Если $\mathbf{c}^0 \notin \operatorname{Int} \mathcal{M}_{n+1}$, то существует $(a_0, \ldots, a_n)$, $\sum_{i=0}^{n} a_i^2 > 0$, обладающий тем свойством, что $\sum_{i=0}^{n} a_i c_i^0 \leqslant 0$ и $\sum_{i=0}^{n} a_i c_i \geqslant 0$ для всех $\mathbf{c} \in \mathcal{M}_{n+1}$. В этом случае $\sum_{i=0}^{n} a_i u_i(t) \geqslant 0$, в то время как $\sum_{i=0}^{n} a_i c_i^0 \leqslant 0$.

Теорема 9.2. *Если $\{u_i\}_0^n$ является $T$-системой, то граница $\mathcal{P}_{n+1}$ состоит из всех многочленов $u \in \mathcal{P}_{n+1}$, которые обращаются в нуль где-либо на интервале $[a, b]$.*

Доказательство. Если многочлен $u \in \mathcal{P}_{n+1}$ обращается в нуль в $t_0 \in [a, b]$, то $(u_0(t_0), u_1(t_0), \ldots, u_n(t_0))$ определяет опорную плоскость к $\mathcal{P}_{n+1}$, содержащую $u$, так что $u$ находится на границе $\mathcal{P}_{n+1}$. Обратное утверждение является следствием свойства, состоящего в том, что если $u(t) > 0$ для всех $t \in [a, b]$, то $u$ принадлежит внутренности $\mathcal{P}_{n+1}$.

## § 10. Теоремы о представлении положительных многочленов

Для классических $T$-систем $u_i(t) = t^i$, $i = 0, 1, \ldots, n$ на $[a, b]$ теорема Карлина и Шепли [1953] утверждает, что каждый многочлен $\mathcal{P}(t)$ (степени $n$), который неотрицателен на $[a, b]$ и имеет самое большее $n - 1$ нулей в $[a, b]$, считая их кратность, допускает единственное представление

$$P(t) = \begin{cases} \alpha \prod_{j=1}^{m} (t - t_{2j-1})^2 + \beta (t - a)(b - t) \prod_{j=1}^{m-1} (t - t_{2j})^2, \\ \qquad\qquad\qquad\qquad\qquad\qquad n = 2m, \\ \alpha (t - a) \prod_{j=1}^{m} (t - t_{2j})^2 + \beta (b - t) \prod_{j=1}^{m} (t - t_{2j-1})^2, \\ \qquad\qquad\qquad\qquad\qquad\qquad n = 2m + 1, \end{cases} \tag{10.1}$$

где $\alpha > 0$, $\beta > 0$ и $a \leq t_1 \leq t_2 \leq \ldots \leq t_{n-1} \leq b$. Более того, $P(t)$ строго положителен на $[a, b]$ тогда и только тогда, когда $a < t_1 < \ldots < t_{n-1} < b$. Используя этот результат, мы можем утверждать, что для системы $\{t^i\}_0^n$, определенной на $[a, b]$, граничные лучи $\mathcal{P}_{n+1}$ образуются теми неотрицательными многочленами, которые имеют $n$ нулей на $[a, b]$, считая их кратность. Действительно, если $P(t)$ не имеет $n$ нулей, то представление (10.1) показывает, что $P(t)$ не является граничным. Обратно, любой многочлен с $n$ нулями должен быть граничным, так как, если $P$ имеет $n$ нулей, равенство $P(t) = P_1(t) + P_2(t)$ имеет место; когда все три многочлена неотрицательны, то $P_1$ и $P_2$ с необходимостью имеют общие нули с $P$. При этих условиях $P_1$ и $P_2$ не могут быть линейно независимыми.

Если мы изучим равенство (10.1), например для $n = 2m$, то мы увидим, что оба многочлена в правой части (10.1) имеют максимальное число нулей (т. е. $2m$), с учетом их кратностей и что два множества корней строго перемежаются, когда $P(t) > 0$. В этом случае оба многочлена проявляют максимально выраженные осцилляции между нулем и многочленом $P(t)$ в том смысле, что каждый из сомножителей обладает максимальным числом различных нулей, и между каждой парой различных нулей и граничными точками $a$ и $b$ они касаются многочлена $P(t)$ по крайней мере один раз.

Этот параграф будет посвящен построению аналога (10.1) для общих $T$-систем. Следствия, характеризующие граничные точки, содержатся в следующем параграфе.

Кроме своего внутреннего интереса, теоремы о представлении, рассматриваемые в этом параграфе, будут использоваться по существу при решении некоторых экстремальных задач, в которых

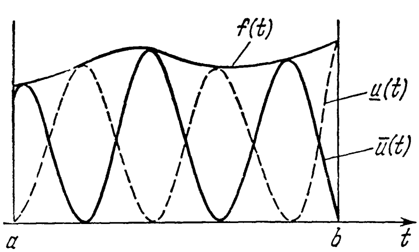


Рис. 5.

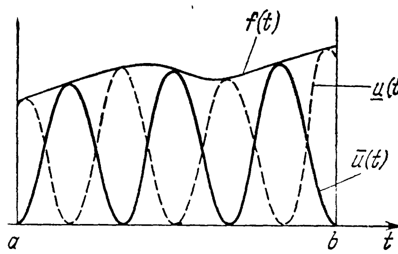


Рис 6.

обобщаются неравенства Бернштейна — Маркова. Результаты такого рода будут изложены в § 5 гл. IX.

Следующая основная теорема устанавливает для произвольной положительной непрерывной функции существование двух специальных многочленов, обладающих замечательными осцилляционными свойствами.

Теорема 10.1. *Пусть* $\{u_i\}_0^n$ *является* $T$-*системой на* $[a, b]$, *и пусть* $f(t)$ *является непрерывной и строго положительной на* $[a, b]$. *Тогда*:

a) *существует единственный многочлен* $\underline{u}(t)$, *удовлетворяющий условиям*:

(i) $f(t) \geqslant u(t) \geqslant 0$;

(ii) $u(t)$ *обращается в нуль на множестве индекса* $n/2$ (*т. е.* $\tilde{Z}(u) = n$);

(iii) $f(t) - u(t)$ *обращается в нуль по крайней мере один раз между каждой парой соседних нулей* $u(t)$ *и по крайней мере один раз между наибольшим нулем и граничной точкой* $b$ *и между наименьшим нулем и граничной точкой* $a$;

(iv) $u(b) > 0$;

(b) *существует единственный многочлен* $\underline{u}(t)$, *удовлетворяющий условиям* (i), (ii), (iii) *и* (iv) $u(b) = 0$.

Общие свойства $\underline{u}$ и $\bar{u}$ отражены на рис. 5 и 6 для $n = 5$ и $n=6$.

Доказательство. Мы разберем только случай $n = 2m$, так как случай $n = 2m + 1$ аналогичен.

Построение полинома $\underline{u}$ в части (a) производится следующим образом. Допустим сначала, что $\{u_i(t)\}_0^n$ является $ET$-системой, и построим для каждого $\xi$ в $m$-мерном симплексе $\Xi^m = \{\xi = (\xi_0, \ldots$

$\ldots, \xi_m) \mid \xi_i \geqslant 0, 1, \ldots, m, \ \sum_{i=0}^{m} \xi_i = b - a\}$ полином $u_\xi(t) = \sum_{i=0}^{n} a_i(\xi) u_i(t)$, который неотрицателен, нормирован посредством равенства $\sum_{i=0}^{n} a_i^2(\xi) = 1$ и обращается в нуль кратности 2 в каждой из $m$ точек:

$$t_i = a + \sum_{k=0}^{i-1} \xi_i, \qquad i = 1, \ldots, m.$$

(Если $p$ точек $t_i$ совпадают, то эта общая точка должна иметь кратность $2p$.) Полином $u_\xi(t)$ задается определителем

$$U^* \begin{pmatrix} t_1, t_1, \ldots, t_i, t_i, t_{i+1}, t_{i+1}, \ldots, t_m, & t_m, & t \\ 0, 1, \qquad \ldots, & & n-1, n \end{pmatrix},$$

умноженным на константу так, что сумма квадратов коэффициентов равна единице. Пусть $\delta_i(\xi) = \min\{\delta \mid \delta f(t) \geqslant u_\xi(t), \ t \in [t_i, t_{i+1}]\}$ для $i = 0, 1, \ldots, m$, где $t_0 = a$ и $t_{m+1} = b$. Тогда, так как

$$a_i(\xi) = \frac{(-1)^i U_i}{\left(\sum_{i=0}^{n} U_i^2\right)^{1/2}},$$

где

$$U_i \equiv U^* \begin{pmatrix} t_1, t_1, & \ldots, & t_m, t_m \\ 0, 1, \ldots, i-1, i+1, & \ldots & n \end{pmatrix},$$

то мы выводим, что функция $\delta_i(\xi)$ непрерывна по $\xi$ и $\delta_i(\xi) = 0$ тогда и только тогда, когда $\xi_i = 0$. Это справедливо, так как, если $\xi^k \to \xi$, то $a_i(\xi^k) \to a_i(\xi)$ для каждого $i$, так что $u_{\xi(k)}(t) \to u_\xi(t)$ равномерно на $[a, b]$. Далее определим

$$F_i(\xi) = \delta_i(\xi) - \min_k \delta_k(\xi), \qquad i = 0, \ldots m,$$

и множество $F_{m+1}(\xi) = F_0(\xi)$. Заметим, что если не существует такой точки $\xi$, что $F_i(\xi) = 0$ для всех $i$, то тогда $\sum_{i=0}^{m} F_i(\xi) > 0$ для всех $\xi$. В этом случае определено непрерывное отображение

$$\xi_i' = \frac{F_{i+1}(\xi)}{\sum_{k=0}^{m} F_k(\xi)} (b-a), \qquad i = 0, \ldots, m,$$

симплекса $\Xi^m$ в себя.

Применяя теорему Брауэра о неподвижной точке, получаем существование точки $\xi$, для которой

$$\xi_i^* = \frac{F_{i+1}(\xi^*)}{\sum_{k=0}^{m} F_k(\xi^*)} (b-a).$$

Но для какого-нибудь $\xi$ и для некоторого $i$ будет выполняться равенство $F_i(\xi) = 0$. Предположим, что $F_i(\xi^*) = 0$. Это допустимо, так как $\xi^*$ является фиксированной точкой, такой, что $F_{i-1}(\xi) = 0$. Из построения следует, что $F_{i-1}(\xi^*) = 0$. Используя это рассуждение еще несколько раз, мы выводим, что $F_i(\xi^*) = 0$ для всякого $i$, что противоречит предположению, что $\sum_{i=0}^{n} F_i(\xi^*) > 0$. Следовательно, существует по крайней мере одна точка $\xi^*$, такая, что $\delta_i(\xi^*) = \delta$ для $i = 0, \ldots, m$. Из $u_{(\xi)^*}(t) \not\equiv 0$ следует, что $\delta > 0$, откуда вытекает, что $t_i$, соответствующие $\xi^*$, являются различными, т. е. $a < t_1 < \ldots$ $\ldots < t_m < b$.

Полином $\underline{u}(t) = \delta^{-1} u_{\xi^*}(t)$ по своему построению удовлетворяет требованиям (i), (ii), (iii) и (iv).

Дальнейший анализ приведенного выше доказательства показывает, что мы можем задать второй полином $\bar{u}(t)$ в $t_0 = a$ и $t_m = b$ с нулями кратности один и $m-1$ других точек $t_1 \leq \ldots \leq t_{m-1}$ каждая с кратностью два. В этом случае результирующий полином $\bar{u}(t)$ удовлетворяет свойствам (i), (ii), (iii), (iv)$'$.

Рассмотрение случая, где $\{u_i\}_0^n$ является только $T$-системой и доказательство единственности будет проведено только в случае многочлена $\underline{u}$ в части (a). Если система $\{u_i\}_0^n$ является только $T$-системой, мы рассмотрим вспомогательную функцию

$$u_i(t; \sigma) = \int_a^b G_\sigma(t, x)\, u_i(x)\, dx, \qquad i = 0, 1, \ldots, n,$$

где

$$G_\sigma(t, x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{t-x}{\sigma}\right)^2\right].$$

С помощью основной композиционной формулы (3.13) гл. I и факта, что $G_\sigma(t; x)$ является $ETP(t)$ (см. также (3.16) гл. I), мы можем заключить, что система $\{u_i(t, \sigma)\}_0^n$ является $ET$-системой на $[a, b]$ и, следовательно, на всяком подынтервале $[a', b']$, где $a < a' < b' < b$. Необходимо ограничить систему $\{u_i(t, \sigma)\}_0^n$ на некоторый подынтервал $[a', b]$ вследствие того, что в конечных точках $t = a$ и $t = b$ мы имеем $\lim\limits_{\sigma \to 0+} u_i(t; \sigma) = 2^{-1} u_i(t)$, $i = 0, 1, \ldots, n$, хотя для $t \in (a, b)$

$\lim\limits_{\sigma\to 0+} u_i(t;\sigma) = u_i(t)$. Теперь для каждого $\sigma > 0$ произведем предшествующее построение многочлена $\underline{u}(t,\sigma)$, удовлетворяющего условиям (i), (ii), (iii) и (iv) на интервале $[a', b']$. Если

$$\underline{u}(t,\sigma) = \sum_{i=0}^{n} a_i(\sigma)\, u_i(t,\sigma),$$

то мы можем выбрать последовательность $\sigma_k \to 0$; пусть $t_1^{(k)}, \dots, t_m^{(k)}$ являются нулями $\underline{u}(t,\sigma_k)$. Дополнительно мы зададим $m+1$ точку $s_1^{(k)}, \dots, s_{m+1}^{(k)}$, каждая из которых лежит между точек $\{t_i^{(k)}\}_{i=1}^{m}$, т. е. $a' \leq s_1^{(k)} < t_1^{(k)} < \dots < t_m^{(k)} < s_{m+1}^{(k)} \leq b'$ и удовлетворяет соотношению

$$f(s_i^{(k)}) - \underline{u}(s_i^{(k)};\ \sigma_k) = 0, \qquad i = 1, 2, \dots, m+1.$$

Так как $f(t) \geqslant \underline{u}(t;\sigma) \geqslant 0$ на $[a', b']$, мы находим, выбрав произвольные различные точки $x_0, \dots, x_n$ на $[a', b']$ и решив систему уравнений

$$\underline{u}(x_j;\sigma) = \sum_{i=0}^{n} a_i(\sigma)\, u_i(x_j;\sigma), \qquad j = 0, \dots, n,$$

относительно $a_i(\sigma)$, $i = 0, \dots, n$, что эти величины равномерно ограничены. Выберем теперь подпоследовательность $\{\sigma_{k'}\}$ из $\{\sigma_k\}$, обладающую свойством, что при $k' \to \infty$ мы имеем

$$\begin{aligned} &a_i(\sigma_{k'}) \to a_i, && i = 0, \dots, n,\\ &s_j^{(k')} \to s_j, && j = 1, \dots, m+1,\\ &t_l^{(k')} \to t_l, && l = 1, \dots, m,\\ &a' \leq s_1 \leq t_1 \leq \dots \leq t_m \leq s_{m+1} \leq b'. \end{aligned}$$

Функция $\underline{u}(t'; a', b') = \sum\limits_{i=0}^{n} a_i u_i(t)$ обращается в нуль на точках $t_l$, $l = 1, \dots, m$, и равняется $f(t)$ на $s_j$, $j = 1, \dots, m+1$. Следовательно, так как $\underline{u}(t; a', b')$ является непрерывной, мы видим, что

$$a' \leq s_1 < t_1 < \dots < t_m < s_{m+1} \leq b',$$

и, следовательно, $\underline{u}(t; a', b')$ удовлетворяет свойствам (i) — (iv) на отрезке $[a, b]$.

Устремив $a' \downarrow a$ и $b' \downarrow b$, получим многочлен, удовлетворяющий условиям (i) — (iv) на всем $[a, b]$.

Вернемся к задаче, что $\underline{u}(t)$ является единственным многочленом, для которого выполняются условия (i) — (iv). Предположим сначала, что мы имеем другой такой же многочлен, скажем $u(t)$. Тогда он имеет $m$ внутренних нулей. Пусть $t_1 < t_2 < \ldots < t_m$ означают нули $\underline{u}(t)$. Без потери общности мы можем предположить, что первый нуль $t'$ функции $u(t)$ меньше $t$ или $t'_1 = t$ и $\underline{u}(t) = \bar{u}(t)$ неотрицательна на некотором интервале $(t - \varepsilon, t_1)$, в противном случае мы можем поменять ролями $\underline{u}(t)$ и $u(t)$. Мы скажем, что $g(t) = \underline{u}(t) - u(t)$ имеем нуль на замкнутом интервале $[c, d]$, если $g(t_0) = 0$ для $t_0 \in (c, d)$, или $g(c) = 0$ и $g(t) \geqslant 0$ для $t \in (c, c + \varepsilon)$, или $g(d) = 0$ и $g(t) \geqslant t$ для $t \in (d - \varepsilon, d)$. При учете нулей, таким образом, ясно, что $g(t)$ имеет по крайней мере два нуля в каждом из интервалов $[t_{i-1}, t_i]$, $i = 1, \ldots, m$, где $t_0 = a$ и по крайней мере один на интервале $[t_m, b]$. В итоге $g(t)$ обращается в нуль по крайней мере $n+1$ раз. Заметим, что каждый неузловой нуль $g(t)$ должен быть сосчитан дважды, так что мы имеем $\tilde{Z}(g) \geqslant n + 1$ и, следовательно, $g(t) = 0$ по теореме 4.2 гл. I.

Следствие 10.1 (а). *Если функция $f(t)$ в теореме 10.1 является полиномом $u(t)$, то тогда существует единственное представление*

$$u(t) = \underline{u}(t) + \bar{u}(t), \quad \bar{u}(t) \geqslant 0, \quad \underline{u}(t) \geqslant 0,$$

*где нули $\underline{u}(t)$ и $\bar{u}(t)$ каждый имеют индекс $n/2$ и строго перемежаются. Многочлен $\bar{u}(t)$ всегда обращается в нуль в точке $t = b$. В случае степенной функции $u_i(t) = t^i$, $i = 0, 1, \ldots, n$, написанное выше представление превращается в (10.1) при условии, что $P(t) = u(t) > 0$ на $[a, b]$.*

Доказательство. Если $f(t)$ является многочленом, то функция $f(t) - \underline{u}(t)$, очевидно, удовлетворяет условиям (i) — (iv), так что из свойства единственности мы заключаем, что

$$f(t) - \underline{u}(t) = \bar{u}(b).$$

Следствие 10.1 (b). *Если система $\{u_i\}_0^{n+1}$, где $u_{n+1}(t) = f(t)$, является Т-системой в добавление к тому, что $\{u_i\}_0^n$ является Т-системой, то разности $f(t) - \underline{u}(t)$ и $f(t) - \bar{u}(t)$ обращаются в нуль на множестве с индексом $(n + 1)/2$.*

Следствие 10.1 (в) означает, что:

(i) *Если $n = 2m$, то тогда $f(t) - \bar{u}(t)$ обращается в нуль ровно один раз между каждой парой соседних нулей $\bar{u}(t)$, и если $t_1, \ldots, t_m$ — нули $\underline{u}(t)$, то $f(t) - \underline{u}(t)$ обращается в нуль ровно один раз на каждом из интервалов $(t_i, t_{i+1})$, $i$ $1, \ldots, m - 1$, и один раз в $[a, t_1)$ и $(t_m, b]$, причем нулем в одном из таких интервалов может быть только крайняя точка $a$ или $b$.*

(ii) *Если $n = 2m + 1$, то тогда $f(t) - \underline{u}(t)$ обращается в нуль ровно один раз между каждой парой нулей $\underline{u}(t)$ и наибольший* *нуль находится в точке $b$. Аналогичное утверждение справедливо для $\bar{u}(t)$.*

В теореме 10.2 ниже мы построим два многочлена, которые изменяются между двумя непрерывными функциями $f$ и $g$ специальным образом. Теорема 10.1 является частным случаем, в котором $g(t) \equiv 0$ и $v(t)$ является некоторым строго положительным многочленом, удовлетворяющим $f(t) > v(t)$. Заметим, что функция $g(t)$ не предполагается многочленом, так что заключение теоремы 10.2 не может быть выведено из теоремы 10.1 простой заменой $f(t)$ на $f(t) - g(t)$.

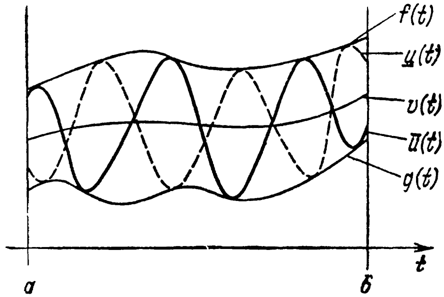


Рис. 7.

Теорема 10.2. *Пусть $\{u_i\}_0^n$ — $T$-система и $f$ и $g$ — две непрерывные функции на $[a, b]$ такие, что существует многочлен $v(t)$, лежащий между $f$ и $g$, т. е.*

$$f(t) > v(t) > g(t), \qquad t \in [a, b].$$

(a) *Тогда существует единственный многочлен $\underline{u}$, удовлетворяющий свойствам:*

(i) $f(t) \geqslant u(t) \geqslant g(t), \quad t \in [a, b]$;

(ii) *существуют $n+1$ точки $s_1 < s_2 < \ldots < s_{n+1}$ такие, что*

$$u(s_{n+1-i}) = \begin{cases} f(s_{n+1-i}), & i = 0, 2, 4, \ldots, \\ g(s_{n+1-i}), & i = 1, 3, 5, \ldots \end{cases} \tag{10.2}$$

(b) *Условие* (ii) *может быть заменено условием* (ii)′, *получающимся из* (10.2) *переменной местами функций $f$ и $g$. Тогда существует единственный многочлен $\bar{u}$, удовлетворяющий* (i) *и* (ii)′.

На рис. 7 приводится изображение многочленов $\underline{u}$, $\bar{u}$ и $v$ для случая $n = 5$.

Доказательство. Мы снова ограничимся только случаем $n = 2m$.

Предполагая сначала, что $\{u_i(t)\}$ является $ET$-системой, мы действуем, как в теореме 10.1. Для каждого $\xi$ в $n$-мерном симплексе $\Xi$ точек $\xi = (\xi_0, \ldots, \xi_n)$ с $\xi_i \geqslant 0$ и $\Sigma\xi_i = b - a$ мы строим многочлен $u_\xi(t) = \sum_{i=0}^{n} a_i(\xi)\, u_i(t)$, где $\sum_{i=0}^{n} a_i^2(\xi) = 1$, который обращается в нуль в каждой из точек

$$t_i = a + \sum_{k=0}^{i-1} \xi_k, \qquad i = 1, \ldots, n.$$

Знак $u_\xi(t)$ задается соглашением, что $u_\xi(t)$ неотрицательно для $t \in (t_i, t_{i+1})$, когда $i$ четно. Теперь для $i = 0, 2, 4, \dots, n$ мы определим

$$\delta_i(\xi) = \min\{\delta | \delta(f(t) - v(t)) \geqslant u_\xi(t), \qquad t \in [t_i, t_{i+1}]\},$$

где $t_0 = a$ и $t_{n+1} = b$, а для $i = 1, 3, \dots, n-1$ определим

$$\delta_i(\xi) = \min\{\delta \,|\, u_\xi(t) \geqslant \delta(g(t) - v(t)), \qquad t \in [t_i, t_{i+1}]\}.$$

Так же, как в теореме 10.1, мы образуем $F_k(\xi) = \delta_k(\xi) - \min \delta_i$. Снова предположение, что $\sum_{k=0}^{n} F_k(\xi) > 0$ для всех $\xi$ приводит к противоречию. Следовательно, существует $\xi^*$, для которого $\delta_i(\xi^*) = \delta$, $i = 0, 1, \dots, n$. Ясно, что $\delta > 0$ и что многочлен $\underline{u}(t) = \delta^{-1} u_{\xi^*}(t) + v(t)$ удовлетворяет условиям теоремы.

Многочлен $\overline{u}(t)$ строится тем же самым способом. Мы определяем $u_\xi(t)$ как отрицательный многочлен, уже вводившийся прежде, и делаем очевидные изменения в определении $\delta_i(\xi)$. Остальные соображения те же самые, что и раньше.

Доказательство случая, когда $\{u_i(t)\}$ — просто $T$-система, и единственности проводится, как в теореме 10.1.

Следствие 10.2. *Если $f$ — многочлен, то $f - \underline{u}$ и $f - \overline{u}$ обращаются в нуль на множестве индекса $n/2$.*

Мы сейчас расширим теорему 10.1 на общий случай, когда $\{u_i\}_0^n$ является $ET$-системой, $f \in C^n$ и разрешается иметь, самое большее, $n-1$ нуль, учитывая их кратность. Если $u_i(t) = t^i$, $i = 0, 1, \dots, n$, на интервале $[a, b]$ и $f$ — неотрицательный многочлен с $r \leqq n-1$ нулями на $[a, b]$, то желаемое расширение достигается просто представлением $f$ в форме $f(t) = Q_r(t) P(t)$, где $Q_r$ содержит все нули $f(t)$ на $[a, b]$ и $P$ является строго положительным многочленом на этом интервале.

Применяя теорему 10.1 к $P$, мы получаем $f(t) = Q_r(t)[\underline{u}(t) + \overline{u}(t)]$, где нули $\underline{u}(t)$ и $\overline{u}(t)$ лежат строго друг между другом и каждый имеет индекс $(n-r)/2$. Этот метод доказательства несправедлив для $T$-систем, так как для них процесс выделения нулей $f(t)$, вообще говоря, невозможен.

Теорема 10.3. *Пусть $\{u_i\}_0^n$ является $ET$-системой и $f(t) \in C_n$ — неотрицательная функция, определенная на $[a, b]$, и пусть $Z^*(f) = r < n$.* (Напомним, что $Z^*(f)$ обозначает число нулей $f$, учитывая их кратность.) *Тогда*:

(a) *существует единственный многочлен $\underline{u}(t)$, удовлетворяющий соотношениям*:

(i) $f(t) \geqslant u(t) \geqslant 0$,

(ii) *$u(t)$ имеет $n$ нулей, учитывая их кратности,*

(iii) *если* $t_{i_1}, \dots, t_{i_{n-r}}$ *—нули* $u(t)$, *которые остаются после удаления* $r$ $(r < n)$ *нулей* $f(t)$, *то* $f(t) - u(t)$ *обращается в нуль по крайней мере дважды (учитывая кратности) в каждом открытом интервале между соседними парами различных* $t_{i_k}$ *и по крайней мере один раз в каждом из интервалов* $[a, t_{i_1})$ *и* $(t_{i_{n-r}}, b]$;

(iv) *нули* $t_{i_1}, \dots, t_{i_{n-r}}$ *условия* (iii) *образуют множество индекса* $(n-r)/2$;

(v) $t_{i_{n-r}} < b$;

(b) *существует единственный многочлен* $\bar{u}$, *удовлетворяющий условиям* (i) — (iv) *и* (v)′ $t_{i_{n-r}} = b$.

Доказательство. Пусть $Q_1, \dots, Q_p$ — различные нули $f(t)$ кратности $m_1, \dots, m_p$, где $\sum_{i=0}^{p} m_i = r < n - 1$; обозначим разность $n -$ через $n'$. Доказательство теперь становится аналогичным доказательству теоремы 10.1, где $n$ заменяется на $n'$. Так как нечетный и четный случаи до некоторой степени одинаковы, мы разберем сейчас нечетный случай $n' = 2m + 1$. Построение $\underline{u}$ в части (a) производится следующим образом. Для каждого $\xi \in \bar{\Xi}^{m'}$ мы строим многочлен $\dot{u}_\xi(t) = \sum_{i=0}^{n} a_i(\xi) u_i(t)$, $\sum_{i=0}^{n} a_i^2(\xi) = 1$, который имеет нули кратностей $m_1, \dots, m_p$ в точках $Q_1, \dots, Q_p$, нули кратности два в $t_{i+1} = a + \sum_{k=0}^{i} \xi_k$, $i = 1, \dots, m$, и нуль кратности один в $t_1 = a$. Теперь определим

$$\delta_i(\xi) = \min\left\{\delta \mid \delta \geqslant \frac{u_\xi(t)}{f(t)},\ t \in [t_i, t_{i+1}]\right\}$$

для $i = 1, 2, \dots, m' + 1$ $(t_{m'+2} = b)$, где отношение $u_\xi/f$ оценивается по правилу Лопиталя, как

$$\frac{u_\xi^{(m_l)}(\theta_l)}{f^{(m_l)}(\theta_l)}$$

в нулях $\theta_1, \dots, \theta_p$ функции $f(t)$.

Рассматривая отношение $u_\xi/f$ сперва вблизи каждой из точек $\theta_1, \dots, \theta_p$, а затем на оставшейся части $[a, b]$, легко видеть, что если $\xi_1^{(k)} \to \xi$, то

$$\frac{u_{\xi^k}(t)}{f(t)} \to \frac{u_\xi(t)}{f(t)}$$

равномерно на $[a, b]$. Следовательно, каждая из $\delta_i(\xi)$ непрерывна по $\xi$ и $\delta_i(\xi) = 0$, тогда и только тогда, когда $\xi_i = 0$. Соображения, приведенные в теореме 10.1, теперь показывают, что для некоторого $\xi$ из внутренности $\Xi^{m'}$ мы имеем $\delta_i(\xi) = \delta > 0$ для $i = 1, \ldots$ $\ldots, m'+1$. Легко проверить, что функция $u(t) = \delta^{-1} u_\xi(t)$ обладает свойствами (i), (ii), (iv) и (v). Чтобы показать, что она также удовлетворяет условию (iii), мы заметим, что $\{t_{i_1}, \ldots, t_{i_{n-r}}\} = \{t_1, \ldots$ $\ldots, t_{m'+1}\}$ и что если $t_i = \theta_j$ для некоторого $j$, то тогда $u_\xi(t)$ имеет нуль в $\theta_j$ кратности, превышающей кратность нулей $f(t)$; так что $\delta$ строго больше, чем $u_\xi(t)[f(t)]^{-1}$ в некоторой окрестности $\theta_j$. Это означает равенство $\delta = u_\xi(t)[f(t)]^{-1}$ для некоторого $t$ в каждом из открытых интервалов $(t_1, t_2), \ldots, (t_{m'}, t_{m'+1})$ и $(t_{m'+1}, b]$. Таким образом, в каждом интервале $(t_i, t_{i+1})$ или $f(t) - \delta^{-1}u_\xi(t)$ обращается в нуль не в нулях $f(t)$, или кратность общего нуля $f(t)$ и $\delta^{-1}u_\xi(t)$ увеличивается на 2. В интервале $(t_{m'+1}, b]$ $f(t) - \delta^{-1}u_\xi(t)$ может обращаться в нуль в $b$ с кратностью только на единицу большей, чем нуль $f(t)$ в этой точке.

Многочлен $\bar{u}(t)$, когда $n' = 2m'+1$, строится аналогичным образом, задавая $b$ вместо $a$ как один из нулей.

Чтобы доказать единственность, мы допустим, что другой многочлен $v(t)$ удовлетворяет тем же самым свойствам, что и $\underline{u}(t)$. Доказательство единственности для $\bar{u}(t)$ проводится аналогичным образом. Без потери общности мы можем также допустить, что первый нуль $f - v$, отличный от нуля $f$, меньше или равен первому нулю $f - \underline{u}$. Мы положим $h(t) = (\underline{u} - v)/f$. Нуль $h(t)$ на одной из точек $\bar{t}_i$, $i = 2, \ldots, n'+1$, обязательно должен быть кратности по крайней мере 2. В этом случае мы выделим один нуль в каждом из интервалов $[t_{i-1}, t_i]$ и $[t_i, t_{i+1}]$, где $t_{m'+2} = b$. При такой процедуре подсчета, учитывая осцилляционные свойства $u$ и $v$, мы выводим, что $h(t)$ имеет по крайней мере три нуля на $[a, t_2]$, по крайней мере два нуля в каждом из интервалов $[t_i, t_{i+1}]$, $i = 2, \ldots, m'$, и по крайней мере один нуль на $[t_{m'+1}, b]$. Ясно, что все эти нули являются отличными от нулей $f$, так что $\underline{u} - v$ имеет по крайней мере $3 + 2(m'-1) + 1 + r = n+1$, учитывая кратности. Из этого следует, согласно теореме 4.3 гл. I, что $\underline{u} = v$.

Следствие 10.3. *Если выполнены условия теоремы* 10.3 *и, кроме того, $f$ является многочленом, то существует единственное представление*

$$f(t) = \underline{u}(t) + \bar{u}(t),$$

*где $\bar{u}$ и $\underline{u}$ — неотрицательные многочлены, имеющие $n$ нулей, которые лежат строго друг между другом, при этом нули $f(t)$ не рассматриваются. Многочлен $\bar{u}$ обращается в нуль в точке $t = b$.*

Доказательство. Если $f(t)$ — многочлен, то мы строим $\underline{u}(t)$, как в теореме 10.3, а затем проверяем, что $f(t) - \underline{u}(t)$ удовлетворяет условиям (i) — (iv) и (v)′, так что $f(t) = \underline{u}(t) + \overline{u}(t)$. Свойство корней лежать строго друг между другом вытекает из построения $\underline{u}(t)$.

Замечание 10.1. Из теоремы 10.3 видно, что результаты теоремы 10.2 могут быть распространены на случай, когда $f - g$ имеет некоторое число нулей.

Замечание 10.2. Результаты теорем 10.1 — 10.3 могут быть обобщены в другом направлении. Например, предположим, что $\{u_i\}_0^n$ — $T$-система на интервале $[c, d]$, где $[a, b] \subset [c, d]$. Для произвольной, но фиксированной непрерывной функции $f(t) > 0$, $t \in [a, b]$, мы рассмотрим полином $u(t)$, удовлетворяющий условию $f(t) \geqslant \geqslant u(t) \geqslant 0$ на $[a, b]$ и который имеет предписанные нули в $k$ различных точках ($k < n$) $z_1, \ldots, z_k$ в $[c, d] \setminus [a, b]$. Теорему 10.1 можно приспособить к этому условию, если величину $n$ заменить на $n-k$. Соображения, аналогичные использовавшимся в теореме 10.1, гарантируют существование точно двух многочленов $\underline{v}$ и $\overline{v}$, обращающихся в нуль в $z_1, \ldots, z_k$ и обладающих следующими свойствами:

(i) $f(t) \geqslant u(t) \geqslant 0$ на $[a, b]$;

(ii) $Z[a, b](u) = n - k$ (см. определение 4.3 гл. I);

(iii) $f(t) - u(t)$ обращается в нуль по крайней мере один раз между каждой парой соседних нулей на $[a, b]$, наибольшим нулем $u(t)$ на $[a, b]$ и граничной точкой $b$, наибольшим нулем и граничной точкой $a$.

Теория представления в § 10 принадлежит Карлину [1963]. В частном случае простых многочленов соответствующая теория была развита Карлином и Шепли [1953]. Более ранние и грубые формулы представлений, выражающие определенные полиномы как суммы квадратов, встречаются у Лукача (см. Полиа, Сеге [1925], т. 2, гл. 5, и Брикман [1960] и ссылки в них).

Сложность аргументов в выводе формулы представления для общих $n$-многочленов $T$-систем проистекает из факта, что мы не можем выделить нули, как в случае простых многочленов. Другая трудность состоит в том, что нужно избавиться от недостаточной дифференцируемости.

## § 11. Крайние точки $\mathcal{P}_{n+1}$

Конструкция множества в начале § 10 позволяет нам охарактеризовать крайние лучи двойственного конуса $\mathcal{P}_{n+1}$ множества $\mathcal{M}_{n+1}$. Вспомним из теоремы 9.2, что граница $\mathcal{P}_{n+1}$, где $\{u_i\}_0^n$ является $T$-системой, совпадает с множеством всех неотрицательных многочленов, обращающихся по крайней мере один раз в нуль. Следующие две теоремы устанавливают факт, что если $\{u_i\}_0^n$ является $ET$-системой, то крайние лучи $\mathcal{P}_{n+1}$ соответствуют неотрицательным полиномам с $n$ нулями, учитывая их кратность.

Теорема 11.1. *Если $\{u_i\}_0^n$ является $T$-системой и $u_i \in C^n$, $i = 0, 1, \ldots, n$, то экстремальные лучи $\mathcal{P}_{n+1}$ содержатся в множестве* *всех неотрицательных многочленов с $n$ нулями, учитывая их кратности.*

Доказательство. Достаточно показать, что если $n$ — произвольный многочлен менее чем с $n$ нулями, учитывая их кратност›, то $u(t)$ может быть представлен в форме $u(t) = v(t) - w(t)$, где $v(t)$ и $w(t)$ линейно независимы и принадлежат $\mathcal{P}_{n+1}$. Чтобы получить это представление, мы положим $r(t)$ фиксированным положительным многочленом. Тогда по теореме 10.1 мы имеем $u(t) + k^{-1}r(t) = \bar{u}(t, k) - \underline{u}(t, k)$ для $k = 1, 2, 3, \ldots$, где $\bar{u}(t, k)$ и $\underline{u}(t, k)$ каждый имеют $n$ нулей, учитывая их кратности. Как и в доказательстве теоремы 10.1, величины $a_i(k)$ и $b_i(k)$, $i = 0, \ldots, n$, встречающиеся в выражениях

$$\bar{u}(t, k) = \sum_{i=0}^{n} b_i(k) u_i(t), \quad \underline{u}(t, k) = \sum_{i=0}^{n} a_i(k) u_i(t)$$

равномерно ограничены по $k$. Мы можем, следовательно, выбрать подпоследовательность $k' \to \infty$ такую, что $b_i(k') \to b_i$ и $a_i(k') \to a_i$ для $i = 0, 1, \ldots, n$. Очевидно, что $r$-е производные ($r = 0, 1, \ldots, n$) $\bar{u}(t, k')$ и $\underline{u}(t, k')$ сходятся к соответствующим производным предельных многочленов, которые мы обозначим через

$$v = \sum_{i=0}^{n} a_i u_i \text{ и } w = \sum_{i=0}^{n} b_i u_i.$$

Многочлены $v(t)$ и $w(t)$ имеют $n$ нулей, учитывая их кратности, так как аппроксимирующие многочлены обладают этим свойством. Более того, $v(t)$ и $w(t)$ линейно независимы, так как, если $v(t) = \alpha w(t)$ для некоторого $\alpha$, то $u(t)$ имеет $n$ нулей. Мы имеем, следовательно, представление $u(t) = v(t) + w(t)$, где $v$ и $w$ линейно независимы, а $v$ и $w$ принадлежат к $\mathcal{P}_{n+1}$. Доказательство закончено.

Теорема 11.2. *Если $\{u_i(t)\}_0^n$ — ЕТ-система, то $u(t) \in \mathcal{P}_{n+1}$ имеет $n$ нулей, учитывая их кратности, если и только если $u(t)$ принадлежит крайнему лучу $\mathcal{P}_{n+1}$.*

Доказательство. Необходимость доказана в теореме 11.1. Чтобы доказать обратное, мы предположим, что $u(t)$ имеет $n$ нулей, учитывая их кратности, и не содержится в крайнем луче. Тогда $u(t) = v(t) + w(t)$, где $v(t) \not\equiv 0$ и $w(t) \not\equiv 0$ принадлежат $\mathcal{P}_{n+1}$ и линейно независимы. Так как $v$ и $w$ обращаются в нуль в $n$ нулях $u(t)$ и $u(t) \not\equiv 0$, то существует некоторое $t_0$, для которого $u(t_0) v(t_0) w(t_0) > 0$. При этом условии, однако, $ET$-свойство системы $\{u_i(t)\}$ приводит к тождеству

$$v(t) \equiv \frac{v(t_0)}{w(t_0)} w(t).$$

так как разность имеет $n + 1$ нуль, учитывая кратности.

Глава III

# ТЕОРЕМА МАРКОВА — КРЕЙНА И ЕЕ ОТВЕТВЛЕНИЯ

В этой главе мы будем в основном заниматься задачей определения верхней и нижней границ для интеграла

$$\int_a^b \Omega(t)\, d\sigma(t),$$

где $\sigma$ пробегает множество всех мер, удовлетворяющих определенным моментным условиям по отношению к $T$-системе $\{u_i\}_0^n$. Как правило, мы будем требовать, чтобы мера $\sigma$ имела заданное число моментов $\mathbf{c}^0 = (c_0^0, c_1^0, \ldots, c_n^0)$, т. е. $\sigma \in V(\mathbf{c}^0)$ (см. (4.4) гл. II). Неравенства, полученные в процессе решения этих экстремальных задач, относятся к категории результатов, объединенных под общим названием чебышевских неравенств. На этот раз мы наложим определенные ограничения на функции $u_0, u_1, \ldots, u_n$ и $\Omega$. В гл. XII—XIV будет предпринято глубокое исследование природы общих чебышевских неравенств. Теория в этих главах будет проиллюстрирована многочисленными конкретными примерами, представляющими интерес в теории вероятностей и математической статистике.

В § 1 этой главы воспроизводятся результаты § 6 гл. II для облегчения ссылок. В § 2 детально разбирается первое доказательство теоремы Маркова — Крейна (теоремы 2.1), которое существенно опирается на теоремы интерполяции из § 6 гл. I. В § 3 предлагается другое, возможно, более ясное доказательство теоремы 2.1.

Экстремальные чебышевские задачи, описанные в §§ 1 и 2, известны достаточно давно и первоначально поставлены для случая $\Omega(t) = 1$. В этом случае задача сводится к определению границ для меры $\sigma$, имеющей заданные моменты. Теорема Маркова — Крейна, сформулированная в § 2, служит обобщением первоначальной теоремы Чебышева. В § 4 рассматриваются различные обобщения; § 5 посвящен описанию другого класса неравенств, в котором ограниченные меры соответствуют моментным точкам на подмножестве из $\mathcal{M}_{n+1}$, а не обязательно одной точке. В § 6 будут охарактеризованы крайние точки множества $V(\mathbf{c}^0)$.

## § 1. Экстремальные значения из $R(\mathbf{c}^0)$

Пусть $\mathbf{c}^0$ содержится внутри моментного пространства $\mathcal{M}_{n+1}$, индуцированного $T$-системой $\{u_i\}_0^n$ на интервале $[a, b]$. *Первая чебышевская задача* состоит в нахождении мер $\sigma$, принадлежащих множеству $V(\mathbf{c}^0) = \left\{\sigma \mid c_i^0 = \int u_i(t)\, d\sigma(t),\ i = 0, 1, \dots, n\right\}$, при которых достигаются экстремальные значения

$$\max_{\sigma \in V(\mathbf{c}^0)} \int_a^b \Omega(t)\, d\sigma(t), \tag{1.1}$$

$$\min_{\sigma \in V(\mathbf{c}^0)} \int_a^b \Omega(t)\, d\sigma(t). \tag{1.2}$$

В случае, когда функции

$$u_0(t),\ u_1(t), \dots, u_n(t),\ u_{n+1}(t) = \Omega(t) \tag{1.3}$$

образуют $T$-систему, эта задача уже была решена в § 6 предыдущей главы, где было рассмотрено одномерное сечение $R(\mathbf{c}^0)$. Сформулируем теперь полученный результат.

Теорема 1.1. *Пусть $\{u_i\}_0^n$ и расширенная система, полученная присоединением функции $\Omega$, являются $T$-системами. Тогда (1.1) достигается только на $\bar{\sigma}$, т. е. на мере, соответствующей верхнему главному представлению точки* $\mathbf{c}^0$, *а (1.2) достигается только на $\underline{\sigma}$, т. е. на мере, соответствующей нижнему главному представлению точки* $\mathbf{c}^0$.

Замечание 1.1. Предположения теоремы 1.1 могут быть ослаблены в том смысле, что $\{u_i\}_0^n$ должна быть $T$-системой, но $\{u_0, u_1, \dots, u_n, \Omega\}$ может предполагаться просто $WT$-системой, т. е.

$$\Delta\begin{pmatrix} u_0, u_1, \dots, u_n, \Omega \\ t_0, t_1, \dots, t_n, t_{n+1} \end{pmatrix} \geqslant 0 \text{ для } a \leqslant t_0 < t_1 < \dots < t_{n+1} \leqslant b \tag{1.4}$$

(заметим, что в (1.4) допускается равенство). В этом случае главные меры $\bar{\sigma}$ и $\underline{\sigma}$ остаются экстремальными мерами для (1.1) и (1.2) соответственно, но при этом могут существовать другие экстремальные меры.

Замечание 1.2. Представляет собой замечательный факт, что максимизирующая и минимизирующая меры $\bar{\sigma}$ и $\underline{\sigma}$ соответственно не зависят от функции $\Omega$ при условии, что $\{u_i\}_0^n$ и $\{u_i\}_0^{n+1}$ ($\Omega(t) = u_{n+1}(t)$) являются $T$-системами.

Замечание 1.3. Предположим, что $\Omega$ — непрерывная функция, определенная на полуоткрытом интервале $[a, b)$, и что функции (1.3) образуют $T$-систему. В этом случае рассмотрим только такие меры $\sigma \in V(\mathbf{c}^0)$, которые не размещают положительную массу в точке $b$. Для таких мер $\sigma$ определим

$$E(\sigma) = \lim_{\beta \to b} \int_a^\beta \Omega(t)\, d\sigma(t).$$

Легко видеть, что $\underline{\sigma}$ минимизирует $E(\sigma)$ на множестве $V(\mathbf{c}^0)$.

## § 2. Теорема Маркова—Крейна

Вернемся теперь к важной задаче определения максимума и минимума значений $\int_a^{\xi+} \Omega(t)\, d\sigma(t)$ для $\sigma \in V(\mathbf{c}^0)$, где $\xi$ фиксирована $(a < \xi < b)$.

Если предположить, что $\Omega(a) > 0$, то решение задачи поиска максимума в случае $\xi = a$ однозначно дается либо мерой $\bar{\sigma}$, либо мерой $\underline{\sigma}$, которые имеют максимальную массу в граничной точке $a$. Когда $\bar{\Omega}(a) > 0$, минимум для $\xi = a$, очевидно, равен нулю и достигается на другой мере, но в общем случае решение может быть не единственным.

Для случая $\xi \neq a$ вышеуказанная задача становится важным обобщением задачи, рассмотренной в § 1. Формально заменим функцию $\Omega(t)$ в (1.1) и (1.2) модифицированной функцией

$$\Omega_\xi(t) = \begin{cases} \Omega(t), & t \leq \xi, \\ 0, & t > \xi. \end{cases}$$

Заметим, что система $u_0, u_1, \ldots, u_n$, дополненная функцией $\Omega$, вообще говоря, не образует ни $T$-систему, ни даже $WT$-систему, хотя функции $u_0, u_1, \ldots, u_n$ могут первоначально быть $T$-системой.

В классической формулировке этой задачи функция $\Omega(t)$ тождественно равна единице, и соответствующая задача сводится к задаче определения границ неубывающей функции $\int_a^\xi d\sigma$, где мера $\sigma$ подчинена определенным моментным условиям. Например, в теории вероятностей часто требуется определить точные границы вероятности

$$\Pr(X \leq \xi) = \int_a^{\xi+} d\sigma,$$

где $X$ — случайная величина с функцией распределения $\sigma$, и первые $n$ моментов $\int_a^b t^i d\sigma(t) = c_i$, $i = 1, \ldots, n$ $(c_0 = 1)$ заданы.

На протяжении этого параграфа будем считать, что функции $u_0, \ldots u_n$, $\Omega$ удовлетворяют условию 6.1 гл. I. Это позволяет использовать теорему интерполяции 6.2 гл. I. Вспомним, что условие

6.1 гл. I требует, чтобы для любого $k=0, \dots, n$ системы функций $\{u_i\}_0^k$ и $\{u_0, \dots, u_k, \Omega\}$ являлись $T$-системами, а функция $\Omega$ строго положительна на $[a, b]$. (В § 4 некоторые из этих ограничений будут ослаблены.)

Теорема 2.1 (Марков — Крейн). *Пусть условие* 6.1 гл. I *выполняется. Далее пусть* $c^0 \in \operatorname{Int} \mathcal{M}_{n+1}$, *а* $\sigma_{\xi_0}$ *обозначает соответствующую каноническую меру, включающую точку* $\xi_0 \in (a, b)$. *Тогда для любой меры* $\sigma \in V(c^0)$, $\sigma \neq \sigma_{\xi_0}$, *оказываются справедливыми следующие неравенства*:

$$\int_a^{\xi_0^-} \Omega(t)\, d\sigma_{\xi_0}(t) < \int_a^{\xi_0^-} \Omega(t)\, d\sigma(t) \leq \int_a^{\xi_0^+} \Omega(t)\, d\sigma(t) < \int_a^{\xi_0^+} \Omega(t)\, d\sigma_{\xi_0}(t), \quad (2.1)$$

*исключая случай, когда мера* $\sigma_{\xi_0}$ *имеет индекс* $(n+2)/2$, *и* $\xi_0$ *является первым корнем меры* $\sigma_{\xi_0}$. *В этом особом случае может иметь место равенство между первыми двумя интегралами в* (2.1), *причем только тогда, когда оба эти интеграла равны нулю.*

Доказательство. Пусть точка $\xi_0$ $(a < \xi_0 < b)$ фиксирована, и пусть $\sigma_{\xi_0}$ обозначает каноническую меру, соответствующую $c^0$ и включающую $\xi_0$. Представим корни меры $\sigma_{\xi_0}$ как $T = \{t_1, \dots, t_p\}$ и обозначим $t_{i_0} = \xi_0$. По теореме 6.2 гл. I и ее следствию, если $\xi_0 \neq t_1$, то существует многочлен $u^*$ со следующими свойствами: $\Omega(t) - u^*(t) \geqslant 0$ для $t \in [a, \xi_0]$, причем равенство имеет место только при $t_1, \dots, t_{i_0-1}$; и $u^*(t) \leq 0$ для $t \in [\xi_0, b]$, причем равенство имеет место только при $t_{i_0}, \dots, t_p$. Для любого многочлена $u$ и меры $\sigma \in V(c^0)$ имеем $\int u\, d\sigma = \int u\, d\sigma_{\xi_0}$. В частности, для $u = u^*$ и меры $\sigma \in V(c^0)$, отличной от меры $\sigma_{\xi_0}$, получим

$$\int_a^{\xi_0^-} \Omega(t)\, d\sigma_{\xi_0}(t) = \int_a^{\xi_0^-} u^*(t)\, d\sigma_{\xi_0}(t) = \int_a^b u^*(t)\, d\sigma_{\xi_0}(t) =$$

$$= \int_a^b u^*(t)\, d\sigma(t) = \int_a^{\xi_0^-} u^*(t)\, d\sigma(t) + \int_{\xi_0^-}^b u^*(t)\, d\sigma(t) < \int_a^{\xi_0^-} u^*(t)\, d\sigma(t) \quad (2.2)$$

(так как $u^*(t) < 0$ на $[\xi_0, b]$, исключая точки $t_{i_0}, \dots, t_p$). Строгое неравенство имеет место на основании леммы 3.1 гл. II, в которой утверждается, что мера $\sigma$ содержит массу на открытом интервале между каждой парой внутренних корней меры $\sigma_{\xi_0}$.

Случай $\xi_0 = t_1$ имеет особенность. Действительно, если $\xi_0 = t_1$, то

$$\int_a^{\xi_0^-} \Omega(t)\, d\sigma_{\xi_0}(t) = 0.$$

Если $\sigma_{\xi_0}$ имеет индекс $(n+1)/2$, то $c^0$ является граничной точкой моментного пространства, индуцированного системой функций $\{u_i\}_0^n$ на интервале $[\xi_0, b]$. По теореме 2.1 гл. II точка $c^0$ имеет единственное представление на этом интервале, так что любая мера на $[a, b]$, отличная от меры $\sigma_{\xi_0}$, должна размещать некоторую массу на полуоткрытом интервале $[a, \xi_0)$, и, значит, левое неравенство в (2.1) выполняется. Если $\sigma_{\xi_0}$ имеет индекс $(n+2)/2$, то существует другое представление точки $c^0$, которому не соответствует масса на $[a, \xi_0)$, так что оба члена в левой части неравенства (2.1) обращаются в нуль.

Аналогичное рассуждение, примененное к многочлену $u = u^{**}$, который построен в теореме 6.2 гл. II, приводит к выводу, что

$$\int_a^{\xi_0^+} \Omega(t)\, d\sigma(t) < \int_a^{\xi_0^+} \Omega(t)\, d\sigma_{\xi_0}(t), \qquad \sigma \in V(c^0), \quad \sigma \neq \sigma_{\xi_0}.$$

## § 3. Другое доказательство теоремы Маркова — Крейна

Аналитический аппарат, примененный в § 2, заимствован из работы Крейна [1951]. В этом параграфе предлагается другое доказательство теоремы 2.1. Хотя это доказательство может показаться несколько более сложным, чем приведенное в § 2, но мы докажем попутно несколько лемм, представляющих и некоторый самостоятельный интерес.

Будем предполагать на протяжении этого параграфа выполненными следующие условия.

Условие 3.1.

(a) $\{u_i\}_0^n$ — $T$-система;

(b) $\{u_0, u_1, \ldots, u_n, \Omega\}$ — $T$-система и $\Omega(t) > 0$, $t \in [a, b]$;

(c) $u_i(t)$, $i=0, 1, \ldots, n$, и $\Omega(t)$ принадлежат классу $C^1 = C^1[a, b]$;

(d) функции

$$v_i(t) = \frac{d}{dt}\left[\frac{u_i(t)}{\Omega(t)}\right], \qquad i = 0, 1, \ldots, n, \tag{3.1}$$

таковы, что $\{(-1)^{n+1} v_0, v_1, \ldots, v_n\}$ является $T$-системой.

Замечание 3.1. Множитель $(-1)^{n+1}$ введен, чтобы быть уверенными, что условия (b) и (d) согласуются. Условие (d) является в известной мере следствием условия (b). Действительно, если $a \leq t_0 < t_1 < \ldots < t_{n+1} \leq b$, то, разделив $i$-й столбец определителя

$$\begin{vmatrix} u_0(t_0) & \ldots & u_0(t_{n+1}) \\ \cdot & \cdot\;\cdot\;\cdot & \cdot \\ u_n(t_0) & \ldots & u_n(t_{n+1}) \\ \Omega(t_0) & \ldots & \Omega(t_{n+1}) \end{vmatrix} > 0$$

на $\Omega(t_i)$ и воспользовавшись $n+1$ раз теоремой о среднем, приходим к выводу, что

$$\begin{vmatrix} \dfrac{u_0(t_0)}{\Omega(t_0)} & v_0(t'_0) \ldots v_0(t'_n) \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \dfrac{u_n(t_0)}{\Omega(t_0)} & v_n(t'_0) \ldots v_n(t'_n) \\ 1 & 0 \ \ \ldots \ \ 0 \end{vmatrix} > 0,$$

где $t_i < t'_i < t_{i+1}$, $i = 0, 1, \ldots, n$. В этом случае определитель

$$V\begin{pmatrix} 0, & 1, & \ldots, & n \\ t'_0, & t'_1, & \ldots, & t'_n \end{pmatrix}$$

имеет знак $(-1)^{n+1}$.

Определение 3.1. Говорят, что измеримая функция $f$ на $[a, b]$ имеет $n$ *существенных перемен знака* $S_\mu(f)$ по отношению к мере $\mu$ на $[a, b]$, если:

1) существуют множества $A_i \subset [a, b]$, $i = 0, 1, \ldots, n$, такие, что $A_i < A_{i+1}$, $i = 0, 1, \ldots, n-1$, $\mu\left(\bigcup\limits_{i=0}^{n} A_i\right) = b - a$, а $f$ имеет противоположные знаки на $A_i$ и $A_{i+1}$, $i = 0, 1, \ldots, n-1$ ($A < B$ будет означать, что из $x \in A$, $y \in B$ вытекает $x < y$, т. е. множество $A$ лежит левее множества $B$);

2) существуют подмножества $A_i^* \subset A_i$, $i = 0, 1, \ldots, n$, такие, что $f(t) \neq 0$ на $A_i^*$ и $\mu(A_i^*) > 0$.

Если функция $f$ не удовлетворяет указанным выше ограничениям для любого конечного $n$, то полагаем $S_\mu(f) = +\infty$. Всякий раз, когда $\mu$ обозначает меру Лебега на $[a, b]$, мы будем просто записывать $S_\mu(f) = S(f)$.

Лемма 3.1. *Пусть условие* 3.1 *выполняется, и предположим, что*

$$\mathbf{c}^0 = (c_0^0, c_1^0, \ldots, c_n^0) \in \operatorname{Int} \mathcal{M}_{n+1}.$$

*Рассмотрим множество всех мер* $\sigma \in V(c_0^0, c_1^0, \ldots, c_n^0, \gamma)$, *т. е. все меры* $\sigma$, *для которых*

$$c_i^0 = \int_a^b u_i(t)\, d\sigma(t), \quad i = 0, 1, \ldots, n; \qquad \gamma = \int_a^b \Omega(t)\, d\sigma(t),$$

*где* $\gamma \in R(\mathbf{c}^0)$ (см. равенство (6.3) гл. II). *Если* $\sigma_1$ *и* $\sigma_2$ — *две различные меры, принадлежащие множеству* $V(c_0^0, c_1^0, \ldots, c_n^0, \gamma)$, *то функция*

$$f(\xi) = \int_a^{\xi+} \Omega(t)\,(d\sigma_1(t) - d\sigma_2(t)), \qquad a \leq \xi \leq b, \tag{3.2}$$

*удовлетворяет условию $S(f) \geqslant n+1$. (Заметим, с целью последующих ссылок, что любая функция $\int\limits_a^{\xi+} \Omega(t)\, d\sigma_i(t)$, $i=1, 2$, является монотонно возрастающей, а также что если $\gamma$ — граничная точка интервала $R(c^0)$, то множество $V(c_0^0, \ldots, c_n^0, \gamma)$ сводится к единственной мере).*

Доказательство. Интегрируя по частям выражение в правой части равенства

$$0 = \int\limits_a^b \frac{u_i(t)}{\Omega(t)} \Omega(t)\,(d\sigma_1 - d\sigma_2), \qquad i = 0, 1, \ldots, n,$$

получим

$$0 = \int\limits_a^b v_i(t)\, f(t)\, dt, \qquad i = 0, 1, \ldots, n. \tag{3.3}$$

Если $S(f) = p \leqq n$, то существуют множества $A_0 < A_1 < \ldots < A_p$ такие, что $f$ имеет противоположные знаки на последовательных $A_i$ и $\mu(\bigcup A_i) = b - a$. Согласно условию 3.1 (d) и теореме 5.2 гл. I можно образовать нетривиальный $v$-многочлен, единственными нулями которого на $(a, b)$ служат узловые нули в точках

$$t_i = \sup\{t \mid t \in A_i\}, \qquad i = 0, 1, \ldots, p-1.$$

Очевидно, $v(t) f(t)$ сохраняет постоянный знак и не равна тождественно нулю, так что $\int\limits_a^b v(t) f(t)\, dt \neq 0$. Это, очевидно, противоречит равенству (3.3), откуда следует, что $S(f) > n$.

Лемма 3.1 может быть геометрически интерпретирована путем сравнения графиков функций

$$f_i(\xi) = \int\limits_a^{\xi+} \Omega(t)\, d\sigma_i(t),$$

$$i = 1, 2.$$

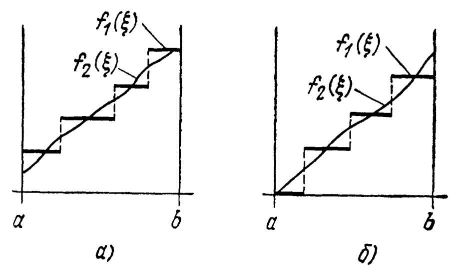



Рис. 8.

Неравенство $S(f) = S(f_1 - f_2) \geqslant n + 1$ показывает, что графики функций $f_1$ и $f_2$ пересекаются по крайней мере $n + 1$ раз. Наиболее интересный и полезный случай (см. лемму 3.2) имеет место, когда одна из мер (скажем, $\sigma_1$) имеет индекс $(n+3)/2$. В этом случае $f_1(\xi)$ — ступенчатая функция, как показано на рис. 8. Основные характерные свойства функций $f_1$ и $f_2$ изображены для случая $n = 4$. График функции $f_2$ начерчен как непрерывная кривая, хотя он также может иметь разрывы.

Заметим, что в каждом случае существует точно $n+2=6$ значений, в которых функция $f(\xi)$ может менять знак. Функцию $f_1$ можно рассматривать как составленную из суммы $n+2$ чередующихся горизонтальных и вертикальных сегментов, которые определяют точки, в которых $f_1$ и $f_2$ могут пересечься. Если кривая $f_2$ не пересекает каждый из этих $n+2$ сегментов, то число перемен знака функции $f(t)=f_1(t)-f_2(t)$ уменьшится по крайней мере на два, так что мы имели бы $S(f)\leqslant n$, что, согласно лемме 3.1, невозможно.

Лемма 3.2. *Пусть условия леммы* 3.1 *выполняются, и пусть для фиксированной* $\xi_0$ $(a<\xi_0<b)$ $\sigma_{\xi_0}(t;\gamma)\in V(c_0^0,\ldots,c_n^0,\gamma)$ *обозначает каноническую меру, включающую корень* $\xi_0$. *Если* $\sigma\in V(c_0^0,\ldots,c_n^0,\gamma)$ *и* $\sigma(t)\not\equiv\sigma_{\xi_0}(t;\gamma)$, *то*

$$\int_a^{\xi_0^-}\Omega(t)\,d\sigma_{\xi_0}(t;\gamma)<\int_a^{\xi_0^-}\Omega(t)\,d\sigma\leqq\int_a^{\xi_0^+}\Omega(t)\,d\sigma<\int_a^{\xi_0^+}\Omega(t)\,d\sigma_{\xi_0}(t;\gamma)\qquad(3.4)$$

*при условии, что* $\xi_0$ *не является первым корнем меры* $\sigma_{\xi_0}(t;\gamma)$. *Если же* $\xi_0$ *— первый корень меры* $\sigma_{\xi_0}(t;\gamma)$, *то два интеграла, стоящие слева в неравенствах* (3.4), *могут обратиться в нуль.*

Доказательство. Из леммы 3.1 известно, что функция

$$f(\xi)=\int_a^{\xi^+}\Omega(t)\,d\sigma(t)-\int_a^{\xi^+}\Omega(t)\,d\sigma_{\xi_0}(t;\gamma)$$

изменяет знак по крайней мере $n+1$ раз. По определению меры $\sigma_{\xi_0}(t;\gamma)$, $g(\xi)=\int_a^{\xi^+}\Omega(t)\,d\sigma_{\xi_0}(t;\gamma)$ есть возрастающая кусочно-постоянная функция, которая может быть представлена как сумма горизонтальных и вертикальных сегментов. Самое большее, $n+2$ из этих сегментов определяют точки, в которых $f(\xi)$ может изменить знак, так что $S(f)\leqq n+2$. Если $\xi_0$ — первый корень меры $\sigma_{\xi_0}(t;\gamma)$ то левое неравенство очевидно. В остальных случаях, если либо левое неравенство, либо правое неравенство нарушается, то $S(f)\leqq n$, что противоречит лемме 3.1.

Следствие 3.2. *Пусть условия леммы* 3.1 *выполняются, и предположим, что* $c^0=(c_0^0,c_1^0,\ldots,c_n^0)\in\operatorname{Int}\mathcal{M}_{n+1}$. *Тогда для заданной, точки* $\xi_0$ $(a<\xi_0<b)$

$$\max_{\sigma\in V(c^0)}\int_a^{\xi_0^+}\Omega(t)\,d\sigma(t)\leqq\max_{\sigma_{\xi_0}(t;\gamma)\in V(c^0)}\int_a^{\xi_0^+}\Omega(t)\,d\sigma_{\xi_0}(t;\gamma),$$

$$\min_{\sigma\in V(c^0)}\int_a^{\xi_0^-}\Omega(t)\,d\sigma(t)\geqslant\min_{\sigma_{\xi_0}(t;\gamma)\in V(c^0)}\int_a^{\xi_0^-}\Omega(t)\,d\sigma_{\xi_0}(t;\gamma).$$

Мера $\sigma_{\xi_0}(t;\gamma)$ определена в лемме 3.2. Для $\gamma$, равной граничной точке $\mathring{R}(\mathbf{c}^0)$, положим $\sigma_{\xi_0}(t;\gamma)$ равной соответствующему единственному верхнему или нижнему главному представлению.

Если добавить к условию 3.1 дополнительное ограничение, то результат следствия 3.2 может быть усилен, и получена теорема 2.1.

Условие 3.2. Системы $\{u_0, u_1, \dots u_r\}$ и $\{u_0, u_1, \dots, u_r, \Omega\}$ являются $T$-системами для всех $r = 0, 1, \dots, n$.

Заметим, что условие 3.1 (b) и условие 3.2 эквивалентны условию 6.1, использовавшемуся при доказательстве теоремы 2.1 в § 2. Дополнительные требования в условии 3.1 (c) и (d) относятся к системе функций (3.1).

Используя только условие 3.1 (a) и (b), т. е. что $\{u_i\}_0^n$ — $T$-система, $\{u_0, \dots u_n, \Omega\}$ — $T$-система и $\Omega(t) > 0$, вместе с дополнительными предположениями относительно системы (3.1), приходим к заключению, сформулированному в следствии 3.2, что экстремальные значения интеграла $\int_a^{\xi} \Omega(t)\, d\sigma(t)$ для $\sigma \in V(\mathbf{c}^0)$ доставляются мерой индекса, самое большее, $(n+3)/2$, содержащей положительную массу в точке $\xi_0$. Для того чтобы сравнить значения этого интеграла для мер индекса $(n+3)/2$ и канонической меры $\sigma_{\xi_0}$, нам в дальнейшем потребуется условие 3.2.

Рассмотрим два различных представления точки $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}$. В частности, пусть

$$c_i^0 = \sum_{j=1}^{p} \lambda'_j u_i(t'_j) = \sum_{j=1}^{q} \lambda''_j u_i(t''_j), \qquad i = 0, 1, \dots, n, \tag{3.5}$$

где первая сумма соответствует канонической мере $\sigma_{\xi_0}$ (индекс $\leq (n+2)/2$), включающей точку $\xi_0 \in (a, b)$, а вторая сумма соответствует мере $\sigma$ индекса $(n+3)/2$, также включающей точку $\xi_0$. Множество $\{t'_j\} \cup \{t''_j\} = \{s_k\}$ содержит ровно $n+2$ различные точки. По лемме 3.1 гл. II это множество имеет по крайней мере $n+2$ точки, и если исследовать различные возможности для $n$ четного или нечетного, то легко убедиться, что число точек не может быть больше чем $n+2$.

Перепишем (3.5) в виде

$$0 = \sum_{k=1}^{n+2} a_k u_i(s_k) = \sum_{k=1}^{n+2} b_k w_i(s_k) \qquad (i = 0, 1, \dots n), \tag{3.6}$$

где $b_k = a_k \Omega(s_k)$ и $w_i(s_k) = [u_i(s_k)]/[\Omega(s_k)]$. Кроме того, зафиксируем знаки $a_k$, потребовав, чтобы коэффициент, соответствующий $\xi_0$, был положительным. Другими словами, фиксируем знаки таким образом, чтобы

$$a_k = \begin{cases} \lambda'_k, & s_k = t'_k, \quad s_k \neq \xi_0, \\ -\lambda''_k, & s_k = t''_k, \quad s_k \neq \xi_0, \\ \lambda'_k - \lambda''_k, & s_k = \xi_0. \end{cases}$$

Этого можно всегда добиться, потому что каноническая мера $\sigma_{\xi_0}$, включающая точку $\xi_0$, размещает (среди мер $\sigma \in V(c^0)$) максимальную массу в точке $\xi_0$ (см. § 4 гл. II).

Производя суммирование по частям, преобразуем (3.6) к виду

$$\sum_{k=1}^{n+1} B_k (w_i(s_k) - w_i(s_{k+1})) = -B_{n+2} w_i(s_{n+2}), \qquad i = 0, 1, \ldots, n, \tag{3.7}$$

где

$$B_k = \sum_{l=1}^{k} b_l, \qquad k = 1, 2, \ldots, n+2. \tag{3.8}$$

Утверждается, что величины $B_1, \ldots, B_{n+1}$ чередуются в знаке. Действительно, записав решение системы уравнений (3.7) относительно $B_k$ в виде отношений определителей (причем правая часть системы (3.7) считается известной), а затем выполнив некоторые элементарные преобразования, приходим к следующей формуле:

$$B_k = (-1)^{n-k} B_{n+2} \frac{W_k(s_1, \ldots, s_{n+2})}{W_{n+2}(s_1, \ldots, s_{n+2})},$$

где $W_k(s_1, \ldots, s_{n+2})$ обозначает определитель

$$W_k(s_1, \ldots, s_{n+2}) =$$
$$= \begin{vmatrix} w_0(s_1) - w_0(s_2), \ldots, w_0(s_{k-1}) - w_0(s_k), & w_0(s_{k+1}), & w_0(s_{k+2}), \ldots, w_0(s_{n+2}) \\ \cdots & \cdots & \cdots \\ w_n(s_1) - w_n(s_2), \ldots, w_n(s_{k-1}) - w_n(s_k), & w_n(s_{k+1}), & w_n(s_{k+2}), \ldots, w_n(s_{n+2}) \end{vmatrix}.$$

Используя следствие 6.1 гл. I, заключаем, что величины $B_1$, $B_2, \ldots, B_{n+1}$ строго чередуются в знаке.

Теперь пусть индекс, соответствующий $\xi_0$, будет равен $k_0$, т. е. $s_{k_0} = \xi_0$. Очевидно,

$$a_{k_0} = B_{k_0} - B_{k_0-1} > 0, \tag{3.9}$$

так как каноническая мера имеет максимальную массу в точке $\xi_0$. Сравнивая полученный результат $B_{k_0} \cdot B_{k_0-1} < 0$ с неравенством (3.9), получим, что

$$B_{k_0} > 0 \text{ и } B_{k_0-1} < 0. \tag{3.10}$$

(Заметим, что для $k_0 = 1$ $B_0$ не было определено.)

Воспользовавшись соотношениями (3.10), (3.8), (3.5) и определением мер $\sigma_{\xi_0}$ и $\sigma$, получим

$$\int_a^{\xi_0^+} \Omega(t)\, d\sigma_{\xi_0}(t) - \int_a^{\xi_0^+} \Omega(t)\, d\sigma(t) > 0, \tag{3.11}$$

в то время как, если $k_0 \geqslant 2$, то

$$\int_a^{\xi_0^-} \Omega(t)\, d\sigma_{\xi_0}(t) - \int_a^{\xi_0^-} \Omega(t)\, d\sigma(t) < 0. \tag{3.12}$$

Другое доказательство теоремы 2.1. Согласно лемме 3.2 достаточно проверить (2.1) для любой меры $\sigma \in V(\mathbf{c}^0)$ индекса $(n+3)/2$. Но рассуждения, с помощью которых мы получили неравенства (3.11) и (3.12), показывают, что и первое из неравенств (2.1) в этом случае справедливо. Левое неравенство также имеет место, за исключением случая, когда $\xi_0$ является первым корнем меры $\sigma_{\xi_0}$. Этот особый случай может быть рассмотрен отдельно точно так же, как и в первоначальном доказательстве (ср. § 2).

## § 4. Обобщения

Материалы этого параграфа служат естественным дополнением результатов § 3. Пример 4.2, приведенный ниже, принадлежит фон Мизесу [1939], который использовал специальное геометрическое доказательство. Мы трактуем этот пример как простое применение теоремы Маркова — Крейна.

I. В первой части этого параграфа воспользуемся процессом возмущения или сглаживания, рассмотренным в § 3 гл. I (см. (3.16) гл. I), и попытаемся ослабить условие 6.1 гл. III теоремы 2.1.

Пусть $\{u_i\}_0^k$ — $WT$-система для $k = 0, 1, \ldots, n$. Для любого $\delta > 0$ определим функции

$$u_i(t;\delta) = \frac{1}{\sqrt{2\pi}\delta} \int_a^b \exp\left(-\frac{1}{2\delta^2}(t-\tau)^2\right) u_i(\tau)\, d\tau, \tag{4.1}$$

$i = 0, 1, \ldots, n$. Используя основную композиционную формулу (3.13) гл. I, находим, что $\{u_i(t;\delta)\}_0^k$ является $ET$-системой на интервале $[a, b]$ при $k = 0, 1, \ldots, n$.

Тот же процесс сглаживания показывает, что если $\{u_0, u_1, \ldots, u_r, \Omega\}$ — слабая $T$-система при $r = 0, 1, \ldots, n$ и $\Omega(t) \geqslant 0$ ($\Omega(t) \not\equiv 0$), то $\{u_0(t;\delta),\ u_1(t;\delta), \ldots, u_r(t;\delta),\ \Omega(t;\delta)\}$ является $ET$-системой для всех $r = 0, 1, \ldots, n$ и $\Omega(t;\delta) > 0$. Функция $\Omega(t;\delta)$ определяется, как в (4.1), только с заменой $u_i$ на $\Omega$. Для дальнейшего полезно отметить, что $u_i(t;\delta)$ и $\Omega(t;\delta)$ сходятся равномерно к $u_i(t)$ и $\Omega(t)$, соответственно, на любом интервале $[\alpha, \beta]$, где $a < \alpha < \beta < b$.

Пусть $\mathbf{c}^0$ будет внутренней точкой $\mathcal{M}_{n+1}$, и пусть мера $\sigma \in V(\mathbf{c}^0)$. Для любого интервала $[\alpha, \beta]$ видоизменим меру $\sigma$, поместив массы, содержащиеся в интервалах $[a, \alpha]$ и $[\beta, b]$, в точки $\alpha$ и $\beta$ соответственно. Полученную меру обозначим через $\sigma(t;\alpha,\beta)$ и, используя ее, образуем моментную точку с координатами

$$c_i^\delta(\alpha,\beta) = \int_a^b u_i(t;\delta)\,d\sigma(t;\alpha,\beta), \qquad i = 0, 1, \ldots, n. \tag{4.2}$$

По отношению к моментному пространству $\mathcal{M}_{n+1}^\delta(\alpha,\beta)$, образованному системой $\{u_i(t;\delta)\}_0^n$ на интервале $[\alpha,\beta]$, мы теперь покажем, что точка с координатами (4.2) является внутренней точкой для всех $\delta$, если $\alpha$ и $\beta$ достаточно близки к граничным точкам $a$ и $b$. Это утверждение будет справедливым, если нам удастся доказать, что любая мера $\sigma \in V(\mathbf{c}^0)$ имеет индекс по крайней мере $(n+1)/2$. Заметим, что нельзя непосредственно применить теорему 2.1 гл. II, так как мы постулировали, что $\{u_i\}_0^n$ является $WT$-системой и не обязана быть $T$-системой. Однако если мера $\sigma$ имеет индекс, самое большее, $n/2$, то можно построить (по теореме 5.1 гл. I) многочлен $\sum_{i=0}^n a_i(\delta)\,u_i(t;\delta)$, удовлетворяющий условию $\sum_{i=0}^n [a_i(\delta)]^2 = 1$, который обращается в нуль на спектре меры $\sigma$. Если выбрать последовательность $\delta_k \to 0$ такую, что $a_i(\delta_k) \to a_i$, $i = 0, 1, \ldots, n$ $\left(\sum_{i=0}^n a_i^2 = 1\right)$, то предельный многочлен $\sum_{i=0}^n a_i u_i(t)$ будет обращаться в нуль на спектре меры $\sigma$, так как

$$\lim_{\delta\to 0} u_i(t,\delta) = \begin{cases} u_i(t), & \text{если } t \in (a,b), \\ \frac{1}{2} u_i(t), & \text{если } t = a \text{ или } b, \end{cases} \qquad i = 0, 1, \ldots, n.$$

Итак, многочлен $\sum_{i=0}^n a_i u_i$ определяет опорную плоскость к $\mathcal{M}_{n+1}$ во внутренней точке $\mathbf{c}^0$, что, очевидно, неверно. Следовательно, моментная точка (4.2) является внутренней точкой пространства $\mathcal{M}_{n+1}^\delta(\alpha, \beta)$ при $\alpha$ и $\beta$, достаточно близких к граничным точкам $a$ и $b$ соответственно. Так как функции $u_0(t;\delta), \ldots, u_n(t;\delta)$, $\Omega(t;\delta)$ удовлетворяют условию 6.1 гл. I, то из теоремы 2.1 можно заключить, что

$$\int_{\alpha}^{\xi_0^-} \Omega(t,\delta)\,d\sigma_{\xi_0}^{\delta}(t;\alpha,\beta) < \int_{\alpha}^{\xi_0^-} \Omega(t,\delta)\,d\sigma(t;\alpha,\beta) \leq$$

$$\leq \int_{\alpha}^{\xi_0^+} \Omega(t,\delta)\,d\sigma(t;\alpha,\beta) < \int_{\alpha}^{\xi_0^+} \Omega(t,\delta)\,d\sigma_{\xi_0}^{\delta}(t;\alpha,\beta),$$

где $\sigma_{\xi_0}^{\delta}(t;\alpha,\beta)$ обозначает каноническую меру, которая включает точку $\xi_0$ и удовлетворяет соотношениям (4.2).

Из непосредственного доказательства компактности (при $\delta\to 0$), использующего, что $\Omega(t;\delta)$ и $u_i(t;\delta)$ равномерно сходятся на $[\alpha,\beta]$ к $\Omega(t)$ и $u_i(t)$ соответственно, следует, что

$$\int_{\alpha}^{\xi_0^-} \Omega(t)\,d\sigma_{\xi_0}(t;\alpha,\beta) \leq \int_{\alpha}^{\xi_0^-} \Omega(t)\,d\sigma(t;\alpha,\beta) \leq$$

$$\leq \int_{\alpha}^{\xi_0^+} \Omega(t)\,d\sigma(t;\alpha,\beta) \leq \int_{\alpha}^{\xi_0^+} \Omega(t)\,d\sigma_{\xi_0}(t;\alpha,\beta), \qquad (4.3)$$

где

$$c_i(\alpha,\beta) = \int_{\alpha}^{\beta} u_i(t)\,d\sigma_{\xi_0}(t;\alpha,\beta), \qquad i = 0, 1, \ldots, n. \qquad (4.4)$$

Теперь заменим нижний предел в интегралах из неравенств (4.3) на граничную точку $a$ и выберем последовательности $\alpha_n \downarrow a$ и $\beta_n \uparrow b$ такие, что мера $\sigma_{\xi_0}(t;\alpha_n,\beta_n)$ сходится слабо к мере $\sigma_{\xi_0}(t)$. Тогда получим

$$\int_{a}^{\xi_0^-} \Omega(t)\,d\sigma_{\xi_0}(t) \leq \int_{a}^{\xi_0^-} \Omega(t)\,d\sigma(t) \leq \int_{a}^{\xi_0^+} \Omega(t)\,d\sigma(t) \leq \int_{a}^{\xi_0^+} \Omega(t)\,d\sigma_{\xi_0}(t), \quad (4.5)$$

где $\sigma_{\xi_0} \in V(c^0)$. Мера $\sigma_{\xi_0}$ имеет индекс либо $(n+1)/2$, либо $(n+2)/2$ и должна обладать максимальной массой в точке $\xi_0$ в соответствии со своей конструкцией. Заметим, что так как $\{u_i\}_0^n$ является только $WT$-системой, то мера с максимальной массой в точке $\xi_0$ может определяться неоднозначно.

Таким образом, мы доказали следующую теорему.

Теорема 4.1. *Пусть $\{u_i\}_0^k$ и $\{u_0, u_1, \ldots, u_k, \Omega\}$ являются $WT$-системами для $k = 0, 1, \ldots, n$, и предположим $\Omega(t) \geq 0$ для $t \in [a, b]$. Если $c^0 \in \operatorname{Int} \mathcal{M}_{n+1}$, то неравенство (4.5) выполняется для любой меры $\sigma \in V(c^0)$, где мера $\sigma_{\xi_0} \in V(c^0)$ имеет индекс $(n+1)/2$ или $(n+2)/2$ и характеризуется максимальной массой в точке $\xi_0$.*

II. Случай, когда $\Omega(t)$ является положительным $u$-многочленом, представляется важным. В этом случае справедливость неравенства (2.1) сохраняется, если предположить существование $n$ многочленов

$$\tilde{u}_k(t)=\sum_{j=0}^{n}\gamma_{k_j}u_j(t),\qquad k=0,1,\ldots,n-1, \tag{4.6}$$

таких, что функции

$$v_k(t)=\frac{d}{dt}\frac{\tilde{u}_k(t)}{\Omega(t)},\qquad k=0,1,\ldots,n-1, \tag{4.7}$$

образуют $T$-систему. (Подчеркнем, что число функций в системе (4.7) равно $n$, а не $n+1$.)

Рассуждение, использовавшееся при доказательстве леммы 3.1, здесь также применимо. Так как $\Omega(t)$ — $u$-многочлен, то имеем

$$\int_a^b \Omega\, d\sigma_{\xi_0}=\int_a^b \Omega\, d\sigma.$$

Пусть $f(\xi;\sigma_{\xi_0},\sigma)=\int_a^{\xi+}\Omega\, d\sigma-\int_a^{\xi+}\Omega\, d\sigma_{\xi_0}$.

Тогда

$$0=\int_a^b \tilde{u}_k\,(d\sigma_{\xi_0}-d\sigma)=\int_a^b v_k(t)\, f(t;\sigma_{\xi_0},\sigma)\, dt.$$

Отсюда следует, как и при доказательстве леммы 3.1, что $f(\xi;\sigma_{\xi_0},\sigma)$ имеет по крайней мере $n$ смен знака. Кроме того, так как мера $\sigma_{\xi_0}$, которая заменяет меру $\sigma_{\xi_0}(t;\gamma)$ в лемме 3.2, имеет индекс $(n+1)/2$ или $(n+2)/2$, то из интерпретации свойства смены знака (ср. рис. 8) следует, что неравенства (3.4) справедливы с заменой $\sigma_{\xi_0}(t;\gamma)$ на $\sigma_{\xi_0}(t)$. Как и прежде, левое неравенство может обратиться в равенство, если $\xi_0$ является первым корнем меры $\sigma_{\xi_0}$, и индекс меры $\sigma_{\xi_0}$ равен $(n+2)/2$. Из проведенных рассуждений следует

Теорема 4.2. *Пусть $\Omega(t)$ — положительный многочлен по $T$-системе $\{u_i\}_0^n$, и предположим, что существуют константы $\gamma_{k_j}$, $j=0,\ldots,n$, $k=0,\ldots,n-1$, такие, что система функций $\{v_i\}_0^{n-1}$, определенная в* (4.6) *и* (4.7), *является $T$-системой. Если $c^0\in \operatorname{Int}\mathcal{M}_{n+1}$ и $\sigma_{\xi_0}$ — каноническая мера, включающая точку $\xi_0\in(a,b)$, то для любой меры $\sigma\in V(c^0)$, которая отлична от меры $\sigma_{\xi_0}$, имеем*

$$\int_a^{\xi_0^-}\Omega(t)\, d\sigma_{\xi_0}<\int_a^{\xi_0^-}\Omega(t)\, d\sigma\le\int_a^{\xi_0^+}\Omega(t)\, d\sigma<\int_a^{\xi_0^+}\Omega(t)\, d\sigma_{\xi_0}, \tag{4.8}$$

*если только* $\xi_0$ *не является первым корнем меры* $\sigma_{\xi_0}$, *и индекс меры* $\sigma_{\xi_0}$ *не равен* $(n+2)/2$. *В этом особом случае первые два члена из* (4.8) *могут быть равны нулю.*

Пример 4.1. Ситуация, описанная выше, была подсказана рассмотрениями $T$-системы $u_i(t)=t^i$, $i=0,1,\ldots,n$. Если $\Omega(t)\equiv 1$, то, положив $\tilde{u}_k(t)=t^{k+1}/(k+1)$, $k=0,1,\ldots,n-1$, получим $v_k(t)=t^k$, $k=0,1,\ldots,n-1$. Для системы $\{t^i\}_0^n$ соотношения (4.8) сводятся к неравенствам

$$\sigma_{\xi_0}\{[a,\xi_0]\}<\sigma\{[a,\xi_0)\}\leq\sigma\{[a,\xi_0]\}<\sigma_{\xi_0}\{[a,\xi_0]\},\tag{4.9}$$

где $\sigma_{\xi_0}$ и $\sigma$ различны, $c_i^0=\int\limits_a^b t^i d\sigma(t)$, $i=0,1,\ldots,n$, а $\sigma_{\xi_0}$ — каноническая мера, соответствующая моментной точке $\mathbf{c}^0=(c_0^0,c_1^0,\ldots,c_n^0)$ и включающая точку $\xi_0$. Если мера $\sigma_{\xi_0}\{[a,\xi_0)\}$ равна нулю, то мера $\sigma\{[a,\xi_0)\}$ тоже может быть равна нулю.

III. Пример 4.1 может быть обобщен на произвольные $WT$-системы с тем, чтобы получить следующий результат.

Теорема 4.3. *Пусть* $\{u_i\}_0^k$ — $WT$-*система для* $k=0,1,\ldots,n$, *и предположим* $u_0(t)\geqslant 0$. *Если* $\mathbf{c}^0\in\operatorname{Int}\mathcal{M}_{n+1}$, *то неравенства* (4.5) *выполняются для любой меры* $\sigma\in V(\mathbf{c}^0)$, *где мера* $\sigma_{\xi_0}\in V(\mathbf{c}^0)$ *имеет индекс* $(n+1)/2$ *или* $(n+2)/2$ *и обладает максимальной массой в точке* $\xi_0$.

Доказательство. Определим $u_i(t;\delta)$ $(i=0,1,\ldots,n)$ в соответствии с (4.1) и положим

$$v_i(t;\delta)=\frac{d}{dt}\,\frac{u_{i+1}(t;\delta)}{u_0(t;\delta)},\quad i=0,1,\ldots,n-1.\tag{4.10}$$

Система функций $\{v_i(t;\delta)\}_0^{n-1}$ является $T$-системой на $[a,b]$. Доказательство этого факта приводится в гл. XI, где представлено подробное исследование таких систем (см. § 3 и теорему 1.2 гл. XI). Оставшаяся часть доказательства теперь следует доказательству теоремы 4.1, за исключением того, что вместо теоремы 2.1 здесь используется теорема 4.2.

Пример 4.2. Предположим, что $g(t)$ и $h(t)$ такие функции на $[a,b]$, что система $\{1,g(t),h(t)\}$ представляет собой $CT$-систему. Это предположение требует, в частности, чтобы функция $g(t)$ была строго возрастающей, и, кроме того, чтобы кривая

$$\Gamma=\{(g(t),h(t))\,|\,t\in[a,b]\},\tag{4.11}$$

была выпуклой. Следовательно, любой отрезок прямой на плоскости пересекает кривую $\Gamma$, самое большее, дважды.

Рассмотрим задачу нахождения точных границ для значений интеграла

$$\int\limits_a^{\xi_0}d\sigma,\tag{4.12}$$

где $\sigma$ — конечная мера на $[a, b]$, удовлетворяющая условиям

$$\int_a^b d\sigma(t) = 1, \quad \int_a^b g(t)\, d\sigma(t) = g_0, \quad \int_a^b h(t)\, d\sigma(t) = h_0, \qquad (4.13)$$

т. е. $\sigma \in V = V(1, g_0, h_0)$. Предполагается, что точка $(1, g_0, h_0)$ лежит внутри моментного пространства $\mathcal{M}_3$, образованного функциями $1$, $g(t)$ и $h(t)$, так что множество $V$ не вырождается в единственную меру.

Обращаясь к теореме 4.3, заключаем, что

$$\int_a^{\xi_0^-} d\sigma_{\xi_0}(t) \le \int_a^{\xi_0^-} d\sigma(t) \le \int_a^{\xi_0^+} d\sigma(t) \le \int_a^{\xi_0^+} d\sigma_{\xi_0}(t), \quad \sigma \in V, \qquad (4.14)$$

где $\sigma_{\xi_0}$ — каноническое представление точки $(1, g_0, h_0)$, сконцентрировавшее массу в точке $\xi_0$.

В данном случае граничные значения в неравенствах (4.14) могут быть вычислены точно. Нижнее главное представление $\underline{\sigma}$ имеет массу в граничной точке $a$ и еще в одной другой точке $t_2 \in (a, b)$. Точка $t_2$ определяется продолжением отрезка прямой, соединяющего точки $(g(a), h(a))$ и $(g_0, h_0)$, до пересечения с кривой $\Gamma$ в точке $(g(t_2), h(t_2))$. Нижняя точка сосредоточения массы $t_1$ верхней главной меры $\overline{\sigma}$ характеризуется условием, что прямая, соединяющая точки $(g(b), h(b))$ и $(g_0, h_0)$, пересекает кривую $\Gamma$ в точке $(g(t_1), h(t_1))$.

Теперь всякий раз, когда $a < \xi_0 < t_1$ $(t_2 < \xi_0 < b)$, каноническая мера $\sigma_{\xi_0}$ имеет массу в точке $\xi_0$ и еще в одной другой точке $\eta$, удовлетворяющей условию $t_2 < \eta < b$ $(a < \eta < t_1)$, в то время как в случае $t_1 < \xi_0 < t_2$ мера $\sigma_{\xi_0}$ размещает массу в точке $\xi_0$ и в обеих граничных точках $a$ и $b$. Всякий раз, когда $\xi_0$ выбирается совпадающей с одной из точек $a$, $t_1$, $t_2$ или $b$, то в качестве меры $\sigma_{\xi_0}$ выбирается главное представление.

Соотношения (4.14) сводятся к следующим:

$$0 \le \sigma\{[a, \xi_0]\} \le 1 - \frac{g_0 - g(\xi_0)}{g(\eta) - g(\xi_0)}, \quad a < \xi_0 < t_1, \qquad (4.15)$$

$$\frac{g_0 - g(\xi_0)}{g(\eta) - g(\xi_0)} \le \sigma\{[a, \xi_0)\} \le 1, \quad t_2 < \xi_0 < b, \qquad (4.16)$$

и

$$p \le \sigma\{[a, \xi_0)\} \le \sigma\{[a, \xi_0]\} \le p + q, \quad t_1 < \xi_0 < t_2, \qquad (4.17)$$

где

$$p = D^{-1} \begin{vmatrix} g_0 - g(b) & g(\xi_0) - g(b) \\ h_0 - h(b) & h(\xi_0) - h(b) \end{vmatrix}, \quad q = D^{-1} \begin{vmatrix} g(a) - g(b) & g_0 - g(b) \\ h(a) - h(b) & h_0 - h(b) \end{vmatrix},$$

$$D = \begin{vmatrix} g(a) - g(b) & g(\xi_0) - g(b) \\ h(a) - h(b) & h(\xi_0) - h(b) \end{vmatrix}.$$

## § 5. Множества мер, заданных моментными неравенствами

Содержание этого параграфа заключается в расширении класса моментных неравенств. Начало исследованиям в этом направлении было положено Розенблюмом [1951]. В §§ 1—4 мы имели дело с задачей нахождения экстремальных значений для некоторого функционала, заданного на множестве мер $V(\mathbf{c}^0)$, соответствующем единственной точке $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}$. В этом параграфе рассмотрим связанные с этими вопросами задачи, когда ограниченное множество состоит из всех таких мер, моменты которых по отношению к системе функций $\{u_i\}_0^n$ принадлежат некоторым подмножествам $\mathcal{M}_{n+1}$. Канонические меры здесь также играют важную роль.

Пусть $u_i$, $0 \leqslant i \leqslant n+1$, — непрерывные функции на конечном интервале $[a, b]$, удовлетворяющие следующему условию.

Условие 5.1.

(а) $\{u_i\}_0^n$ и $\{u_i\}_0^{n+1}$ — $T$-системы;

(b) для $i = 0, 1, \ldots, n+1$ определители $n+1$-го порядка, образованные функциями $u_0, \ldots, u_{i-1}, \ u_{i+1}, \ldots, u_{n+1}$, имеют все один строго определенный знак $\varepsilon_i$, т. е. существует число $\varepsilon_i = +1$ или $-1$ такое, что

$$\varepsilon_i \begin{vmatrix} u_0(t_0) & u_0(t_1) & \cdots & u_0(t_n) \\ \cdot & \cdot & \cdot & \cdot \\ u_{i-1}(t_0) & u_{i-1}(t_1) & \cdots & u_{i-1}(t_n) \\ u_{i+1}(t_0) & u_{i+1}(t_1) & \cdots & u_{i+1}(t_n) \\ \cdot & \cdot & \cdot & \cdot \\ u_{n+1}(t_0) & u_{n+1}(t_1) & \cdots & u_{n+1}(t_n) \end{vmatrix} > 0, \quad i = 0, 1, \ldots, n+1$$

для всех $a \leqslant t_0 < t_1 < \ldots < t_n \leqslant b$,

с) $u_i(t) > 0, \ t \in [a, b], \ i = 0, 1, \ldots, n+1$.

Теорема 5.1. *Предположим, что для системы функций $\{u_i\}_0^{n+1}$ выполняется условие 5.1, и $V$ обозначает класс мер на $[a, b]$, удовлетворяющих условию*

$$\int_a^b u_i d\sigma \begin{cases} = d_i, & i \in A, \\ \leqslant d_i, & i \in B, \\ \geqslant d_i, & i \in C, \end{cases}$$

*где $A$, $B$ и $C$ образуют разбиение множества $\{0, 1, \ldots, n\}$, а $d_i$ — заданное множество положительных констант.*

*Если $V$ не пусто, то*

(1) $\underline{\alpha} = \min\limits_{\sigma \in V} \int u_{n+1} d\sigma$

*и минимум достигается единственным образом в $V$ на мере $\sigma$*

*индекса, не большего чем* $(n+1)/2$, *которая не включает точку* $b$, *если индекс равен* $(n+1)/2$.

(2) *Если* $A \cup B$ *не пусто, то*

$$\bar{\alpha} = \max_{\sigma \in V} \int u_{n+1} d\sigma < \infty,$$

*и максимум достигается единственным образом в* $V$ *на мере* $\bar{\sigma}$ *индекса, не большего чем* $(n+1)/2$, *которая включает точку* $b$, *если индекс равен* $(n+1)/2$.

Доказательство. Рассмотрим только вторую часть теоремы. Так как $A \cup B$ не пусто, то $\int u_i d\sigma \leq d_i$ для некоторого $i \in A \cup B$ и всех мер $\sigma \in V$. Учитывая условие 5.1 (с), легко видеть, что $\int d\sigma \leq k$ для всех мер $\sigma \in V$. Так как функция $u_{n+1}(t)$ ограничена на $[a, b]$, то отсюда следует, что $\bar{\alpha} < \infty$. Теперь можно определить последовательность $\sigma^{(k)}$ такую, что $\lim_{k \to \infty} \int u_{n+1} d\sigma^{(k)} = \bar{\alpha}$. Применение теоремы выбора Хелли позволяет установить тот факт, что значение $\bar{\alpha}$ достигается на некоторой мере $\bar{\sigma} \in V$. Теорема 1.1 и свойство, согласно которому граничные точки в $\mathcal{M}_{n+1}$ обладают единственными представлениями индекса, не большего чем $n/2$ (см. теорему 2.1 гл II), показывают, что мера $\bar{\sigma}$ имеет индекс, не превышающий $(n+1)/2$, а если этот индекс равен $(n+1)/2$, то $\bar{\sigma}$ совпадает с верхним главным представлением.

Единственность меры $\bar{\sigma}$, доставляющей минимальное значение $\bar{\alpha}$, доказывается с помощью леммы 5.1, которая приводится ниже.

Прежде всего определим

$$\underline{c}_i = \underline{c}_i(c_0, \ldots, c_{i-1}, c_{i+1}, \ldots, c_{n+1}) = \inf_{\sigma \in V_i(\mathrm{c})} \int u_i d\sigma,$$

$$\bar{c}_i = \bar{c}_i(c_0, \ldots, c_{i-1}, c_{i+1}, \ldots, c_{n+1}) = \sup_{\sigma \in V_i(\mathrm{c})} \int u_i d\sigma,$$

где $V_i(\mathrm{c}) = \{\sigma \mid \int u_j d\sigma = c_j,\ 0 \leq j \leq n+1,\ j \neq i\}$.

Лемма 5.1. *Пусть для функций* $u_i$, $0 \leq i \leq n+1$ *выполняются условия* 5.1 (а) *и* (с). *Если для фиксированных* $i$ *и* $j$ *условие* (b) *также удовлетворяется, то величины* $(-1)^{j-i} \varepsilon_i \varepsilon_j \underline{c}_i$ *и* $(-1)^{j-i} \varepsilon_i \varepsilon_j \bar{c}_i$ *являются строго убывающими функциями координат* $c_j$, $j \neq i$.

Доказательство. Докажем только утверждение, касающееся $\underline{c}_i$. Легко может быть показано, что $\underline{c}_i$ — конечно и минимум достигается для некоторой меры $\mu$ индекса, не большего чем $(n+1)/2$. Так как $\{u_0, u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, \varepsilon_i u_{n+1}\}$ и $\{u_0, \ldots, u_{i-1}, u_{i+1}, \ldots, \varepsilon_i u_{n+1}, (-1)^{n-i+1} \varepsilon_i u_i\}$ представляют собой $T$-системы, то мера $\mu$ всякий раз, когда ее индекс равен $(n+1)/2$, является верхним главным представлением при $(-1)^{n-i+1} \varepsilon_i = -1$ и нижним главным представлением при $(-1)^{n-i+1} \varepsilon_i = 1$. Теперь предположим, что все координаты $c_0, \ldots, c_{i-1}, c_{i+1}, \ldots, c_{n+1}$, определяющие $\underline{c}_i$, фиксированы, за исключением $j$-й координаты $(j \neq i)$, которая возрастает до $c'_j$. Пуст $\mu_1$ — экстремальная мера, которая доставляет новое минимальное значение $\underline{c}'_i$. Несложное рассуждение показывает, что общее число точек сосредоточения масс мер $\mu$ и $\mu_1$ не превышает $n+1$. Пусть

$$\int u_k d\mu - \int u_k d\mu_1 = \sum_{p=1}^{n+1} \gamma_p u_k(s_p), \quad k = 0, 1, \ldots, n+1, \tag{5.1}$$

где последовательность $\{s_p\}_1^{n+1}$ состоит из точек сосредоточения масс мер $\mu$ и $\mu_1$, расположенных в возрастающем порядке, и, если необходимо, добавляются дополнительные точки, с тем, чтобы получить в сумме $n+1$ точку. Коэффициенты $\gamma_p$ $(p = 1, \ldots, n+1)$ служат соответствующими весами с надлежащими знаками ($\gamma_p = 0$ для любой из дополнительных точек.) Обращаясь к равенству (5.1), видим, что первый столбец в написанном ниже определителе представляет собой линейную комбинацию остальных столбцов, так что

$$\begin{vmatrix} 0 & u_0(s_1) & u_0(s_2) & \ldots & u_0(s_{n+1}) \\ 0 & u_1(s_1) & u_1(s_2) & \ldots & u_1(s_{n+1}) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_j - c'_j & & & & \\ 0 & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \underline{c}_i - \underline{c}'_i & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & u_{n+1}(s_1) & u_{n+1}(s_2) & \ldots & u_{n+1}(s_{n+1}) \end{vmatrix} = 0.$$

Разлагая этот определитель по первому столбцу и вспоминая, что алгебраические дополнения ненулевых элементов этого столбца имеют знаки $\varepsilon_j$ и $\varepsilon_i$ соответственно, можно вывести свойство монотонности, сформулированное в лемме.

Для того чтобы установить утверждение теоремы 5.1 об единственности, предположим, напротив, что две меры $\bar{\sigma}$ и $\sigma$ из класса $V$ доставляют максимальное значение $\bar{\alpha}$ интегралу. Так как для фиксированного множества моментов $\bar{\alpha}$ порождается единственной мерой, то отсюда следует, что $\int u_j d\bar{\sigma} \neq \int u_j d\sigma$ для неко орого $j = 0, 1, \ldots, n$. Однако если одна координата отличается, то, согласно лемме 5.1, мы можем увеличить значение $\bar{\alpha}$, что противоречит смыслу $\bar{\alpha}$. Доказательство теоремы 5.1 на этом завершается.

Сохранив обозначение теоремы 5.1, положим

$$\underline{D} = \left\{ i \mid \int u_i d\underline{\sigma} = d_i, \quad 0 \leq i \leq n \right\},$$
$$\overline{D} = \left\{ i \mid \int u_i d\bar{\sigma} = d_i, \quad 0 \leqslant i \leqslant n \right\}.$$

Следствие 5.1а. *Пусть* $V_1$ $(V_1 \supset V)$ *представляет собой класс мер* $\sigma$, *удовлетворяющих условию*

$$\int u_i d\sigma \begin{cases} = d_i, & i \in A \cap \underline{D}, \\ \leqslant d_i, & i \in B \cap \underline{D}, \\ \geqslant d_i, & i \in C \cap \underline{D}. \end{cases}$$

*Тогда экстремальные меры* $\underline{\sigma}$ *и* $\bar{\sigma}$, *фигурирующие в теореме* 5.1, *обладают следующими свойствами:*

(1) $\underline{\sigma}$ — *единственная мера в* $V_1$, *на которой достигается значение*

$$\underline{\beta} = \min_{\sigma \in V_1} \int u_{n+1}\, d\sigma;$$

(2) *если* $A \cup B$ *не пусто, то* $\bar{\bar{\sigma}}$ — *единственная мера в* $V_2$, *на которой достигается значение*

$$\bar{\beta} = \max_{\sigma \in V_2} \int u_{n+1}\, d\sigma,$$

*где* $V_2$ *определяется аналогично* $V_1$ *с использованием* $\bar{D}$ *вместо* $\underline{D}$.

Доказательство. Докажем только вторую часть теоремы. Если $A \cup B$ не пусто, а $(A \cup B) \cap \bar{D}$ пусто, то $A$ и подавно пусто и $\int u_i d\bar{\sigma} < d_i$ для всех $i \in B$. В этом случае $\lambda\bar{\sigma} \in V$ при $1 + \varepsilon > \lambda > 1$ ($\varepsilon$ достаточно малое) и, следовательно, мера $\bar{\sigma}$ не доставляет максимум интегралу $\int u_{n+1} d\sigma$ в классе мер $V$. Далее, если $(A \cup B) \cap \bar{D}$ не пусто, то легко показать, по аналогии с доказательством теоремы 5.1, что значение $\bar{\beta}$ достигается на некоторой мере. Пусть $\sigma$ представляет собой любую меру, которая доставляет максимальное значение $\bar{\beta}$ интегралу $\int u_{n+1} d\sigma$. Эта мера должна иметь индекс, не больший чем $(n+1)/2$, и соответствовать верхнему главному представлению, если индекс равен $(n+1)/2$. Если $i \in \bar{D}$, то, согласно лемме 5.1, $\int u_i d\bar{\sigma} = \int u_i d\sigma$. Далее, если $\sigma \in V$, то $\bar{\beta} = \bar{\alpha}$ и $\sigma = \bar{\sigma}$. С другой стороны, если $\sigma \notin V$, то для некоторого $i \notin \bar{D}$ и $i \in B \cup C$ выполняется $\int u_i d\bar{\sigma} < d_i < \int u_i d\sigma$, когда $i \in B$, и знаки неравенства меняют направление, когда $i \in C$. По лемме 5.1 мы можем, следовательно, увеличить одно из значений $\bar{\beta}$ или $\bar{\alpha}$, что, очевидно, невозможно. Поэтому $\int u_i d\sigma = \int u_i d\bar{\sigma}$ для $i = 0, 1, \ldots, n$, что означает $\sigma = \bar{\sigma}$.

Следствие 5.1b. *Пусть* $\bar{\sigma}$, $\underline{\sigma}$, $A$, $B$, $C$, $\underline{D}$ *и* $\bar{D}$ *те же, что и в следствии* 5.1a, *и пусть* $\{v_0, \ldots, v_n\}$ *заданы так, что* $\{v_0, \ldots, v_n, u_{n+1}\}$ *удовлетворяют условию* 5.1. *Если* $e_i = \int v_i d\underline{\sigma}$ и $f_i = \int v_i d\bar{\sigma}$ *при* $i = 0, 1, \ldots, n$, *то* $\underline{\sigma}\,(\bar{\sigma})$ — *единственная мера в* $W_1\,(W_2)$, *на которой достигается значение*

$$\underline{\gamma} = \min_{\sigma \in W_1} \int u_{n+1} d\sigma \quad \left(\bar{\gamma} = \max_{\sigma \in W_2} \int u_{n+1} d\sigma\right),$$

*где* $W_1$ — *класс мер* $\sigma$, *удовлетворяющих условию*

$$\int v_i d\sigma \begin{cases} = e_i, & i \in A \cap \underline{D}, \\ \leqslant e_i, & i \in B \cap \underline{D}, \\ \geqslant e_i, & i \in C \cap \underline{D}, \end{cases}$$

*а* $W_2$ — *класс мер* $\sigma$, *удовлетворяющих условию*

$$\int v_i d\sigma \begin{cases} = f_i, & i \in A \cap \bar{D}, \\ \leqslant f_i, & i \in B \cap \bar{D}, \\ \geqslant f_i & i \in C \cap \bar{D}, \end{cases}$$

Доказательство. Докажем только утверждение, касающееся $\bar{\sigma}$. Как и при доказательстве следствия 5.1a, можно показать, что $\bar{\gamma} < \infty$, и это значение достигается на некоторой мере $\sigma$ индекса, не большего чем $(n+1)/2$. Если $\int v_i d\sigma \neq \int v_i d\bar{\sigma}$ для некоторого $i = 0, 1, \ldots, n$, то мы можем изменить любую из этих величин, увеличив при этом одно из значений $\bar{\gamma}$ или $\bar{\beta}$. Следовательно, $\int v_i d\sigma = \int v_i d\bar{\sigma}$ для $i = 0, 1, \ldots, n$, что означает $\sigma = \bar{\sigma}$.

Предположим теперь, что имеется возможность присоединить к системе функций $\{u_0, \ldots, u_{n+1}\}$ другую непрерывную функцию $v_i$, удовлетворяющую следующему условию.

Условие 5.2.

(a) $\{u_0, \ldots, u_i, v_i, u_{i+1}, \ldots, u_n\}$ и $\{u_0, \ldots, u_i, v_i, u_{i+1}, \ldots, u_{n+1}\}$ — $T$-системы;

(b) определители $n+1$-го порядка, образованные из любого упорядоченного множества $n+1$ функций, выбранных из системы $\{u_0, \ldots, u_i, v_i, u_{i+1}, \ldots, u_n\}$, имеют строго определенный знак (см. условие 5.1 (b)).

Теорема 5.2. *Если функции* $u_0, \ldots, u_i, v_i, u_{i+1}, \ldots, u_{n+1}$ *удовлетворяют условиям* 5.1 *и* 5.2, *а* $V$, $\underline{\sigma}$ *и* $\bar{\sigma}$ *определены как в теореме* 5.1, *то при условии, что* $A \cup B$ *не пусто, экстремумы*

$$\underline{\delta}_i = \min_{\sigma \in V} (-1)^{n-1} \int v_i d\sigma \ \text{и} \ \bar{\delta}_i = \max_{\sigma \in V} (-1)^{n-i} \int v_i d\sigma$$

*реализуются единственным образом в $V$ на мерах $\underline{\sigma}$ и $\bar{\sigma}$ соответственно.*

Доказательство. Ограничимся рассмотрением только величины $\underline{\delta}_i$. Если $\underline{\delta}_i < (-1)^{n-i} \int v_i d\underline{\sigma}$, то применение леммы 5.1 к функциям $u_0, \ldots, u_i, v_i, u_{i+1}, \ldots, u_{n+1}$ показывает, что величина $\underline{\alpha}$, определенная в теореме 5.1, могла бы быть уменьшена. Пусть $\bar{\sigma}$ обозначает другую меру ($\sigma \neq \underline{\sigma}$), которая доставляет значение $\underline{\delta}_i$. Тогда, так как $\{u_0, \ldots, u_n\}$ и $\{u_0, \ldots, u_n, (-1)^{n-i} v_i\}$ — $T$-системы, то мера $\sigma$ имеет индекс, не больший чем $(n+1)/2$, и соответствует нижнему главному представлению, если этот индекс равен $(n+1)/2$. Следовательно, $\int u_i d\sigma \neq \int u_i d\underline{\sigma}$ для некоторого $i = 0, \ldots, n$. В таком случае, применяя лемму 5.1 к системе $\{u_0, \ldots, u_i, v_i, u_{i+1}, \ldots$ $\ldots, u_n\}$, можно уменьшить значение $\delta_i$, что приводит к противоречию с его определением.

## § 6. Крайние точки множества функций распределения с заданными моментами

В конечномерных евклидовых пространствах вещественная линейная функция, определенная на ограниченном замкнутом выпуклом множестве достигает своего максимума, по меньшей мере, в крайней точке выпуклого множества. Так как

$$L(\sigma) = \int_a^b \Omega(t)\, d\sigma(t)$$

представляет собой линейный функционал по $\sigma$, то экстремальные значения функционала $L(\sigma)$ для $\sigma \in V(\mathbf{c})$ должны также доставляться крайними точками множества $V(\mathbf{c})$. В этом параграфе мы охарактеризуем крайние точки множества $V(\mathbf{c})$ и рассмотрим некоторые смежные вопросы.

Пусть $\{u_i\}_0^n$ — $T$-система на $[a, b]$, и определим $\mathscr{D}(\mathbf{c})$ как подкласс мер $\sigma \in V(\mathbf{c})$, имеющих конечный спектр. Пусть $S(\sigma)$ для $\sigma \in \mathscr{D}(\mathbf{c})$ обозначает число точек в спектре меры $\sigma$.

Теорема 6.1. *Крайними точками $\mathscr{E}(\mathbf{c})$ множества $\mathscr{D}(\mathbf{c})$ служат те меры $\sigma$, для которых $S(\sigma) \leqslant n+1$.*

Доказательство. Сначала предположим, что $\sigma \in \mathscr{D}(\mathbf{c})$ и $S(\sigma) > n+1$. Положив

$$I(t) = \begin{cases} 1, & t \geqslant 0, \\ 0, & t < 0, \end{cases}$$

запишем

$$\sigma(t) = \sum_{j=1}^{S(\sigma)} \lambda_j I(t - t_j),$$

где

$$\sum_{j=1}^{S(\sigma)} \lambda_j u_i(t_j) = c_i, \qquad i = 0, 1, \ldots, n, \tag{6.1}$$

и $\lambda_j > 0$ для $j = 1, \ldots, S(\sigma)$. Для достаточно малого и положительного $\varepsilon$ можно записать

$$\sigma = \frac{\sigma + \varepsilon\mu}{2} + \frac{\sigma - \varepsilon\mu}{2},$$

где $\mu(t) = \sum_{j=1}^{S(\sigma)} \delta_j I(t - t_j)$ и $\delta_j$ ($\delta_j$ не все равны нулю) удовлетворяют соотношениям

$$\sum_{j=1}^{S(\sigma)} \delta_j u_i(t_j) = 0, \qquad i = 0, 1, \ldots, n.$$

Другими словами, $\mu$ — обобщенная мера, которая сосредоточивает значения $\delta_j$ в точках $t_j$, $j = 1, 2, \ldots, S(\sigma)$. Существование нетривиального решения, очевидно, гарантировано, так как, по предположению, число переменных $S(\sigma)$ превышает число уравнений $n + 1$. Ясно, что меры $\sigma + \varepsilon\mu$ и $\sigma - \varepsilon\mu$ различны и обе принадлежат классу $\mathcal{D}(c)$ для достаточно малого $\varepsilon$. Таким образом, мы представили меру $\sigma$ в виде выпуклой комбинации двух различных элементов класса $\mathcal{D}(c)$, поэтому $\sigma$ не является крайней точкой.

Чтобы установить обратное, заметим, что представление меры $\sigma = \alpha\sigma_1 + (1 - \alpha)\sigma_2$ означает, что спектры мер $\sigma_1$ и $\sigma_2$ содержатся в спектре $\sigma$. Однако если $S(\sigma) \leqslant n + 1$, то решение системы уравнений (6.1) единственно, так что, если $S(\sigma_1) < S(\sigma)$, то $\alpha=0$ и $\sigma_2 = \sigma$, в то время как, если $S(\sigma_1) = S(\sigma)$, то $\sigma_1 = \sigma$. Следовательно, если $S(\sigma) \leqslant n + 1$, то мера $\sigma$ является крайней точкой множества $\mathcal{D}(c)$, что завершает доказательство.

Любая мера $\sigma \in \mathcal{D}(c)$ может быть представлена в виде

$$\sigma(t) = \sum_{j=1}^{S(\sigma)} \lambda_j I(t - t_j).$$

Если $S(\sigma) \geqslant n + 1$, то размерность пространства решений $\lambda_1, \lambda_2, \ldots, \lambda_{S(\sigma)}$ системы уравнений

$$\sum_{j=1}^{S(\sigma)} \lambda_j u_i(t_j) = c_i, \qquad i = 0, 1, \ldots, n, \tag{6.2}$$

очевидно, равна $S(\sigma) - (n+1)$. Если $u(t) = \sum_{i=0}^{n} a_i u_i(t)$ — положительный многочлен, то

$$\left(\min_{t \in [a,b]} u(t)\right) \int_a^b d\sigma \leq \int_a^b u\, d\sigma = \sum_{i=0}^{n} a_i c_i$$

для всех $\sigma \in \mathcal{D}(\mathbf{c})$ (в действительности, для всех $\sigma \in V(\mathbf{c})$). Следовательно, множество неотрицательных решений системы уравнений (6.2) выпукло, замкнуто и ограничено. Это множество натянуто на свои крайние точки, и каждая крайняя точка допускает, самое большее, $n+1$ точек сосредоточения масс по теореме 6.1. Кроме того, любая мера $\sigma \in \mathcal{D}(\mathbf{c})$ представима в виде

$$\sigma(t) = \sum_{i=1}^{k} \alpha_i \sigma_i(t), \qquad \alpha_i > 0, \tag{6.3}$$

где $S(\sigma_i) \leqslant n+1$, $i = 1, \ldots, k$, и $\sum_{i=1}^{k} \alpha_i = 1$. Тем самым доказана следующая теорема.

Теорема 6.2. *Множество $\mathcal{D}(\mathbf{c})$ натянуто на свои крайние точки.*

Обобщение этого результата на выпуклое множество $V(\mathbf{c})$ всех мер $\sigma$, соответствующих моментной точке $\mathbf{c} = (c_0, c_1, \ldots, c_n)$, требует более глубокого анализа, чем тот, который был приведен выше для множества $\mathcal{D}(\mathbf{c})$. Необходимо отметить, что под термином «натянуто» в формулировке теоремы 6.2 мы подразумеваем, что любая мера $\sigma \in \mathcal{D}(\mathbf{c})$ может быть записана в виде конечной выпуклой комбинации крайних точек, как это сделано в выражении (6.3). Представление для меры $\sigma \in \mathcal{D}(\mathbf{c})$, приведенное в выражении (6.3), может быть записано в виде

$$\sigma(t) = \int_{0}^{1} \sigma_\alpha(t)\, d\alpha, \tag{6.4}$$

где $\sigma_\alpha(t) = \sigma_i(t)$ на интервале длины $\alpha_i$, $i = 1, 2, \ldots, k$. Можно показать, но мы не будем здесь этого делать, что выражение (6.4) справедливо для любой непрерывной справа меры $\sigma \in V(\mathbf{c})$, где для любого $\alpha$ $\sigma_\alpha \in \mathfrak{E}(\mathbf{c})$ является непрерывной справа ступенчатой функцией, самое большее, с $n+1$ скачками, а интеграл в правой части выражения (6.4) существует в смысле интеграла Лебега.

Соотношение (6.4) выполняется в ситуациях гораздо более общих, чем те, которые мы здесь рассматривали. Действительно, ни чебышевские свойства системы функций $\{u_i\}_0^n$, ни конечность интервала $[a, b]$ не являются существенными для справедливости формулы (6.4). Модифицированная формула представления типа (6.4) принадлежит Малголланду и Роджерсу [1958]. Эти вопросы, хотя и в более узкой постановке, связаны с некоторыми работами Чокета [1956], Мейера [1964] и других авторов, посвященными краевым задачам теории потенциалов.

Мы докажем элементарными методами следующую теорему.

Теорема 6.3. *Если $\{u_i\}_0^n$ — $T$-система, то множество $V(\mathbf{c})$ служит слабым замыканием множества $\mathcal{D}(\mathbf{c})$, т. е. для любой меры $\sigma \in V(\mathbf{c})$ существует последовательность мер $\sigma_k \in \mathcal{D}(\mathbf{c})$,*

$k = 1, 2, \ldots,$ *такая, что*

$$\lim_{k\to\infty}\int_a^b g(t)\,d\sigma_k(t) = \int_a^b g(t)\,d\sigma(t) \tag{6.5}$$

*для всех непрерывных функций* $g$, *т. е. для* $g \in C[a, b]$.

Доказательство. Без потери общности можно предположить, что $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$, так как в противном случае множество $V(\mathbf{c})$ сводится к единственной мере с конечным спектром (см. теорему 2.1 гл. II).

Пусть $m = \int_a^b d\sigma(t)$, и определим

$$t_i^{(k)} = \inf\left\{t \mid \sigma\{[a, t]\} \geqslant \frac{mi}{k}\right\}, \qquad i = 0, 1, \ldots, k$$

(заметим, что $t_0^{(k)} = a$). Тогда последовательность мер, заданная как

$$\mu_k(t) = \begin{cases} \dfrac{mi}{k}, & t_{i-1}^{(k)} \leq t < t_i^{(k)}, \quad i = 1, \ldots, k, \\ m, & t_k^{(k)} \leq t \leq b, \end{cases}$$

удовлетворяет условию $\lim\limits_{k\to\infty}\int_a^b g(t)\,d\mu_k(t) = \int_a^b g(t)\,d\sigma(t)$ для всех функций $g \in C[a, b]$. Отметим, что меры $\mu_k$ не обязательно принадлежат множеству $V(\mathbf{c})$. Для каждого $k$ затем определим

$$c_i^{(k)} = \int_a^b u_i(t)\,d\mu_k(t), \qquad i = 0, 1, \ldots, n,$$

а после этого зададим $\alpha_k$ и $\mathbf{b}^{(k)}$ соотношениями

$$(1 + \alpha_k)\,\mathbf{c} - \alpha_k \mathbf{c}^{(k)} = \mathbf{b}^{(k)}, \qquad \alpha_k > 0, \quad |\mathbf{b}^{(k)} - \mathbf{c}| = \varepsilon. \tag{6.6}$$

Здесь число $\varepsilon$ задается достаточно малым, таким, что $\varepsilon$-окрестность точки $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$ содержится в $\mathcal{M}_{n+1}$.

Так как $\mathbf{b}^{(k)}$ выражается через меру $\rho_k$ с конечным спектром, то видим, что мера

$$\sigma_k = \frac{\rho_k + \alpha_k \mu_k}{1 + \sigma_k}$$

принадлежит множеству $\mathcal{D}(\mathbf{c})$.

Пусть $u(t) = \sum\limits_{i=0}^{n} a_i u_i(t)$ — строго положительный многочлен. Так как моментные точки $\mathbf{b}^{(k)}$ ограничены, то получим неравенство

$$K \geqslant \sum_{i=0}^{n} a_i b_i^{(k)} = \int_a^b u(t)\,d\rho_k \geqslant \left(\min_{a \leqslant t \leqslant b} u(t)\right) \int_a^b d\rho_k,$$

которое показывает, что интегралы $\int_a^b d\rho_k$ равномерно ограничены. Так как $\lim_{k\to\infty} \mathrm{c}^{(k)} = \mathrm{c}$, то на основании (6.6) можно также прийти к выводу, что $\alpha_k \to \infty$. Объединяя эти утверждения, легко получить, что

$$\lim_k \int_a^b g(t)\, d\sigma_k(t) = \int_a^b g(t)\, d\sigma(t)$$

для всех функций $g \in C[a, b]$. Следовательно, множество $V(\mathrm{c})$ является слабым замыканием множества $\mathcal{D}(\mathrm{c})$.

С помощью вышеприведенной теоремы легко доказывается следующее утверждение.

Теорема 6.4. *Если $g$ — непрерывная функция на $[a, b]$ и точка $\mathrm{c} \in \mathcal{M}_{n+1}$, то*

$$\sup_{\sigma\in V(\mathrm{c})} \int_a^b g\, d\sigma = \sup_{\sigma\in \mathfrak{E}(\mathrm{c})} \int_a^b g\, d\sigma, \tag{6.7}$$

*где $\mathfrak{E}(\mathrm{c})$ — совокупность крайних мер множества $\mathcal{D}(\mathrm{c})$. Соотношение (6.7) остается справедливым, если супремум заменить на инфимум.*

Глава IV

# ДОПОЛНЕНИЯ И КЛАССИЧЕСКИЕ МОМЕНТНЫЕ ПРОСТРАНСТВА

В этой главе мы конкретизируем и уточняем теорию, изложенную в предшествующих двух главах, для специального случая важной $T$-системы $\{t^i\}_0^n$ на интервале $[0, 1]$. Ограничение рассмотрения интервалом $[a, b]$ не приводит к потере общности, так как любой конечный интервал $[a, b]$ может быть рассмотрен аналогично. Мы будем интерпретировать понятия опорной плоскости, канонической и главной мер, максимальной массы и другие геометрические величины, введенные в гл. II и III.

## § 1. Классический критерий для моментных точек

В лемме 9.2 гл. II мы доказали для произвольной $T$-системы $\{u_i\}_0^n$ на конечном интервале $[0, 1]$, что $\mathbf{c} \in \mathcal{M}_{n+1}$, если и только если неравенство $\sum_{i=0}^{n} a_i c_i \geqslant 0$ имеет место всякий раз, когда неравенство $\sum_{i=0}^{n} a_i u_i(t) \geqslant 0$ выполняется для всех $t \in [a, b]$, и что $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$, если и только если неравенство $\sum_{i=0}^{n} a_i c_i > 0$ имеет место всякий раз, когда $\sum_{i=0}^{n} a_i u_i(t) \geqslant 0$ и $\sum_{i=0}^{n} a_i^2 > 0$.

Далее, для $T$-системы $\{t^i\}_0^n$ на интервале $[0, 1]$ при $n = 2m$ было показано в следствии 10.3 гл. II, что любой неотрицательный многочлен $\sum_{i=0}^{n} a_i t^i$ может быть представлен как сумма вида $R_m^2(t) + t(1-t) R_{m-1}^2(t)$, где $R_m(t)$ обозначает многочлен степени $m$. Следовательно, на основании леммы 9.2 гл. II, становится очевидным, что $\mathbf{c}$ принадлежит $\mathcal{M}_{n+1}$, если и только если неравенство

$\sum_{i=0}^{n} a_i c_i \geqslant 0$ выполняется для всех наборов $\{a_0, \ldots, a_n\}$, где множество $\{a_i\}$ состоит из коэффициентов многочлена либо вида $R_m^2(t)$, либо вида $t(1-t)R_{m-1}^2(t)$. Кроме того, заключаем, что $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$, если и только если неравенство $\sum_{i=0}^{n} a_i c_i > 0$ выполняется для всех нетривиальных многочленов этих специальных видов. (Аналогичные утверждения устанавливаются в случае, когда $n = 2m+1$.) Теперь, записывая $R_m(t) = \sum_{i=0}^{m} \alpha_i t^i$ и $R_{m-1}(t) = \sum_{i=0}^{m-1} \beta_i t^i$ и обращаясь к лемме 9.2 гл. II, получим следующую теорему.

Теорема 1.1. *Для системы функций* $\{t^i\}_0^n$ *на интервале* $[0, 1]$ *имеем:*

(1) *Если* $n = 2m$, *то* $\mathbf{c} \in \mathcal{M}_{n+1}$, *если и только если две квадратичные формы*

$$\sum_{i,j=0}^{m} c_{i+j}\alpha_i\alpha_j \quad и \quad \sum_{i,j=0}^{m-1} (c_{i+j+1} - c_{i+j+2})\beta_i\beta_j \tag{1.1}$$

*положительно полуопределены. Кроме того,* $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$, *если и только если обе квадратичные формы являются положительно определенными.*

(2) *Если* $n = 2m+1$, *то* $\mathbf{c} \in \mathcal{M}_{n+1}$, *если и только если две квадратичные формы*

$$\sum_{i,j=0}^{m} c_{i+j+1}\alpha_i\alpha_j \quad и \quad \sum_{i,j=0}^{m} (c_{i+j} - c_{i+j+1})\beta_i\beta_j \tag{1.2}$$

*положительно полуопределены. Кроме того,* $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$, *если и только если эти квадратичные формы положительно определены.*

Теорема 1.1 вводит классический тест, которому последовательность $\{c_i\}_0^n$ должна удовлетворять, для того чтобы отвечать определению моментной точки. Для дискуссий по истории вопроса и первоисточникам отсылаем читателя к Шохату и Тамаркину [1943], а также к книге Ахиезера [1965]. В этих монографиях также обсуждаются обобщения на случай бесконечной моментной последовательности.

Теперь введем специальное обозначение для определителей Ганкеля, связанных с квадратичными формами, указанными в (1.1) и (1.2). Пусть

$$\underline{\Delta}_{2m} = \begin{vmatrix} c_0 & c_1 & \cdots & c_m \\ c_1 & c_2 & \cdots & c_{m+1} \\ \cdot & \cdot & \cdots & \cdot \\ c_m & c_{m+1} & \cdots & c_{2m} \end{vmatrix}, \quad \underline{\Delta}_{2m+1} = \begin{vmatrix} c_1 & c_2 & \cdots & c_{m+1} \\ c_2 & c_3 & \cdots & c_{m+2} \\ \cdot & \cdot & \cdots & \cdot \\ c_{m+1} & c_{m+2} & \cdots & c_{2m+1} \end{vmatrix}, \tag{1.3}$$

$$\overline{\Delta}_{2m} = \begin{vmatrix} c_1 - c_2 & c_2 - c_3 & \dots & c_m - c_{m+1} \\ c_2 - c_3 & c_3 - c_4 & \dots & c_{m+1} - c_{m+2} \\ \dots & \dots & \dots & \dots \\ c_m - c_{m+1} & c_{m+1} - c_{m+2} & \dots & c_{2m-1} - c_{2m} \end{vmatrix},$$

$$\overline{\Delta}_{2m+1} = \begin{vmatrix} c_0 - c_1 & c_1 - c_2 & \dots & c_m - c_{m+1} \\ c_1 - c_2 & c_2 - c_3 & \dots & c_{m+1} - c_{m+2} \\ \dots & \dots & \dots & \dots \\ c_m - c_{m+1} & c_{m+1} - c_{m+2} & \dots & c_{2m} - c_{2m+1} \end{vmatrix}.$$

Индексы в $\underline{\Delta}_{2m}$, $\underline{\Delta}_{2m+1}$ и т. д. указывают на порядок высшего момента, встречающегося в каждом определителе. В следующем параграфе будет видно, что знак черты снизу и сверху не противоречит нашему предыдущему употреблению этого символа.

Часто оказывается полезным следующее следствие из теоремы 1.1.

Следствие 1.1 *Точка* c *является внутренней по отношению к моментному пространству* $\mathcal{M}_{n+1}$, *если и только если определители* $\underline{\Delta}_0, \overline{\Delta}_0, \underline{\Delta}_1, \dots, \underline{\Delta}_n, \overline{\Delta}_n$ *положительны.*

Доказательство. Следствие вытекает из теоремы 1.1 ввиду очевидного геометрического факта, что если $(c_0, \dots, c_{n-1}, c_n) \in \in \operatorname{Int} \mathcal{M}_{n+1}$, то это означает, что $(c_0, \dots, c_r) \in \operatorname{Int} \mathcal{M}_{r+1}$, $0 \leq r \leq n$.

## § 2. Опорные гиперплоскости и ортогональные многочлены

Поставим в соответствие каждому из определителей $\underline{\Delta}_{2m}, \underline{\Delta}_{2m+1}$, $\overline{\Delta}_{2m}$ и $\overline{\Delta}_{2m+1}$ многочлен, полученный путем замещения элементов последнего столбца функциями $1, t, t^2, t^3, \dots$ А именно, введем многочлены вида

$$\underline{\Delta}_{2m}(t) = \begin{vmatrix} c_0 & c_1 & \dots & c_{m-1} & 1 \\ c_1 & c_2 & \dots & c_m & t \\ \dots & \dots & \dots & \dots & \dots \\ c_m & c_{m+1} & \dots & c_{2m-1} & t^m \end{vmatrix},$$

$$\underline{\Delta}_{2m+1}(t) = \begin{vmatrix} c_1 & c_2 & \dots & c_m & 1 \\ c_2 & c_3 & \dots & c_{m+1} & t \\ \dots & \dots & \dots & \dots & \dots \\ c_{m+1} & c_{m+2} & \dots & c_{2m} & t^m \end{vmatrix},$$

$$\overline{\Delta}_{2m+2}(t) = \begin{vmatrix} c_1 - c_2 & c_2 - c_3 & \dots & c_m - c_{m+1} & 1 \\ c_2 - c_3 & c_3 - c_4 & \dots & c_{m+1} - c_{m+2} & t \\ \dots & \dots & \dots & \dots & \dots \\ c_{m+1} - c_{m+2} & c_{m+2} - c_{m+3} & \dots & c_{2m} - c_{2m+1} & t^m \end{vmatrix},$$

$$\overline{\Delta}_{2m+1}(t)=\begin{vmatrix} c_0-c_1 & c_1-c_2 & \dots & c_{m-1}-c_m & 1 \\ c_1-c_2 & c_2-c_3 & \dots & c_m-c_{m+1} & t \\ \dots & \dots & \dots & \dots & \dots \\ c_m-c_{m+1} & c_{m+1}-c_{m+2} & \dots & c_{2m-1}-c_{2m} & t^m \end{vmatrix}.$$

Заметим, что индекс $m$ указывает степень многочленов.

Теорема 2.1. *Если $\sigma$ есть мера с моментами $c_i$, $i=0, 1, 2,\dots$, то многочлены $\{\underline{\Delta}_{2k}(t)\}$, $k=0, 1, 2,\dots$, ортогональны по отношению к $\sigma$, т. е.*

$$\int_0^1 \underline{\Delta}_{2k}(t)\,\underline{\Delta}_{2l}(t)\,d\sigma(t)=0 \text{ для } k\neq l \text{ и } k, l, \geqslant 0.$$

*Многочлены $\{\underline{\Delta}_{2k+1}(t)\}$, $\{\overline{\Delta}_{2k+2}(t)\}$ и $\{\overline{\Delta}_{2k+1}(t)\}$, $k=0, 1, 2, \dots$, также ортогональны по отношению к мерам $t\,d\sigma(t)$, $t(1-t)\,d\sigma(t)$ и $(1-t)\,d\sigma(t)$ соответственно.*

Доказательство. Необходимо только заметить, что

$$\int_0^1 t^i\underline{\Delta}_{2k}(t)\,d\sigma(t)=\begin{vmatrix} c_0 & \dots & c_{k-1} & c_i \\ c_1 & \dots & c_k & c_{i+1} \\ \dots & \dots & \dots & \dots \\ c_k & \dots & c_{2k-1} & c_{k+i} \end{vmatrix}=0, \qquad i<k.$$

Три других случая проверяются аналогично.

Четыре системы многочленов, в том виде как они приведены, не являются ортонормированными. Вместо этого мы имеем

$$\begin{aligned} &\int_0^1 [\underline{\Delta}_{2k}(t)]^2 d\sigma(t)=\underline{\Delta}_{2k}\underline{\Delta}_{2k-2}, \\ &\int_0^1 [\underline{\Delta}_{2k+1}(t)]^2 t\,d\sigma(t)=\underline{\Delta}_{2k+1}\underline{\Delta}_{2k-1}, \\ &\int_0^1 [\overline{\Delta}_{2k+2}(t)]^2 t(1-t)\,d\sigma(t)=\overline{\Delta}_{2k+2}\overline{\Delta}_{2k}, \\ &\int_0^1 [\overline{\Delta}_{2k+1}(t)]^2 (1-t)\,d\sigma(t)=\overline{\Delta}_{2k+1}\overline{\Delta}_{2k-1}, \end{aligned} \tag{2.1}$$

что устанавливается простым вычислением с использованием ортогональности.

Очевидно, если точка $(c_0, \dots, c_n)$ является внутренней по отношению к $\mathcal{M}_{n+1}$ для любого $n$, то все четыре системы бесконечны. Если мера $\sigma(t)$ содержит только конечное число точек сосредоточения масс, то эти четыре системы конечны. В частности, если $\sigma(t)$ есть мера, состоящая из $r$ сосредоточения масс, то многочлены $\underline{\Delta}_{2k}(t)$, $\underline{\Delta}_{2k+1}(t)$, $\overline{\Delta}_{2k+2}(t)$, $\overline{\Delta}_{2k+1}(t)$ имеют точно степень $k$ при $k=0, 1, \ldots, r$ и тождественно равны нулю для $k \geqslant r+1$. Многочлен $\underline{\Delta}_{2r}(t)$ обращается в нуль на спектре меры $\sigma$. Аналогичные утверждения применимы для многочленов $\underline{\Delta}_{2r+1}(t)$, $\overline{\Delta}_{2r+2}(t)$ и $\overline{\Delta}_{2r+1}(t)$.

Определение 2.1. Системы многочленов (возможно, конечные), которые *ортонормированы* по отношению к $d\sigma$, $t\,d\sigma$, $t(1-t)\,d\sigma$ и $(1-t)\,d\sigma$, причем нормированы так, что старший коэффициент положителен, будут обозначаться как $\{\underline{P}_k\}$, $\{\underline{Q}_k\}$, $\{\overline{P}_k\}$ и $\{\overline{Q}_k\}$ соответственно.

Всякий раз, когда знаменатели не равны нулю, имеем

$$\underline{P}_k(t)=\frac{\underline{\Delta}_{2k}(t)}{\sqrt{\underline{\Delta}_{2k}\underline{\Delta}_{2k-2}}},\quad \underline{Q}_k(t)=\frac{\underline{\Delta}_{2k+1}(t)}{\sqrt{\underline{\Delta}_{2k+1}\underline{\Delta}_{2k-1}}},$$

$$\overline{P}_k(t)=\frac{\overline{\Delta}_{2k+2}(t)}{\sqrt{\overline{\Delta}_{2k+2}\overline{\Delta}_{2k}}},\quad \overline{Q}_k(t)=\frac{\overline{\Delta}_{2k+1}(t)}{\sqrt{\overline{\Delta}_{2k+1}\overline{\Delta}_{2k-1}}}.$$

Мы будем теперь интерпретировать эти многочлены в терминах опорных плоскостей для моментных пространств. С этой целью докажем вспомогательную лемму, которая справедлива для произвольной $T$-системы и поэтому будет установлена в своей общей формулировке.

Пусть $\overline{c}_{n+1}$ и $\underline{c}_{n+1}$ будут верхним и нижним значениями интервала, образованного величиной $\int\limits_a^b u_{n+1}\,d\sigma$, когда $\sigma$ изменяется на множестве всех мер, представляющих $(c_0, c_1, \ldots, c_n)$.

Лемма 2.1. *Если $\{u_i\}_0^n$ и $\{u_i\}_0^{n+1}$ являются $T$-системами на $[a, b]$, то опорные многочлены пространства $\mathcal{M}_{n+2}$ при*

$$\overline{\mathbf{c}}=(c_0, \ldots, c_n, \overline{c}_{n+1}) \quad \text{и} \quad \underline{\mathbf{c}}=(c_0, \ldots, c_n, \underline{c}_{n+1})$$

*обращаются в нуль в корнях верхнего и нижнего главных представлений $(c_0, c_1, \ldots, c_n)$ соответственно.*

Доказательство. Повторив рассуждения, приведенные в § 6 гл. II, получим, что

$$\underline{c}_{n+1}=\sum_{j=1}^{p}\lambda_j^* u_{n+1}(t_j^*) \quad \text{и} \quad \overline{c}_{n+1}=\sum_{j=1}^{q}\alpha_j^* u_{n+1}(s_j^*),$$

где $\{t_j^*; \lambda_j^*\}_1^p$ и $\{s_j^*; \lambda_j^*\}_1^q$ образуют нижнее и верхнее главные представления $(c_0, \ldots, c_n)$ соответственно. Так как оба множества $\{t_j^*\}$ и $\{s_j^*\}$ имеют индекс $(n+1)/2$, то опорные многочлены должны обращаться в нуль, как указывалось. Например, если $(a_0, \ldots a_{n+1})$

определяет опорную плоскость в $\mathcal{M}_{n+1}$ в точке $\underline{c}$, то

$$0 = \sum_{i=0}^{n} a_i c_i + a_{n+1} \underline{c}_{n+1} = \sum_{j=1}^{p} \lambda_j^* \sum_{i=0}^{n+1} a_i u_i (t_j^*).$$

Но так как $\sum_{i=0}^{n+1} a_i u_i (t) \geqslant 0$ для всех $t \in [a, b]$ и $\lambda_j^* > 0$, то приходим к выводу, что $\sum_{i=0}^{n+1} a_i u_i (t_j^*) = 0$ для $j = 1, \ldots, p$, что и требовалось доказать.

Следствие 2.1. *Если $\{u_i\}_0^n$ является ET-системой порядка 2, то опорные многочлены при $\bar{c}$ и $\underline{c}$ определяются однозначно с точностью до постоянного множителя.*

Лемма 2.2. *Пусть точка* $\mathbf{c} = (c_0, c_1, \ldots, c_n)$ *является внутренней для моментного пространства $\mathcal{M}_{n+1}$, образованного системой $\{t^i\}_0^n$ на* $[0, 1]$. *Корни верхнего и нижнего главных представлений точки* $\mathbf{c}$ *совпадают соответственно с множеством нулей многочленов*:

(1) $(1 - t)\, \bar{Q}_m (t)$ *и* $t Q_m (t)$, *если* $n = 2m$;

(2) $t (1 - t)\, \bar{P}_m (t)$ *и* $\underline{P}_{m+1} (t)$, *если* $n = 2m + 1$.

Доказательство. Рассмотрим только верхнее главное представление, когда $n = 2m + 1$. В этом случае, если $R_{m-1}$ есть произвольный многочлен степени $m - 1$, то из свойств ортогональности многочлена $\bar{P}_m$ следует, что

$$\int_0^1 R_{m-1} (t)\, \bar{P}_m (t)\, t (1 - t)\, d\bar{\sigma} = 0, \tag{2.2}$$

где $\bar{\sigma}$ обозначает верхнюю главную меру, соответствующую точке с. Эта мера имеет $m$ внутренних корней $s_2^*, \ldots, s_{m+1}^*$, поэтому, если мы определяем многочлен $R_{m-1}$ так, чтобы он обращался в нуль в любых точках, кроме корня $s_i^*$, то условие (2.2) сводится к равенству

$$R_{m-1} (s_i^*)\, \bar{P}_m (s_i^*)\, s_i^* (1 - s_i^*)\, \alpha_i^* = 0$$

и, следовательно, $\bar{P}_m (s_i^*) = 0$. Итак, многочлен $t (1 - t)\, \bar{P}_m (t)$ обращается в нуль в корнях $0 = s_1^* < s_2^* < \ldots < s_{m+1}^* < s_{m+2}^* = 1$ меры $\bar{\sigma}$.

Следствие 2.2 а. *Для T-системы $\{t^i\}_0^n$ на интервале* $[0, 1]$ *опорными многочленами пространства $\mathcal{M}_{n+2}$ в граничных точках* $\bar{\mathbf{c}} = (c_0, \ldots, c_n, \bar{c}_{n+1})$ *и* $\underline{\mathbf{c}} = (c_0, \ldots c_n, \underline{c}_{n+1})$ (см. лемму 2.1) *служат* *многочлены $\overline{S}_{n+1}(t)$ и $\underline{S}_{n+1}(t)$ соответственно, где*

$$\underline{S}_{n+1}(t) = \begin{cases} t\,[\underline{Q}_m(t)]^2, & n = 2m, \\ [\underline{P}_{m+1}(t)]^2, & n = 2m+1, \end{cases}$$
$$\overline{S}_{n+1}(t) = \begin{cases} (1-t)\,[\overline{Q}_m(t)]^2, & n = 2m, \\ t(1-t)\,[\overline{P}_m(t)]^2, & n = 2m+1. \end{cases} \tag{2.3}$$

Доказательство. Этот результат следует из лемм 2.1, 2.2 и следствия 2.1. Например, если $n = 2m$, то опорный многочлен (ср. со сноской на стр. 54. гл. II) при $\overline{c}$ обращается в нуль в корнях меры $\overline{\sigma}$, которые по лемме 2.2 точно совпадают с нулями многочлена $(1-t)\,\overline{Q}_m(t)$. Тогда опорный многочлен должен быть пропорционален многочлену $(1-t)\,[\overline{Q}_m(t)]^2$.

Следствие 2.2 b. *Если $(c_0, \ldots, c_n) \in \operatorname{Int} \mathcal{M}_{n+1}$, то $\overline{c}_{n+1}$ и $\underline{c}_{n+1}$ являются соответственно решениями уравнений $\overline{\Delta}_{n+1} = 0$ и $\underline{\Delta}_{n+1} = 0$, в которых момент $c_{n+1}$ является переменным.*

Доказательство. Мы рассмотрим только случай $n = 2m$ для $\overline{c}_{n+1}$. Согласно следствию 2.2 а многочлен $\overline{S}_{n+1}(t) = (1-t)\,[\overline{Q}_m(t)]^2$ является опорным для пространства $\mathcal{M}_{n+2}$ при $\overline{\mathbf{c}} = (c_0, \ldots, c_n, \overline{c}_{n+1})$, так что $\int\limits_0^1 [\overline{Q}_m(t)]^2\,(1-t)\,d\overline{\sigma} = 0$. Следовательно, из (2.1) получаем $0 = \int\limits_0^1 [\overline{\Delta}_{2m+1}(t)]^2\,(1-t)\,d\overline{\sigma} = \overline{\Delta}_{2m+1}\overline{\Delta}_{2m-1}$. Согласно следствию 1.1 $\overline{\Delta}_{2m-1} > 0$, и, следовательно, $\overline{\Delta}_{2m+1} = 0$.

Опираясь на лемму 2.2, можно показать, что корни верхнего и нижнего канонических представлений точки $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$ являются нулями некоторых квазиортогональных многочленов. Для любой меры $\sigma$ на интервале $[0, 1]$ связанные с ней ортогональные многочлены $\{\underline{P}_k(t)\}$ с точностью до постоянного множителя определяются требованием ортогональности к любому многочлену степени, самое большее, $k-1$. *Квазиортогональный* многочлен степени $k$ есть многочлен, ортогональный к любому многочлену степени, самое большее, $k-2$. Согласно этому определению любой квазиортогональный многочлен $q_k(t)$ степени $k$ имеет вид

$$q_k(t) = \underline{P}_k(t) - \alpha \underline{P}_{k-1}(t), \tag{2.4}$$

так как в разложении $q_k(t) = \sum\limits_0^k a_i \underline{P}_i(t)$ коэффициенты $a_0, a_1, \ldots, a_{k-2}$ должны быть равны нулю, а $a_k \neq 0$, потому что $q_k(t)$ — многочлен степени $k$.

Лемма 2.3. *Предположим, что* $\mathbf{c} = (c_0, \ldots, c_n) \in \operatorname{Int} \mathcal{M}_{n+1}$. *Тогда*

(1) *если* $n = 2m$, *то существует взаимно однозначное соответствие между квазиортогональными многочленами*

$$t\underline{Q}_m(t) - \gamma(1-t)\overline{Q}_m(t), \quad 0 < \gamma < \infty, \tag{2.5}$$

*и нижними каноническими представлениями моментной точки* $\mathbf{c}$ *индекса* $(n+2)/2$ *такое, что нули многочлена* (2.5) *являются корнями соответствующего канонического представления. Каждый из внутренних корней есть строго возрастающая непрерывная функция параметра* $\gamma \in (0, \infty)$.

*Подобное соответствие существует и между нулями многочлена*

$$t(1-t)[\underline{Q}_m(t) - \gamma\overline{Q}_m(t)], \qquad -\infty < \gamma < 0, \tag{2.6}$$

*и верхними каноническими представлениями индекса* $(n+2)/2$;

2) *для* $n = 2m+1$ *утверждения, соответствующие утверждениям п.*(1), *справедливы в том смысле, что нули многочленов*

$$t[\underline{P}_{m+1}(t) - \gamma(1-t)\overline{P}_m(t)], \qquad 0 < \gamma < \infty, \tag{2.7}$$

$$(1-t)[\underline{P}_{m+1}(t) - \gamma t\overline{P}_m(t)], \qquad -\infty < \gamma < 0, \tag{2.8}$$

*отвечают нижнему и верхнему каноническим представлениям* $\mathbf{c}$ *соответственно.*

Доказательство. Рассмотрим нижние канонические представления для случая $n = 2m$. Как и в § 5 гл. II, положим, что

$$0 = t_1^* < t_2^* < \ldots < t_{m+1}^*, \quad s_1^* < s_2^* < \ldots < s_{m+1}^* = 1 \tag{2.9}$$

обозначают корни нижнего и верхнего главных представлений соответственно. Эти корни перемежаются следующим образом: $t_i^* < s_i^* < t_{i+1}^* < s_{i+1}^*$, $i = 1, \ldots, m$. Вспомним свойство, согласно которому корни $\xi_1 < \xi_2 < \ldots < \xi_{m+1}$ любого нижнего канонического представления индекса $(n+2)/2$ удовлетворяют условию $t_i^* < \xi_i < s_i^*$, $i = 1, 2, \ldots, m+1$.

Теперь по лемме 2.2 корни (2.9) являются нулями многочленов $t\underline{Q}_m(t)$ и $(1-t)\overline{Q}_m$ соответственно. Следовательно, так как каждый из многочленов $\underline{Q}_m$ и $\overline{Q}_m$ положителен для значений $t$, превышающих $t_{m+1}^*$ и $s_m^*$ соответственно, то значения функции

$$\gamma(\xi) = \frac{\xi\underline{Q}_m(\xi)}{(1-\xi)\overline{Q}_m(\xi)}$$

непрерывно заполняют интервал от 0 до $+\infty$, когда $\xi$ пробегает каждый из $m+1$ интервалов $(t_i^*, s_i^*)$, $i = 1, \ldots, m+1$. Кроме того, функция $\gamma(\xi)$ должна быть строго возрастающей, так как предположение $\gamma(\xi_1) = \gamma(\xi_2)$ для $\xi_1$ и $\xi_2$ в том же интервале означает, что многочлен $t\underline{Q}_m(t) - \gamma(\xi_1)(1-t)\overline{Q}_m(t)$ имеет по крайней мере $m+2$ нулей.

Теперь параметризуем нижние канонические представления $\sigma_\xi$ моментной точки c, имеющей индекс $(n+2)/2$, по их наибольшему корню $\xi \in (t^*_{m+1}, 1)$. Многочлен $S_{m+1}(t) = t\underline{Q}_m(t) - \gamma(\xi)(1-t)\overline{Q}_m$, очевидно, обращается в нуль при $\xi$ и имеет нуль в каждом из других $m$ интервалов $(t^*_i, s^*_i)$, $i = 1, \ldots, m$.

Мы хотим показать, что эти нули являются в действительности другими корнями канонической меры $\sigma_\xi$. В самом деле, благодаря свойствам ортогональности многочленов $\underline{Q}_m$ и $\overline{Q}_m$ имеем

$$\int_0^1 R_{m-1}(t)\,[t\underline{Q}_m(t) - \gamma(\xi)(1-t)\overline{Q}_m(t)]\,d\sigma_\xi = 0$$

для любого многочлена $R_{m-1}(t)$ степени, самое большее, $m-1$. Если мы определим многочлен $R_{m-1}(t)$ таким образом, чтобы он обращался в нуль в корнях $\xi_1, \ldots, \xi_m$, за исключением $\xi_i$, то приходим к выводу, что многочлен $S_{m+1}(t)$ обращается в нуль при $\xi_i$, как и требовалось.

Функция $\gamma(\xi)$, когда $\xi$ пробегает интервал $(t^*_{m+1}, 1)$, описывает взаимно однозначное соответствие между многочленами (2.5) и нижними каноническими представлениями. Кроме того, каждый из корней канонических представлений является строго возрастающей функцией $\gamma \in (0, \infty)$.

Другие случаи, включающие многочлены (2.6), (2.7) и (2.8), могут быть трактованы таким же образом.

Геометрические интерпретации, содержащиеся в этом параграфе, основаны на книге Карлина и Шепли [1953]. Классическое аналитическое изложение этих идей в завершенном виде представлено у Шохата и Тамаркина [1943]. Обсуждение вопросов, связанных с каноническими мерами и квазиортогональными многочленами, в неявном виде встречается также у Шохата и Тамаркина.

## § 3. Максимальная масса $\rho(t_0)$

Для $\mathbf{c} = (c_0, \ldots, c_n) \in \operatorname{Int} \mathcal{M}_{n+1}$ была введена (§ 4 гл. II) величина

$$\rho(t_0) = \min \int_0^1 u(t)\,d\sigma(t), \tag{3.1}$$

где $\sigma \in V(\mathbf{c})$, т. е. $\sigma$ удовлетворяет условию $\int u_i d\sigma = c_i$ для $i = 0, 1, \ldots, n$, а минимум отыскивается на множестве всех неотрицательных многочленов $u(t) = \sum_{i=0}^{n} a_i u_i(t)$, удовлетворяющих условию

$u(t_0) = 1$. В § 4 гл. II было установлено, что $\rho(t_0)$ представляет собой массу, приписываемую точке $t_0$ в каноническом представлении c, которое включает $t_0$. Кроме того, минимум в (3.1) достигается при многочлене $u(t)$, который обращается в нуль в корнях (исключая $t_0$) этого канонического представления. В случае системы $\{t^i\}_0^n$, которая является $ET$-системой, мы доказали, что минимизирующий многочлен является единственным при условии, что $t_0$ не является внутренним корнем одного из двух главных представлений точки c. Если $t_0 \in (0, 1)$ является корнем главного представления $(c_0, \ldots, c_n)$, то мин мизирующий многочлен единствен в классе всех неотрицательных многочленов степени, самое большее, $n-1$.

Для удобства вспомним некоторые определения (см. §§ 5, 6 гл. II). Для $n = 2m$ корни $\{s_i^*\}_1^{m+1}$ и $\{t_i^*\}_1^{m+1}$ верхнего и нижнего главных представлений моментной точки $c \in \operatorname{Int} \mathcal{M}_{n+1}$ удовлетворяют соотношениям

$$0 = t_1^* < s_1^* < t_2^* < s_2^* < \ldots < t_{m+1}^* < s_{m+1}^* = 1,$$

а верхние и нижние канонические интервалы (определение 6.1) были определены как $K_i = (s_i^*, t_{i+1}^*)$, $i = 1, \ldots, m$, и $J_i = (t_i^*, s_i^*)$, $i = 1, 2, \ldots, m+1$, соответственно. Для $n = 2m+1$ соответствующие корни главных представлений удовлетворяют условию $0 = s_1^* < t_1^* < s_2^* < \ldots < t_{m+1}^* < s_{m+2}^* = 1$, а верхние и нижние канонические интервалы имеют вид $K_i = (s_i^*, t_i^*)$ и $J_i = (t_i^*, s_{i+1}^*)$, $i = 1, 2, \ldots, m+1$.

Пусть $J \equiv \bigcup J_i$ и $K \equiv \bigcup K_i$. Наша цель в данный момент состоит в том, чтобы явно выразить $\rho(t_0)$ через ортонормированные многочлены, введенные в § 2. Это и составляет содержание следующей теоремы.

Теорема 3.1 (Шенберг и Сеге [1960]).

(1) *Если $n = 2m$, то*

$$\rho^{-1}(t_0) = \begin{cases} t_0(1-t_0)\sum_{k=0}^{m-1} \overline{P}_k^2(t_0), & t_0 \in \overline{K}, \\ \sum_{k=0}^{m} \underline{P}_k^2(t_0), & t_0 \in \overline{J}. \end{cases}$$

(2) *Если $n = 2m+1$, то*

$$\rho^{-1}(t_0) = \begin{cases} (1-t_0)\sum_{k=0}^{m} \overline{Q}^2(t_0), & t_0 \in \overline{K}, \\ t_0\sum_{k=0}^{m} \underline{Q}_k^2(t_0), & t_0 \in \overline{J}, \end{cases}$$

*($\overline{J}$ и $\overline{K}$ обозначают замыкания интервалов $J$ и $K$ соответственно).*

Замечание 3.1. Результат, в точности соответствующий теореме 3.1, был установлен Шенбергом и Сеге [1960]. Более ранние несовершенные варианты формулировок этого результата рассмотрены у Сеге [1959] и Крейна [1951]. Относящиеся к этой задаче результаты имеются у Брикмана [1960]. Доказательства, данные Шенбергом и Сеге, носили аналитический характер, в то время как доказательство, представленное ниже, является по существу геометрическим.

Доказательство. Рассмотрим случай $n = 2m+1$ с $t_0 \in J$. Как упоминалось выше, минимум $\rho(t_0)$ достигается для единственного неотрицательного нормированного многочлена, который обращается в нуль в корне (исключая $t_0$) канонического представления $\sigma_{t_0}$. Следовательно, минимизирующий многочлен имеет вид $P(t) = tR_m^2(t)$. Если мы выражаем $R_m(t)$ через систему многочленов $\underline{Q}_k(t)$, т. е. $R_m(t) = \sum\limits_{k=0}^{m} \gamma_k \underline{Q}_k(t)$, то, используя свойство ортонормированности многочленов $\underline{Q}_k(t)$ по отношению к мере $t\,d\sigma$, найдем, что

$$\int_0^1 P(t)\, d\sigma(t) = \sum_{k=0}^{m} \gamma_k^2. \tag{3.2}$$

Применяя неравенство Шварца, получаем

$$1 = P(t_0) = t_0 \left| \sum_{k=0}^{m} \gamma_k \underline{Q}_k(t_0) \right|^2 \leq t_0 \sum_{k=0}^{m} \gamma_k^2 \sum_{k=0}^{m} \underline{Q}_k^2(t_0), \tag{3.3}$$

где знак равенства имеет место тогда и только тогда, когда

$$\gamma_k = \frac{\underline{Q}_k(t_0)}{\sqrt{t_0} \sum\limits_{0}^{m} \underline{Q}_k^2(t_0)}.$$

С учетом (3.2) из неравенства (3.3) следует, что $\rho^{-1}(t_0) = t_0 \sum\limits_{k=0}^{m} \underline{Q}_k^2(t_0)$. В других случаях используются аналогичные методы.

Когда $t_0$ есть корень одного из главных представлений и не является граничной точкой, то значение $\rho(t_0)$ может быть выражено в несколько другой форме. В частности, пусть $s_i^*$ и $t_i^*$ обозначают корни верхнего и нижнего главных представлений соответственно, которые не равны нулю или единице. Тогда, если $n = 2m$, то

$$\rho^{-1}(t_0) = \begin{cases} (1-t_0) \sum\limits_{k=0}^{m-1} \overline{Q}_k^2(t_0), & \text{если} \quad t_0 = s_i^*, \qquad i = 1, \ldots, m, \\ t_0 \sum\limits_{k=0}^{m-1} \underline{Q}_k^2(t_0), & \text{если} \quad t_0 = t_i^*, \qquad i = 2, \ldots, m+1, \end{cases}$$

а если $n = 2m+1$, то

$$\rho^{-1}(t_0) = \begin{cases} t_0(1-t_0)\sum_{k=0}^{m-1} \overline{P}_k^2(t_0), & \text{если} \quad t_0 = s_i^*, \quad i = 2, \ldots, m+1, \\ \sum_{k=0}^{m} \underline{P}_k^2(t_0), & \text{если} \quad t_0 = t_i^*, \quad i = 1, \ldots, m+1. \end{cases}$$

Чтобы убедиться в справедливости этих формул, заметим, например, для $n = 2m$ и $t_0 = s_i^*$ ($i = 1, \ldots, m$), что единственный минимизирующий многочлен степени $n-1$ с необходимостью обращается в нуль в $m-1$ других внутренних точках верхнего главного представления и в граничной точке $t=1$. Таким образом, минимизирующий многочлен имеет вид $(1-t)R_{m-1}^2(t)$. Далее следует анализ, который приведен в теореме 3.1. Остальные случаи совершенно аналогичны.

Мы теперь получим ряд различных представлений для величины $\rho(t_0)$.

(1) Эквивалентная характеризация величины $\rho(t_0)$ в терминах задачи максимизации состоит в следующем. Из (3.1) легко видеть, что $\rho^{-1}(t_0) = \max P(t_0)$, где максимум берется по всем неотрицательным многочленам на $[0, 1]$, нормированным так, что $\int_0^1 P(t)\,d\sigma(t) = 1$, а мера $\sigma$ образует моменты $c_i = \int_0^1 t^i d\sigma(t)$, $i = 0, 1, \ldots, n$.

(2) Величину $\rho(t_0)$ можно также вычислить, используя произвольную систему $\{f_i(t)\}_0^m$ «линейно независимых» многочленов, для которых линейной оболочкой служит то же самое пространство, что и для степенных функций $\{t^i\}_0^m$. Например, если минимизирующий многочлен выражается в виде $R_m^2(t) = \left(\sum_0^m \gamma_i f_i(t)\right)^2$, то величину $\int_0^1 R_m^2(t)\,d\sigma(t)$ можно минимизировать при условии $R_m^2(t_0) = 1$, используя технику множителей Лагранжа, чтобы получить равенство $\rho^{-1}(t_0) = f'M^{-1}f$, где $f' = (f_0(t_0), \ldots, f_m(t_0))$, $M = (m_{ij})$ и $m_{ij} = \int_0^1 f_i(t)\,f_j(t)\,d\sigma(t)$.

Минимальное значение достигается, когда

$$(\gamma_0, \ldots, \gamma_m) = \frac{\pm f'M^{-1}}{f'M^{-1}f}.$$

(3) Другое представление для величины $\rho(t_0)$ выражается в терминах некоторых определителей, включающих «воспроизводящее ядро».

Рассмотрим случай $n = 2m$. Введем следующую функцию:

$$K(x, y) = \frac{(-1)}{\Delta_{2m}} \begin{vmatrix} 0 & 1 & y & \ldots & y^m \\ 1 & c_0 & c_1 & \ldots & c_m \\ x & c_1 & c_2 & \ldots & c_{m+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x^m & c_m & c_{m+1} & \ldots & c_{2m} \end{vmatrix}.$$

Тогда $K(x, y)$ обладает тем свойством, что для любого многочлена $\pi_k(y)$ степени $k \leqslant m$

$$\int_0^1 K(x, y)\, \pi_k(y)\, d\sigma(y) = \pi_k(x). \tag{3.4}$$

Достаточно проверить (3.4) для случая степенных функций $y^k$, $k = 0, 1, \ldots, m$. Тогда мы получим

$$\int_0^1 K(x, y)\, y^k d\sigma(y) = \frac{(-1)}{\Delta_{2m}} \begin{vmatrix} 0 & c_k & \ldots & c_{k+m} \\ 1 & c_0 & \ldots & c_m \\ x & c_1 & \ldots & c_{m+1} \\ \cdot & \cdot & \cdot & \cdot \\ x^m & c_m & \ldots & c_{2m} \end{vmatrix}.$$

Определитель легко сводится к $x^k$, если разложить его по элементам первого столбца.

Если мы представим определитель из выражения для $K(x, y)$ как функцию от $y$ в терминах многочленов $\underline{P}_k(y)$, которые ортонормированы по отношению к мере $\sigma$, то получим $K(x, y) = \sum_{k=0}^{m} g_k(x)\, \underline{P}_k(y)$. Положив $\pi_i(y) = \underline{P}_i(y)$ в (3.4), заключаем, что

$$\underline{P}_i(x) = \int_0^1 K(x, y)\, \underline{P}_i(y)\, d\sigma(y) = g_i(x), \qquad i = 1, \ldots, m,$$

так что $K(x, y) = \sum_{k=0}^{m} \underline{P}_k(x)\, \underline{P}_k(y)$. Из теоремы 3.1 следует, что если $n = 2m$ и $t_0$ принадлежит $\overline{J}$, то

$$\rho^{-1}(t_0) = \frac{(-1)}{\Delta_{2m}} \begin{vmatrix} 0 & 1 & t_0 & \ldots & t_0^m \\ 1 & c_0 & c_1 & \ldots & c_m \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ t_0^m & c_m & c_{m+1} & \ldots & c_{2m} \end{vmatrix}.$$

Аналогичные выражения, включающие три другие системы многочленов, также могут быть получены.

## § 4. Геометрия классических чебышевских многочленов

С точностью до некоторых усилений §§ 4—8 следуют в основном работе Карлина и Шепли [1953]. В этом параграфе звездочка (*) будет указывать величины, имеющие распределение вида

$$d\sigma^*(t) = \frac{dt}{\pi \sqrt{t(1-t)}}.$$

Момент порядка $k$ меры $\sigma^*(t)$ легко вычисляется и равен

$$c_k^* = \frac{2k-1}{2k} c_{k-1}^* = \binom{2k}{k} 2^{-2k}, \tag{4.1}$$

а ортогональные системы многочленов $\{\underline{\Delta}_{2k}^*(t)\}$ и $\{\overline{\Delta}_{2k}^*(t)\}$ (см. теорему 2.1) оказываются чебышевскими многочленами первого и второго родов соответственно, т. е.

$$\mathrm{const}\cdot\underline{\Delta}_{2k}^*(t) = T_k(t) = \cos k\theta,$$

$$\mathrm{const}\cdot\overline{\Delta}_{2k}^*(t) = U_{k-1}(t) = \frac{\sin k\theta}{\sin\theta},$$

где $\theta = \cos^{-1}(2t-1)$. Системы многочленов $\{\underline{\Delta}_{2k+1}^*(t)\}$ и $\{\overline{\Delta}_{2k+1}^*(t)\}$ суть многочлены Якоби, которые обозначаются как $P_k^{(-1/2,1/2)}$ и $P_k^{(1/2,-1/2)}$, т. е.

$$\mathrm{const}\cdot\underline{\Delta}_{2k+1}^*(t) = \frac{\cos\frac{1}{2}(2k+1)\theta}{\cos\frac{1}{2}\theta},$$

$$\mathrm{const}\cdot\overline{\Delta}_{2k+1}^*(t) = \frac{\sin\frac{1}{2}(2k+1)\theta}{\sin\frac{1}{2}\theta}.$$

Эти многочлены, рассматриваемые как опорные гиперплоскости к моментным пространствам, обладают рядом интересных геометрических свойств. Мы воспользуемся записью $(x_0, \ldots, x_n)$ для обозначения произвольной точки в евклидовом $(n+1)$-мерном пространстве $E_{n+1}$.

Теорема 4.1. *Для любого $n$ два линейных многообразия, определенные пересечением гиперплоскости $x_0 = 1$ с опорными гиперплоскостями пространства $\mathcal{M}_{n+1}$ в точках $(c_0^*, \ldots, c_{n-1}^*, \underline{c}_n^*)$ и $(c_0^*, \ldots, c_{n-1}^*, \overline{c}_n^*)$, параллельны; это свойство однозначно характеризует $\sigma^*$.*

Доказательство. Свойство параллельности имеет место тогда и только тогда, когда первые производные от соответствующих опорных многочленов пропорциональны, т. е.

$$\begin{aligned} [\underline{\Delta}_n(t)^2]' &= a_n[t(1-t)\overline{\Delta}_n(t)^2]', \quad \text{если } n \text{ четное,} \\ [t\underline{\Delta}_n(t)^2]' &= b_n[(1-t)\overline{\Delta}_n(t)^2]', \quad \text{если } n \text{ нечетное,} \end{aligned} \tag{4.2}$$

где $a_n$ и $b_n$ — коэффициенты пропорциональности.

В случае $(c_0, \ldots, c_{n-1}) = (c_0^*, \ldots, c_{n-1}^*)$ непосредственная подстановка показывает, что все четыре производные пропорциональны величине $\sin n\theta/\sin\theta$, что и доказывает первую часть теоремы.

Чтобы установить единственность, рассмотрим только случай $n = 2m$ и применим запись $\underline{R}_m(t)$ для $\underline{\Delta}_{2m}(t)$ и $\overline{R}_{m-1}(t)$ для $\overline{\Delta}_{2m}(t)$, чтобы указать более определенно на степени этих многочленов. Многочлены в равенствах (4.2) имеют степень $2m-1$ и с необходимостью обращаются в нуль в каждом из корней многочленов $\underline{R}_m$ и $\overline{R}_{m-1}$, которые все вещественны и различны, так как они соответствуют внутренним корням в нижнем и верхнем главных представлениях $(c_0^*, \ldots, c_{n-1}^*)$ (см. лемму 2.2). Следовательно, имеем

$$2\underline{R}_m\overline{R}_m' = d_m\underline{R}_m\overline{R}_{m-1}, \tag{4.3}$$

$$2t(1-t)\,\overline{R}_{m-1}\overline{R}_{m-1}' + (1-2t)\,\overline{R}_{m-1}^2 = e_m\underline{R}_m\overline{R}_{m-1},$$

где $d_m$ и $e_m$ — константы. Исключая поочередно $\overline{R}_{m-1}$ и $\underline{R}_m$, получаем

$$\begin{gathered} 2t(1-t)\,\underline{R}_m'' + (1-2t)\,\underline{R}_m' = \gamma_m\underline{R}_m, \\ 2t(1-t)\,\overline{R}_{m-1}'' + 3(1-2t)\,\overline{R}_{m-1}' = \delta_{m-1}\overline{R}_{m-1}, \end{gathered} \tag{4.4}$$

где $\gamma_m$ и $\delta_{m-1}$ зависят от $d_m$ и $e_m$. Таким образом, мы видим, что многочлены $m$-й степени $\underline{R}_m$ и $\overline{R}_m$ являются собственными функциями двух дифференциальных операторов. Единственность для каждого $m$ теперь следует в силу того, что первое дифференциальное уравнение в (4.4) единственным образом (с точностью до постоянного множителя) определяет чебышевские многочлены первого рода, а второе уравнение однозначно характеризует чебышевские многочлены второго рода (дальнейшие подробности, относящиеся к этому вопросу, см. у Сеге [1959], стр. 60).

Теорема 4.1 по существу формулирует на геометрическом языке два известных экстремальных свойства чебышевских многочленов. Мы сформулируем эти результаты в виде следующей теоремы.

Теорема 4.2.(1) *Пусть $\mathcal{P}$ обозначает класс многочленов степени $m$ со старшим коэффициентом $2^{m-1}$. Тогда*

$$\min_{\mathcal{P}} \max_{t\in[0,1]} |P_m(t)|^2 = \max_{t\in[0,1]} |T_m(t)|^2,$$

*где $T_m(t)$ есть чебышевский многочлен первого рода.*

(2) *Если $\mathcal{R}$ обозначает класс многочленов степени $m-1$ со старшим коэффициентом $2^{m-1}$, то*

$$\min_{\mathcal{R}} \max_{t\in[0,1]} t(1-t)\,|P_{m-1}(t)|^2 = \max_{t\in[0,1]} t(1-t)\,|U_{m-1}(t)|^2,$$

*где $U_{m-1}(t)$ является чебышевским многочленом второго рода.*

Доказательство. Мы ограничимся рассмотрением части (1), так как доказательство части (2) проводится аналогично. Заметим, что расстояние в направлении $x_n$ от точки $\mathbf{x}_0 \in R^n$ до гиперплоскости, заданной уравнением $(\mathbf{y}, \mathbf{x}_0) = 0$, выражается как $(\mathbf{y}, \mathbf{x}_0)/y_n$, если $y_n \neq 0$. Пусть $R_m(t) = \sum_{i=0}^{m} \alpha_i t^i$, $\alpha_m \neq 0$, — произвольный многочлен степени $m$. Расстояние в направлении $x_{2m}$ от точки

$\mathbf{x}(t) = (1, t, \ldots, t^{2m})$ до гиперплоскости, соответствующей $R_m^2(t)$, равно $R_m^2(t)/\alpha_m^2$. Так как $\mathbf{x}(t)$ описывает множество экстремальных точек сечения пространства $\mathcal{M}_{n+1}$ с $c_0 = 1$, в то время как гиперплоскость является граничной или опорной к нему, то максимум величины $R_m^2(t)/\alpha_m^2$, когда $t$ меняется от $t$ до 1, не может быть меньше, чем ширина в направлении $x_{2m}$. Из свойства параллельности плоскостей, сформулированного в теореме 4.1, заключаем, что $T_m$ доставляет нижнюю границу для максимального расстояния. Это эквивалентно утверждению части (1).

Любой момент меры $\sigma^*$ есть среднее значений, принимаемых ранее указанными моментами; таким образом, справедлива следующая теорема.

Теорема 4.3. *Для любого $n$*

$$c_n^* = \frac{1}{2}(\bar{c}_n^* + \underline{c}_n^*). \tag{4.5}$$

*Это свойство однозначно определяет меру $\sigma^*$, если $c_0^* = 1$.*

Доказательство. Пусть $\bar{\sigma}$ и $\underline{\sigma}$ — верхняя и нижняя главные меры $(c_0^*, \ldots, c_{n-1}^*)$. Если $n = 2m$, то по следствию 2.2 b и (2.1) имеем

$$\int_0^1 t^m \underline{\Delta}_{2m}^*(t)\, d\underline{\sigma} = 0, \tag{4.6}$$

$$\int_0^1 (t^m - t^{m+1})\, \bar{\Delta}_{2m}^*(t)\, d\bar{\sigma} = 0, \tag{4.7}$$

причем первое из этих соотношений включает $\underline{c}_n^*$, а второе — $\bar{c}_n^*$. Используя равенство (4.3), можно из (4.7) получить, что

$$\int_0^1 (t^m - t^{m+1})\, [\underline{\Delta}_{2m}^*(t)]'\, d\bar{\sigma} = 0. \tag{4.8}$$

Левые части равенств (4.8) и (4.6), записанные как определители, отличаются только элементами последних столбцов. Вычитая равенство (4.8) из равенства (4.6), предварительно умноженного на $m$, получим

$$\begin{vmatrix} c_0^* & \cdots & c_{m-1}^* & mc_m^* \\ c_1^* & \cdots & c_m^* & (m+1)\,c_{m+1}^* - c_m^* \\ \cdot & \cdots & \cdot & \cdots \\ c_k^* & \cdots & c_{m+k-1}^* & (m+k)c_{m+k}^* - kc_{m+k-1}^* \\ \cdot & \cdots & \cdot & \cdots \\ c_m^* & \cdots & c_{n-1}^* & m(\underline{c}_n^* + \bar{c}_n^*) - mc_{n-1}^* \end{vmatrix} = 0.$$

В этом определителе последние два столбца пропорциональны, за исключением последней строки, но в то же время из соотношения (4.1) имеем, что

$$(2m-1)\,c^*_{m+k-1} = 2\,(m+k)\,c^*_{m+k} - 2kc^*_{m+k-1}.$$

Следовательно, равенство с определителем сводится к

$$\underline{\Delta}^*_{n-2}\left[m\,(\underline{c}^*_n + \bar{c}^*_n) - mc^*_{n-1} - \frac{2m-1}{2}\,c^*_{n-1}\right] = 0,$$

и так как $\underline{\Delta}^*_{n-2} > 0$, то получаем

$$\frac{1}{2}\,(\underline{c}^*_n + \bar{c}^*_n) = \frac{2n-1}{2n}\,c^*_{n-1} = c^*_n,$$

что и требовалось доказать. Случай нечетного $n$ может быть рассмотрен аналогично.

Теорема 4.4. *Для любого* $n$ $\underline{\Delta}^*_n = \overline{\Delta}^*_n$; *это свойство определяет* $\sigma^*$ *однозначно, если* $c_0 = 1$.

Доказательство. Так как $(c^*_0, \ldots, c^*_{n-1}) \in \operatorname{Int}\mathcal{M}_n$, то из следствия 2.2 b известно, что решениями уравнений $\underline{\Delta}^*_n = 0$ и $\overline{\Delta}^*_n = 0$ относительно $c_n$ служат $\underline{c}^*_n$ и $\bar{c}^*_n$ соответственно. Данные уравнения линейны относительно $c^*_n$, и поэтому можно записать

$$\underline{c}^*_n = c^*_n - \frac{\underline{\Delta}^*_n}{\underline{\Delta}^*_{n-2}}, \tag{4.9}$$

$$\bar{c}^*_n = c^*_n + \frac{\overline{\Delta}^*_n}{\overline{\Delta}^*_{n-2}}. \tag{4.10}$$

Заметим, что эти уравнения действительно зависят от $c^*_n$, как и должно было бы быть. Далее из (4.9), (4.10) и предыдущей теоремы получим

$$\frac{\overline{\Delta}^*_n}{\overline{\Delta}^*_{n-2}} - \frac{\underline{\Delta}^*_n}{\underline{\Delta}^*_{n-2}} = \bar{c}^*_n + \underline{c}^*_n - 2c^*_n = 0. \tag{4.11}$$

Непосредственная проверка показывает, что $\underline{\Delta}^*_1 = \overline{\Delta}^*_1 = 1/2$ и $\underline{\Delta}^*_2 = \overline{\Delta}^*_2 = 1/8$. По индукции можно показать, что $\underline{\Delta}^*_n = \overline{\Delta}^*_n$ для любого $n$. Напротив, если любая последовательность $1, c_1, c_2, \ldots$ обладает тем свойством, что $\underline{\Delta}_n = \overline{\Delta}_n$ для любого $n$, то равенство (4.11), записанное без звездочек, в сочетании с предыдущей теоремой показывает, что $\{c_k\}$ должна быть моментной последовательностью по мере $\sigma^*$.

Теперь приведем численные результаты, которые зависят конечно от конкретного интервала ортогональности $[0, 1]$. Предшествующие теоремы этого параграфа в равной мере справедливы (после очевидных корректив) для любого конечного интервала.

Теорема 4.5. *Ширина в направлении $c_n$ сечения пространства $\mathcal{M}_{n+1}$, определенная при $c_0 = 1$, равна* $2^{-2n+2}$.

Доказательство. Старший коэффициент ортонормированного чебышевского многочлена степени $m$, вычисленный непосредственно, оказывается равным $2^{(4m-1)/2}$. Однако из определения многочлена $\underline{\Delta}_{2k}^*(t)$ он также равен $(\underline{\Delta}_{2m-2}^*/\underline{\Delta}_{2m}^*)^{1/2}$. Согласно равенствам (4.9) и (4.10), а также в соответствии с последней теоремой имеем $(n=2m)$

$$\bar{c}_{2m}^* - \underline{c}_{2m}^* = \frac{2\underline{\Delta}_{2m}^*}{\underline{\Delta}_{2m-2}^*} = 2^{-4m+2}.$$

Подобная же процедура устанавливает равенство (4.12) для нечетного $n$. Требуемый результат тогда следует из теоремы 4.1.

Теорема 4.6.

$$\underline{\Delta}_n^* = \overline{\Delta}_n^* = 2^{-n(n+1)/2}, \quad \underline{c}_n^* = \left[\binom{2n}{n} - 2\right] 2^{-2n}, \quad \bar{c}_n^* = \left[\binom{2n}{n} + 2\right] 2^{-2n}.$$

Доказательство. Первое равенство получается индукцией из (4.12). Два других равенства следуют из (4.12), (4.5) и (4.1).

## § 5. Симплексы $S^n$ и $B^n$

**А. Симплекс $S^n$.** Два симплекса $M^n$ и $P^n$, которые являются сечениями пространств $\mathcal{M}_{n+1}$ и $\mathcal{P}_{n+1}$, полученными путем наложения условий нормировки $\int_0^1 d\sigma(t) = 1$ и $\int_0^1 P(t)dt = 1$ соответственно, тесно связаны с другой парой многогранных симплексов.

Первый многогранный симплекс, описанный около симплекса $M^n$, есть выпуклая оболочка его вершин $\mathbf{x}^k$, $k = 0, 1, \ldots, n$, точные координаты которых вычисляются по следующей формуле:

$$x_i^{(k)} = \frac{k!\,(n-i)!}{n!\,(k-i)!} = \frac{\binom{k}{i}}{\binom{n}{i}}, \qquad i = 0, 1, \ldots, n, \tag{5.1}$$

где, как обычно, $\binom{k}{i} = 0$, если $i > k$. Мы будем обозначать этот симплекс через $S^n$.

Теорема 5.1. *$S^n$ содержит $M^n$.*

Доказательство. В основе доказательства лежит тождество

$$t^i = \sum_{k=0}^{n} \binom{n}{k} t^k (1-t)^{n-k} \frac{\binom{k}{i}}{\binom{n}{i}}, \qquad i = 0, 1, \ldots, n. \tag{5.2}$$

Приняв (5,2), непосредственной подстановкой убеждаемся, что

$$t^i = \sum_{k=0}^{n} \xi_k x_i^{(k)}, \qquad i = 0, 1, \ldots, n, \tag{5.3}$$

где $\xi_k = \binom{n}{k} t^k (1-t)^{n-k} \geqslant 0$ и $\sum_{k=0}^{n} \xi_k = 1$. Это равенство показывает, что выпуклое множество, натянутое на крайние точки $\mathbf{x}^{(k)}$ симплекса $S^n$, содержит крайние точки $(1, t, \ldots, t^n)$ симплекса $M^n$, и, следовательно, $M^n \subset S^n$.

Чтобы проверить справедливость тождества (5.2), используем тривиальное соотношение $\binom{n}{k}\binom{k}{i} = \binom{n-i}{k-i}\binom{n}{i}$. Выражение из право части (5.2) может быть представлено в виде $\sum_{k=0}^{n} \binom{n-i}{k-i} t^k (1-t)^{n-k}$ и затем легко сводится к $t^i$.

Данное распределение $\sigma(t)$ с условием нормировки $\int_0^1 d\sigma = 1$ и определенные величины $\lambda_{nk}(\sigma)$ были введены Хаусдорфом, который использовал последовательные правые разности $\Delta^l c_k(\sigma) = (-1)^l \times \times \sum_{i=0}^{l} (-1)^i \binom{l}{i} c_{k+i}(\sigma)$ от последовательности $\{c_i(\sigma)\}$ моментов меры $\sigma$. В частности,

$$\lambda_{nk}(\sigma) = (-1)^{n-k} \binom{n}{k} \Delta^{n-k} c_k(\sigma) =$$

$$= \binom{n}{k} \sum_{i=0}^{n-k} (-1)^i \binom{n-k}{i} c_{k+i}(\sigma). \tag{5.4}$$

Эти числа имеют простую геометрическую интерпретацию. Действительно, сумма в выражении (5.4) может быть записана как

$$\lambda_{nk}(\sigma) = \int_0^1 \binom{n}{k} t^k (1-t)^{n-k} d\sigma. \tag{5.5}$$

Из равенств (5.1)—(5.3) следует, что

$$c_i = \sum_{k=0}^{n} \lambda_{nk}(\sigma)\, x_i^{(k)}, \quad i = 0, 1, \ldots, n, \tag{5.6}$$

где, согласно (5.5), $\lambda_{nk} \geqslant 0$ и $\sum_{k=0}^{n} \lambda_{nk} = 1$. Таким образом, $\lambda_{nk}(\sigma)$

оказываются *барицентрическими координатами* точки $(c_0, \ldots, c_n)$ в симплексе $S^n$.

**В. Симплекс $B^n$.** Вернувшись к дуальному пространству, рассмотрим многочлены $B_{nk}(t) = \binom{n}{k} t^k (1-t)^{n-k}$, $k = 0, 1, \ldots, n$. По теореме 11.2 гл. II эти многочлены определяют крайние лучи пространства многочленов $\mathcal{P}_{n+1}$, и легко проверить, что $y^{(k)} \in P^n$, где

$$y_i^{(k)} = \begin{cases} 0, & i = 0, \ldots, k-1, \\ (-1)^{i-k} \dfrac{(n+1)!}{(n-i)!(i-k)!k!}, & i = k, \ldots, n, \end{cases}$$

являются коэффициентами многочлена $(n+1) B_{nk}(t)$. Матрицы с элементами $y_i^{(k)}$ и $x_i^{(k)}$ являются по существу обратными, так как

$$(\mathbf{y}^{(k)}, \mathbf{x}^{(l)}) = \begin{cases} n+1, & \text{если } k = l, \\ 0, & \text{если } k \neq l. \end{cases} \tag{5.7}$$

Симплекс, натянутый на $n+1$ точек $\mathbf{y}^{(k)}$, обозначим через $B^n$. Заметим, что $B^n$ вписывается в $P^n$.

Если обозначить конусы, индуцированные симплексами $B^n$ и $S^n$ через $\mathcal{B}_{n+1}$ и $\mathcal{Y}_{n+1}$ соответственно, то можно сформулировать следующую теорему.

Теорема 5.2. *$\mathcal{B}_{n+1}$ и $\mathcal{Y}_{n+1}$ являются дуальными выпуклыми конусами* (см. § 9 гл. II).

Доказательство. Дуальный конус к $\mathcal{Y}_{n+1}$, определяемый как $\mathcal{B} = \{\mathbf{y} \mid (\mathbf{y}, \mathbf{x}) \geqslant 0$ для всех $\mathbf{x} \in Y_{n+1}\}$, очевидно, представляет собой множество $\{\mathbf{y} \mid (\mathbf{y}, \mathbf{x}^{(k)}) \geqslant 0, \ldots, n\}$. Из (5.7) очевидно, что это множество содержит $B^n$. Таким образом, имеем $\mathcal{B}_{n+1} \subset \mathcal{B}$. Обратно, если $\mathbf{y}$ такой, что $(\mathbf{y}, \mathbf{x}^{(k)}) = a_k \geqslant 0$ для $k = 0, 1, \ldots, n$, и через $\mathbf{A}$ мы обозначим матрицу, у которой $k$-я строка равна $\mathbf{x}^{(k)}$, то $\mathbf{A}\mathbf{y}' = (a_0, \ldots, a_n)'$ ($\mathbf{y}'$ есть вектор-столбец, определенный из $\mathbf{y}$). Это означает, что $\mathbf{y}'$ определяется однозначно, так как $|\mathbf{A}| \neq 0$. Однако

$$\mathbf{z} = \sum_{k=0}^{n} \frac{a_k}{n+1} \mathbf{y}^{(k)}$$

также удовлетворяет той же самой системе уравнений. Отсюда заключаем, что

$$\mathbf{y} = \sum_{k=0}^{n} \frac{a_k}{n+1} \mathbf{y}^{(k)}.$$

Это в свою очередь означает, что $\mathbf{y}$ принадлежит $\mathcal{B}_{n+1}$. Следовательно, $\mathcal{B} \subset \mathcal{B}_{n+1}$ и $\mathcal{B}_{n+1}$ является дуальным конусом к конусу $\mathcal{Y}_{n+1}$.

Так как конус $\mathcal{Y}_{n+1}$ замкнут, то он должен быть дуальным к конусу $\mathcal{B}_{n+1}$ (см. лемму 9.1 гл. II).

Следующая теорема показывает, что симплекс $B^n$ в известном смысле почти заполняет симплекс $P^n$.

Теорема 5.3. *Если* $\mathbf{y}$ — *внутренняя точка* $P^n$, *то для достаточно большого* $m$ *точка*

$$\mathbf{y}_{(m)} = (y_0, y_1, \ldots, y_n, 0, \ldots, 0) \qquad (m+1 \text{ координат})$$

*принадлежит* $B^m$.

Доказательство. Так как $\mathbf{y} = (y_0, y_1, \ldots, y_n) \in \operatorname{Int} P^n$, то многочлен $\sum_{i=0}^{n} y_i t^i$ не обращается в нуль на $[0, 1]$, так что для некоторого положительного $\delta$ $\sum_{i=0}^{n} y_i t^i \geqslant \delta > 0$, $t \in [0, 1]$. Однако так как

$$x_i(m, t) = \frac{(tm)(tm-1)\ldots(tm-i+1)}{m(m-1)\ldots(m-i+1)}$$

сходятся (равномерно по $t$) к $t^i$, когда $m \to \infty$, получаем, что $\sum_{i=0}^{n} y_i x_i(m, t) \geqslant \delta/2$ для достаточно большого $m$. Поэтому, положив $t = k/m$ или $tm = k$, видим, что $(\mathbf{y}_{(m)}, \mathbf{x}_{(m)}^{(k)}) \geqslant 0$ для достаточно большого $m$, где $x_{(m)}^{(k)}$, $k = 0, 1, \ldots, m$, являются вершинами симплекса $S^m$. Как и в доказательстве теоремы 5.2, это означает, что $\mathbf{y}_{(m)} \in B^m$ для достаточно большого $m$.

Теорема 5.4. *Для данной последовательности* $\{c_i\}_0^\infty$ *усеченная точка* $\mathbf{c}_{(m)} = (c_0, \ldots, c_m)$ *принадлежит* $S^m$ *для всех* $m$, *если и только если* $\mathbf{c}_{(m)}$ *принадлежит* $M^m$ *для всех* $m$.

Доказательство. Если $\mathbf{c}_{(m)} \in M^m$ для всех $m$, то $\mathbf{c}_{(m)} \in S^m$ для всех $m$ по теореме 5.1. Теперь предположим, что $\mathbf{c}_{(m)} \in S^m$ для всех $m$ и $\mathbf{c}_m \notin M^n$ для некоторого фиксированного $n$. Тогда существует положительный многочлен степени $n$ с коэффициентами $y = (y_0, \ldots, y_n)$, $\sum_{i=0}^{n} y_i^2 > 0$ такой, что $\sum_{i=0}^{n} y_i c_i < 0$. Но по теореме 5.3 $(y_0, \ldots, y_n, 0, \ldots, 0) = \sum_{k=0}^{m} \lambda_k \mathbf{y}_{(m)}^{(k)}$ для достаточно большого $m$, где $\mathbf{y}_{(m)}^{(k)}$ — вершины симплекса $B^m$ и $\lambda_k \geqslant 0$.

Следовательно, по теореме 5.2 $\sum_{i=0}^{n} y_i c_i = \sum_{k=0}^{m} \lambda_k (y_{(m)}^{(k)}, \mathbf{c}_{(m)}) \geqslant 0$. Из этого противоречия и следует утверждение теоремы.

Переводя теоремы 5.3 и 5.4 с геометрического языка, получим следующие две хорошо известные теоремы. Первая из них является переформулировкой теоремы 5.3, в то время как вторая следует из теоремы 5.4 и выражений (5.4) и (5.6), связывающих величины $\Delta^k c_i$ с барицентрическими координатами точки $\mathbf{c}$ в $S^{k+1}$.

Теорема 5.5. *Любой многочлен* $\sum_{i=0}^{n} y_i t^i$, *который строго положителен на* $[0, 1]$, *может быть представлен как конечная сумма с неотрицательными коэффициентами многочленов вида*

$$B_{mk}(t) = \binom{m}{k} t^k (1-t)^{m-k} \qquad k = 0, 1, \ldots, m,$$

*при условии, что* $m$ *достаточно большое.*

Теорема 5.6. *Необходимое и достаточное условие для того, чтобы* $(c_0, c_1, \ldots, c_n, \ldots)$ *были моментами некоторой функции распределения на* $[0, 1]$, *состоит в том, что последовательность должна быть вполне монотонной, т. е. что* $c_i$ *и их последовательные разности* $(-1)^k \Delta^k c_i$ *всех порядков должны быть неотрицательными.*

## § 6. Объем симплекса

Вспомним, что $M^n$ является сечением моментного пространства $\mathcal{M}_{n+1}$, определенным условием нормировки $c_0 = \int_0^1 d\sigma(t) = 1$.

Наша цель в настоящий момент состоит в расчете объема моментного сечения $M^n$ (который мы будем обозначать $\operatorname{Vol}(M^n)$). Это оказывается необходимым при решении изопериметрических задач (см., например, Шенберг [1954 b]). В теореме 6.1, приведенной ниже, мы выведем формулу для объема сечения, соответствующего $M^n$ для общих $T$-систем. Теорема 6.2 устанавливает более частный результат для случая $u_i(t) = t^i$, $i = 0, 1, \ldots, n$.

Пусть $\{u_i\}_0^n$ является $T$-системой на $[0, 1]$ такой, что $u_0, u_1, \ldots, u_n$ — непрерывно дифференцируемые функции. Предположим, кроме того, что $u_0(t) \equiv 1$ и $u_i(0) = 0$, $i = 1, 2, \ldots, n$.

Если система $\{u_i\}_0^n$ не обладает этими свойствами, но $u(t) \equiv 1$ есть многочлен, то можно преобразовать систему $\{u_i\}_0^n$ в новую систему $\{v_i\}_0^n$, а именно, $v_i = \sum_{j=0}^{n} a_{ij} u_j$, где $\mathbf{A} = \|a_{ij}\|$ есть соответствующая матрица размером $(n+1) \times (n+1)$, $|\mathbf{A}| > 0$, и выполняются условия $v_0(t) \equiv 1$, $v_i(0) = 0$, $i = 1, 2, \ldots, n$. Этого можно добиться, если определить первую строку как $(a_{00}, a_{01}, \ldots, a_{0n})$, где $\sum_{j=0}^{n} a_{0j} u_j(t) \equiv 1$. Остальные строки конструируются так, чтобы они были ортогональны вектору $(u_0(0), \ldots, u_n(0))$ и взаимно линейно независимы. Легко видеть, что при этих условиях определитель $|\mathbf{A}|$ не равен нулю, а в таком случае мы можем зафиксировать его знак и считать положительным. Введенная система $\{v_i\}_0^n$, очевидно, обладает требуемыми свойствами, т. е. $v_0(t) \equiv 1$ и $v_i(0) = 0$, $i = 1, 2, \ldots, n$. Из тождества $|v_i(t_j)| = |\mathbf{A}|\,|u_i(t_j)|$ следует, что $\{v_i\}_0^n$ является $T$-системой.

Теорема 6.1. *Если $\{u_i\}_0^n$ есть $T$-система из класса $C^1$, удовлетворяющая условиям $u_0(t) \equiv 1$ и $u_i(0) = 0$, $i = 1, \ldots, n$, то для $n = 2m - 1$*

$$\mathrm{Vol}(M^n) = \frac{1}{n!} \int \cdots \int\limits_{0<t_1<\ldots<t_m<1} D(t_1, \ldots, t_m)\, dt_1 \ldots dt_m,$$

*где*

$$D(t_1, \ldots, t_m) = \begin{vmatrix} u_0(t_1) & u_0'(t_1) & u_0(t_2) & u_0'(t_2) & \ldots & u_0(t_m) & u_0'(t_m) \\ u_1(t_1) & u_1'(t_1) & u_1(t_2) & u_1'(t_2) & \ldots & u_1(t_m) & u_1'(t_m) \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ u_{2m-1}(t_1) & u_{2m-1}'(t_1) & u_{2m-1}(t_2) & u_{2m-1}'(t_2) & \ldots & u_{2m-1}(t_m) & u_{2m-1}'(t_m) \end{vmatrix}$$

*и*

$$D(t_1, \ldots, t_m) = \begin{vmatrix} u_1(t_1) & u_1'(t_1) & u_1(t_2) & u_1'(t_2) & \ldots & u_1(t_m) & u_1'(t_m) \\ u_2(t_1) & u_2'(t_1) & u_2(t_2) & u_2'(t_2) & \ldots & u_2(t_m) & u_2'(t_m) \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ u_{2m}(t_1) & u_{2m}'(t_1) & u_{2m}(t_2) & u_{2m}'(t_2) & \ldots & u_{2m}(t_m) & u_{2m}'(t_m) \end{vmatrix},$$

*если $n = 2m$.*

Доказательство. Если $n=2m-1$, то каждая точка $\mathbf{c} \in \mathrm{Int} M^n$ имеет единственное нижнее главное представление

$$c_i = \sum_{j=1}^{m} \lambda_j u_i(t_j), \qquad i = 1, 2, \ldots, 2m-1, \tag{6.1}$$

где $0 < t_1 < t_2 < \ldots < t_m < 1$ и $\sum\limits_{j=1}^{m} \lambda_j = 1$. Следовательно, мы имеем, что

$$\mathrm{Vol}(M^n) = \int \cdots \int\limits_{M^n} dc_1 \ldots dc_m = \int\limits_{U^m} \int\limits_{\Xi^{m-1}} |J|\, d\lambda_2 \ldots d\lambda_m dt_1 \ldots dt_m,$$

где

$$U^m = \{(t_1, \ldots, t_m) \mid 0 < t_1 < \ldots < t_m < 1\},$$

$$\Xi^{m-1} = \left\{(\lambda_2, \ldots, \lambda_m) \mid \lambda_i \geqslant 0,\ \sum_{2}^{m} \lambda_i \leqslant 1\right\}$$

и $J$ — якобиан преобразования (6.1), причем $\lambda_1$ заменяется на $1-\sum_2^m \lambda_i$. Легко проверить, что якобиан $J$ равен

$$J=\left(1-\sum_2^m \lambda_i\right)\left(\prod_{i=2}^m \lambda_i\right) D(t_1, \ldots, t_m).$$

Требуемый результат теперь следует после интегрирования по переменным $\{\lambda_i\}_{i=2}^m$. Если $n=2m$, то анализ проводится аналогично. Преобразование (6.1) остается тем же самым, за исключением того, что в этом случае мы имеем $\sum_{i=1}^m \lambda_i \leq 1$, $\Xi^{m-1}$ заменяется на $\Xi^m$, а член $\prod_1^m \lambda_i$ выносится за якобиан.

Теорема 6.1 обобщает соответствующую теорему для степенных функций, приведенную у Карлина и Шепли [1953]. В работе Шенберга [1954 b] приведена дискуссия по расчету объемов некоторых выпуклых тел, который связан в свою очередь с изопериметрической задачей, имеющей отношение к кривым в евклидовых пространствах и объему, заключенному в их выпуклой оболочке.

Теорема 6.2 *Если $u_i(t)=t^i$, $i=0, \ldots n$, то*

$$\operatorname{Vol}(M^n)=\prod_{k=1}^n \frac{\Gamma(k)\Gamma(k)}{\Gamma(2k)}=\prod_{k=1}^n \mathrm{B}(k, k),$$

*где $\Gamma$ обозначает обычную гамма-функцию, а $\mathrm{B}(p, q)$ — бета-функция указанных параметров.*

Доказательство. Если $n=2m-1$, то определитель $D(t_1, \ldots, t_m)$, стоящий под знаком интеграла в формулировке теоремы 6.1, равен

$$\prod_{j=2}^m \prod_{k=1}^{j-1}(t_k-t_j)^4.$$

Чтобы убедиться в этом, заменим $2i$-й столбец, $i=1, \ldots, m$, на столбец с элементами $1, s_i, s_i^2, \ldots, s_i^{2m-1}$, где $t_i<s_i<t_{i+1}$, и обратим внимание на то, что если получившийся определитель продифференцировать последовательно по переменным $s_1, s_2, \ldots, s_m$ и вычислить при $t_1, \ldots, t_m$, то получим первоначальный определитель $D(t_1, \ldots, t_m)$. Следовательно, мы можем вычислить вспомогательный определитель Вандермонда, а затем произвести дифференцирование, чтобы получить желаемый результат. Таким образом,

$$\operatorname{Vol}(M^n)=\frac{1}{n!m!}\int_0^1 \ldots \int_0^1 \prod_{k<j}(t_k-t_j)^4\, dt_1 \ldots dt_m.$$

По результату Сельберга [1944] значение вышеприведенного интеграла равно

$$\prod_{k=1}^{m} \frac{\Gamma(2k+1)\,[\Gamma(2k-1)]^2}{2\Gamma(2m+2k-2)},$$

откуда, как легко видеть, заключение теоремы получается индукцией по $m$. Случай $n = 2m$ рассматривается аналогично.

## § 7. Свойства симметрии моментных пространств

**А. Граница сечения $M^n$.** Прежде чем изучать различные свойства симметрии, присущие сечению $M^n$, сделаем несколько замечаний, относящихся к его границе. Во-первых, предположим, что моментное пространство $\mathcal{M}_{n+1}$ образуется $T$-системой $\{u_i\}_0^n$ на $[a, b]$, причем такой, что функция $u(t) \equiv 1$ является многочленом Большинство последующих результатов будет относиться к системе $\{t^i\}_0^n$ на $[0, 1]$, где могут быть сделаны более определенные утверждения.

Как и в классическом случае, пусть $M^n$ является сечением $\mathcal{M}_{n+1}$, заданным мерами с общей массой, равной единице. Мы классифицируем точки $\mathbf{c} = (c_0, c_1, \ldots, c_n) \in M^n$ в соответствии с их индексом $I(\mathbf{c})$, т. е. определим $M_k^n = \{\mathbf{c} \mid \mathbf{c} \in M^n,\ 2I(\mathbf{c}) - 1 = k\}$, $k = 0, 1, \ldots, n$.

Множество $M_n^n$ состоит из внутренних точек $M^n$. Для $0 \leqslant k < n$ точки множества $M_k^n$ принадлежат границе $M^n$, так что по теореме 2.1 гл. II любая точка $\mathbf{c} \in M_k^n$ имеет единственное представление. Полезно разбить каждое из множеств $M_k^n$, $k = 0, 1, \ldots, n-1$, на два множества $\underline{M}_k^n$ и $\overline{M}_k^n$, состоящих соответственно из тех точек, представления которых не включают граничную точку $b$, и тех точек, представления которых содержат точку $b$ Множество $M_k^n$ называют $k$-мерной гранью $M^n$, в то время как $\overline{M}_k^n$ и $\underline{M}_k^n$ называются верхней и нижней $k$-мерными гранями соответственно.

Лемма 7.1. *Множества $\overline{M}_k^n$ и $\underline{M}_k^n$, $0 \leqslant k < n$, являются максимально связными компонентами множества $M_k^n$ и имеют топологическую размерность $k$.*

Доказательство. Рассмотрим только нижнюю $k$-мерную грань $\underline{M}_k^n$ для случая $k = 2l - 1$. Любая точка $\mathbf{c} = (c_0, c_1, \ldots, c_n) \in \underline{M}_{2l-1}^n$ имеет единственное представление

$$c_i = \sum_{j=1}^{l} \lambda_j u_i(t_j), \qquad i = 0, 1, \ldots, n, \tag{7.1}$$

где $\lambda = (\lambda_1, \ldots, \lambda_l)$ содержится в $\Xi^{l-1} = \left\{\lambda \mid \lambda_i > 0,\ i = 1, 2, \ldots, l,\ \sum_{i=1}^{l} \lambda_i = 1\right\}$ и $\mathbf{t} = (t_1, t_2, \ldots, t_l)$ принадлежит $U^l = \{\mathbf{t} \mid a < t_1 < t_2 < \ldots < t_l < b\}$. Отображение (7.1) прямого произведения $U^l \times \Xi^{l-1}$ на $\underline{M}_k^n$ является взаимно однозначным и непрерывным в обе стороны Следовательно, множество $\underline{M}_k^n$ является связным и имеет топологическую размерность $k = 2l - 1$ Подобное доказательство может быть использовано для других $k$-мерных граней

Отделимость множеств $\underline{M}_k^n$ и $\overline{M}_k^n$ может быть выведена из соотношений

$$[\underline{M}_k^n] \setminus \underline{M}_k^n = \bigcup_{i=0}^{k-1} M_i^n = [\overline{M}_k^n] \setminus \overline{M}_k^n,$$

так как в этом случае $[\underline{M}_k^n] \cap \overline{M}_k^n = \varnothing = [\overline{M}_k^n] \cap \underline{M}_k^n$ [1]).

**B. Обратное распределение**. Для удобства зафиксируем интервал $[a, b] = [0, 1]$ и пусть $\mathscr{D}$ представляет собой множество неубывающих и непрерывных справа функций $\sigma$, определенных на $(-\infty, \infty)$ и таких, что $\sigma(t) = 1$ для $t \geqslant 1$ и $\sigma(t) = 0$ для $t < 0$. Каждому из $\mathscr{D}$ поставим в соответствие обычным путем меру, и наоборот.

Зададим отображение $g$: $D \to D$ следующей формулой

$$(g\sigma)(t) = \begin{cases} \sup\{s \mid \sigma(s) \leqslant t\}, & 0 \leqslant t < 1, \\ \sigma(t), & t < 0 \text{ или } t \geqslant 1 \end{cases}$$

По этому определению $g(g\sigma) = \sigma$, так что функция $g\sigma$ по существу является обратной к $\sigma$.

Для частного случая $u_i(t) = t^i$, $i = 0, 1, \ldots, n$, соответствующие этим функциям первые моменты связаны посредством тождества $c_1(\sigma) = 1 - c_1(g\sigma)$, где $c_1(\sigma) = \int\limits_0^1 t d\sigma(t)$. Оказывается, что подобные соотношения не имеют силы для моментов более высокого порядка

Рассмотрим ступенчатую функцию

$$\sigma(t) = \sum_{i=1}^{m} \lambda_i I(t - t_i), \qquad 0 \leqslant t_1 < t_2 < \ldots < t_m \leqslant 1,$$

где $I(t - t_i) = \begin{cases} 1, & t \geqslant t_i, \\ 0, & t < t_i, \end{cases}$ и $\sum\limits_{i=1}^{m} \lambda_i = 1$. При отображении $g$ ступенчатая функция $\sigma$ переводится в ступенчатую функцию $g\sigma$ со скачками $t_1, t_2 - t_1, t_3 - t_2, \ldots$ $\ldots, t_m - t_{m-1}, 1 - t_m$ в точках $0, \lambda_1, \lambda_1 + \lambda_2, \ldots, \lambda_1 + \ldots + \lambda_{m-1}, 1$. Следовательно, приходим к выводу, что индекс функции $g\sigma$ равен индексу функции $\sigma$ и что функция $g\sigma$ имеет скачок при $t = 1$, если и только если функция $\sigma$ его не имеет. Так как каждая точка $\mathbf{c}$, принадлежащая границе $M^n$, определяется единственной функцией $\sigma \in \mathscr{D}$, то отображение $g$, следовательно, индуцирует взаимно однозначное отображение $g_0$ границы $M^n$ на само себя, определенное посредством $g_0(\mathbf{c}) = \mathbf{c}(g\sigma)$, где $\mathbf{c}(g\sigma)$ является моментной точкой, образованной функцией $g\sigma$.

Теорема 7.1. *Для T-системы $\{u_i\}_0^n$ на $[0, 1]$ такой, что $u(t) \equiv 1$ является многочленом, отображение $g_0$ переставляет верхнюю и нижнюю k-мерные грани на $M^n$, т. е. $g_0(\underline{M}_k^n) = \overline{M}_k^n$ и $g_0(\overline{M}_k^n) = \underline{M}_k^n$.*

**C. Обращение единичного интервала**. Хотелось бы отметить, что функция $g_0$ была определена только на границе $M^n$. Один путь индуцирования взаимно однозначного отображения всего пространства $M^n$ на само себя состоит в рассмотрении взаимно однозначного непрерывного отображения $h$ единичного интервала $[0, 1]$ на самого себя. Для любого такого отображения $h$ существует соответ-

[1]) Здесь через $[M]$ обозначается замыкание множества $M$, ниже для этой цели будет также ставиться черта сверху, здесь это не делается, чтобы не внести путаницы.

ствующее взаимно однозначное отображение пространства $\mathscr{D}$ на само себя, которое переводит $\sigma$ в меру $\sigma_h$, определенную как $\sigma_h(A) = \sigma(h^{-1}(A))$ для любого множества $A \subset [0, 1]$. Если это отображение на $\mathscr{D}$ должно индуцировать отображение пространства $M^n$ на само себя, то необходимо тогда потребовать, чтобы индуцированное отображение $F_h$ удовлетворяло условию $F_h(\mathbf{c}(\sigma)) = \mathbf{c}(\sigma_h)$ или

$$F_h\left(\int u_0 d\sigma, \ldots, \int u_n d\sigma\right) = \left(\int u_0 d\sigma_h, \ldots, \int u_n d\sigma_h\right) =$$

$$= \left(\int u_0(h(t))\, d\sigma(t), \ldots, \int u_n(h(t))\, d\sigma(t)\right). \tag{7.2}$$

Отсюда мы можем сделать вывод, что $F_h(\alpha\mathbf{c}(\sigma) + \beta\mathbf{c}(\mu)) = \alpha F_h(\mathbf{c}(\sigma)) + \beta F_h(\mathbf{c}(\mu))$, так что отображение $F_h$ линейно на $M^n$.

В оставшейся части этого параграфа мы теперь ограничимся рассмотрением системы функций $\{t^i\}_0^n$ на интервале $[0, 1]$. В этом случае, предположив, что $F_h$ линейная, и используя равенство

$$F_h(1, t, \ldots, t^n) = (1, h(t), \ldots, h^n(t)),$$

получим $h(t) = \sum_{k=0}^{n} a_k t^k$, где $a_k$, $k = 0, \ldots, n$,—константы. Далее, для того чтобы $(1, h(t), \ldots, h^n(t))$ содержалась в $M^n$, потребуем, чтобы $a_k = 0$ для $k = 2, 3, \ldots, n$. Следовательно, единственная функция $h$ на $[0, 1]$, которая приводит к нетривиальному отображению $F_h$, дается формулой

$$h_0(t) = 1 - t. \tag{7.3}$$

В этом случае мы имеем $\sigma_{h_0}(t) = 1 - \sigma(1 - t - 0)$. Предельная операция, указанная символом «—0» служит для того, чтобы сохранить непрерывность справа. Моментная точка $\mathbf{c}(\sigma)$ тогда преобразуется в точку с координатами

$$c_k(\sigma_{h_0}) = \int_0^1 t^k d\sigma_{h_0}(t) = \int_0^1 (1 - t)^k d\sigma = \sum_{i=0}^{k} (-1)^i \binom{k}{i} c_i(\sigma) = (-1)^k \Delta^k c_0(\sigma).$$

Теорема 7.2. *Функция $h_0(t) = 1 - t$ отображает $M^n$ в соответствии с формулой*

$$\mathbf{c}(\sigma_{h_0})' = \mathbf{A}_n \mathbf{c}(\sigma)', \tag{7.4}$$

*где $\mathbf{A}_n$ — $(n+1) \times (n+1)$-матрица с элементами*

$$(\mathbf{A}_n)_{ij} = (-1)^j \binom{i}{j}, \quad i, j = 0, 1, \ldots, n, \tag{7.5}$$

$\mathbf{c}(\sigma) = (c_0(\sigma), \ldots, c_n(\sigma))$ *со штрихом обозначает результат транспонирования, или, иначе, вектор-столбец.*

Барицентрические координаты $\lambda_{nk}(\sigma) = \int_0^1 \binom{n}{k} t^k (1-t)^{n-k} d\sigma$, определенные равенством (5.5), удовлетворяют условию $\lambda_{nk}(\sigma_{h_0}) = \lambda_{n\, n-k}(\sigma)$, так что

$$\mathbf{c}(\sigma_{h_0}) = \sum_{k=0}^{n} \lambda_{nk}(\sigma_{h_0})\, \mathbf{x}^{(k)} = \sum_{k=0}^{n} \lambda_{nk}(\sigma)\, \mathbf{x}^{(n-k)}$$

(см. равенства (5.1) и (5.6)), и отображение $F_{h_0}$, определенное на $M^n$, может рассматриваться как преобразование точки $\mathbf{x}^{(k)}$ в $\mathbf{x}^{(n-k)}$. Если функция $\sigma$ имеет скачки $\lambda_1, \ldots, \lambda_m$ в точках $t_1, \ldots, t_m$ соответственно, то $\sigma_{h_0}$—ступенчатая функция со скачками $\lambda_m, \ldots, \lambda_1$ в точках $1-t_m, \ldots, 1-t_1$. Следовательно, можно прийти к такому выводу

Теорема 7.3. *Матрица* $\mathbf{A}_n$ *преобразует k-мерные грани* $M^n$ *в соответствии с формулами*

$$\mathbf{A}_n(\underline{M}_k^n) = \overline{M}_k^n, \quad \mathbf{A}_n(\overline{M}_k^n) = \underline{M}_k^n, \quad \text{если } k \text{—четное},$$

$$\mathbf{A}_n(\underline{M}_k^n) = \underline{M}_k^n, \quad \mathbf{A}_n(\overline{M}_k^n) = \overline{M}_k^n, \quad \text{если } k \text{—нечетное}.$$

Пусть $\bar{\mathbf{c}} = (c_0, \ldots, c_{n-1}, \bar{c}_n) \in M^n$, и предположим, что $\underline{\mathbf{c}} = (c_0, \ldots, c_{n-1}, \underline{c}_n)$ и $\bar{\mathbf{c}} = (c_0, \ldots, c_{n-1}, \bar{c}_n)$ в $M^n$ определены посредством

$$\bar{c}_n = \max\left\{c_n \,\Big|\, \int t^i d\sigma = c_i, \qquad i = 0, 1, \ldots, n-1\right\},$$

$$\underline{c}_n = \min\left\{c_n \,\Big|\, \int t^i d\sigma = c_i, \qquad i = 0, 1, \ldots, n-1\right\}.$$

Тогда можно получить следующее следствие.

Следствие 7. 3. *Для n нечетного* $\mathbf{A}_n\mathbf{c}' = \overline{\mathbf{A}_n\mathbf{c}'}$ и $\mathbf{A}_n\bar{\mathbf{c}}' = \underline{\mathbf{A}_n\mathbf{c}'}$, в то время как для $n$ четного $\mathbf{A}_n\underline{\mathbf{c}}' = \underline{\mathbf{A}_n\mathbf{c}'}$ и $\mathbf{A}_n\bar{\mathbf{c}}' = \overline{\mathbf{A}_n\mathbf{c}'}$.

В пространстве многочленов $P_n$ индуцированное преобразование задается как

$$P(t) \to P(1-t),$$

где $P(t) = \sum_{i=0}^{n} a_i t^i$. Если мы обозначим индуцированное преобразование на $P^n$ через $G$, то можно записать

$$G(\mathbf{a}) = (b_0, b_1, \ldots, b_n),$$

где $b_i = \sum_{k=i}^{n} (-1)^i \binom{k}{i} a_k$.

Теорема 7. 4. *Функция G, индуцированная функцией* $h_0(t) = 1-t$, *преобразует точки из* $P^n$ *в соответствии с формулой*

$$G(\mathbf{a}) = \mathbf{a}\mathbf{A}_n,$$

где $\mathbf{a} = (a_0, a_1, \ldots, a_n)$, а $\mathbf{A}_n$ определена в (7.5).

Возвращаясь к определителям Ганкеля, можно установить следующие соотношения.

Теорема 7.5. *При отображении* (7.4)

$$\underline{\Delta}_{2k}(\mathbf{A}_n\mathbf{c}') = \underline{\Delta}_{2k}(\mathbf{c}), \qquad \overline{\Delta}_{2k}(\mathbf{A}_n\mathbf{c}') = \overline{\Delta}_{2k}(\mathbf{c}),$$

$$\underline{\Delta}_{2k+1}(\mathbf{A}_n\mathbf{c}') = \underline{\Delta}_{2k+1}(\mathbf{c}), \qquad \overline{\Delta}_{2k+1}(\mathbf{A}_n\mathbf{c}') = \overline{\Delta}_{2k+1}(\mathbf{c}),$$

где $\underline{\Delta}_{2k}(\mathbf{c})$ *обозначает определитель* (1.3), *построенный по отношению к точке* $\mathbf{c} = (c_0, c_1, \ldots, c_n) \in M^n$, *и т. д.*

Доказательство. Пусть $\underline{\mathbf{D}}_{2k}(\mathbf{c})$ обозначает матрицу, соответствующую определителю $\underline{\Delta}_{2k}(\mathbf{c})$. Путем непосредственной проверки из равенства (7.5) следует, что $\mathbf{A}_k\underline{\mathbf{D}}_{2k}(\mathbf{c}) = \underline{\mathbf{D}}_{2k}(\mathbf{A}_n\mathbf{c}')\mathbf{A}_k'$, где $\mathbf{A}_k'$ обозначает результат транспонирования $\mathbf{A}_k$. Следовательно, $\underline{\Delta}_{2k}(\mathbf{A}_n\mathbf{c}') = \underline{\Delta}_{2k}(\mathbf{c})$ Другие утверждения теоремы доказываются аналогично.

Используя теорему 7.5, можно показать, что матрица $\mathbf{A}_n$ сохраняет ширину в направлении $x_n$ сечения $M^n$ во всех точках, другими словами, справедливо следующее следствие.

Следствие 7.5. *Выполняется соотношение*

$$\bar{c}_n(\sigma) - \underline{c}_n(\sigma) = \bar{c}_n(\sigma_{h_0}) - \underline{c}_n(\sigma_{h_0}). \tag{7.6}$$

Доказательство. Если $(c_0(\sigma), \ldots, c_{n-1}(\sigma))$ — граничная точка $M^{n-1}$, то обе части равенства (7.6) обращаются в нуль. Для $(c_0(\sigma), \ldots, c_{n-1}(\sigma))$ внутри $M^{n-1}$ равенство (7.6) следует из теоремы 7.5 и соотношений

$$\underline{c}_n = c_n - \frac{\underline{\Delta}_n}{\underline{\Delta}_{n-2}}.$$

$$\bar{c}_n = c_n + \frac{\bar{\Delta}_n}{\bar{\Delta}_{n-2}}.$$

которые могут быть получены из того, что $\underline{c}_n$ и $\bar{c}_n$ служат решениями уравнений $\underline{\Delta}_n = 0$ и $\bar{\Delta}_n = 0$ соответственно (см доказательство теоремы 4.4).

**D. Моментные пространства симметричных распределений**. Пусть $M_S^n$ обозначает моментное пространство, образованное из системы функций $\{t^i\}_0^n$ такими распределениями $\sigma \in \mathscr{D}$, которые удовлетворяют условию $\sigma(t) = 1 - \sigma(1 - t - 0)$. Очевидно, $M_S^n$ состоит из таких точек $\mathbf{c} = (c_0, c_1, \ldots, c_n)$, что $\mathbf{c}' = \mathbf{A}_n\mathbf{c}'$

Теорема 7.6. *Множество $M_S^n$ выпукло, замкнуто, ограничено и имеет размерность* $[n/2]$.

Доказательство. Заметим, что $M_S^n$ в точности совпадает с пересечением $M^n$ и гиперплоскостей, заданных линейной системой уравнений $\mathbf{c}\mathbf{A}_n' = \mathbf{c}$ Решения этих уравнений могут быть представлены для каждого нечетного момента в виде выражений, зависящих от четных моментов более низкого порядка, т. е.

$$c_1 = \frac{1}{2}$$

$$c_3 = \frac{3}{2}c_2 - \frac{1}{4},$$

$$c_5 = \frac{5}{2}c_4 - \frac{5}{2}c_2 + \frac{1}{2} \text{ и т. д.,} \tag{7.7}$$

которые представляют собой ровно $n - [n/2]$ независимых линейных связей (заметим, что если бы мы выбрали $[-1, 1]$ как основной интервал, то соотношения (7.7) просто устанавливали бы, что $c_i = 0$, $i = 1, 3, 5, \ldots$). Следовательно, так как $M^n$ является выпуклым телом в $E^n$, то $M_S^n$ — выпуклое множество размерности $[n/2]$, а так как $M^n$ замкнуто и ограничено, то $M_S^n$ также замкнуто и ограничено.

Вышеприведенная теорема наводит на мысль, что имеется соответствие между пространством $M_S^n$ и пространством $M^{[n/2]}$ В действительности имеет место следующая теорема.

Теорема 7.7 *Существует взаимно однозначное линейное преобразование $T$ пространства $M_S^n$ на $M^{[n/2]}$.*

Доказательство. Прежде всего отметим, что замена основного интервала представляет собой взаимно однозначное линейное преобразование моментных пространств В частности, существует взаимно однозначное преобразование $V$ пространства $M_S[0, 1]$ на пространство $M_S^n[-1, 1]$. Взаимно однозначное соответствие между симметричными распределениями на этих двух интервалах может быть определено следующим образом:

$$\varphi \to \psi: \int_0^1 f(t)\, d\varphi = \int_{-1} f(t^2)\, d\psi,$$

$$\psi \to \varphi: \int_{-1}^1 g(t)\, d\psi = \frac{1}{2}\left[\int_0^1 g(\sqrt{t})\, d\varphi + \int_0^1 g(-\sqrt{t})\, d\varphi\right],$$

где равенства имеют место для всех непрерывных функций $f$ и $g$ на $[0, 1]$ и $[-1, 1]$ соответственно Моменты от $\psi$ и $\varphi$ связаны посредством соотношения

$$c_{2k}(\psi) = c_k(\varphi), \quad c_{2k+1}(\psi) = 0, \qquad k = 0, 1, 2,$$

Требуемое отображение тогда получается композицией отображений $T = WV$, где $W$ отображает $M_S^n[-1,1]$ на $M^{[n/2]}$ и задается как

$$W(c_0(\psi), c_1(\psi), \ldots, c_n(\psi)) = (c_0(\varphi), c_1(\varphi), \ldots c_{[n/2]}(\varphi))$$

В теореме 1.2 гл. II было показано, что сечение $M^n$ есть выпуклая оболочка кривой $C_{n+1}$, образованной точкой $\mathbf{c}(t) = (1, t, \ldots, t^n)$, когда $t$ изменяется на интервале $[0, 1]$. Для $M_S^n$ кривая $C_S^n$, образованная посредством

$$\mathbf{c}_S(t) = \frac{1}{2}[\mathbf{c}(t) + \mathbf{c}(1-t)], \qquad t \in \left[0, \frac{1}{2}\right],$$

играет ту же роль, что и $C_{n+1}$

Теорема 7.8. *Выпуклой оболочкой $\mathcal{C}(C_S^n)$ кривой $C_S^n$ служит $M_S^n$.*

Доказательство. Очевидно, $\mathcal{C}(C_S^n) \subset M_S^n$. Если $\mathbf{c} \in M_S^n$, то $\mathbf{c}$ имеет конечное представление, образованное распределением $\sigma$. Точка $\mathbf{c}$, которая образуется посредством распределения $2^{-1}[\sigma(t) + \sigma(1-t-0)]$, принадлежит тогда выпуклой оболочке кривой $C_S^n$

Следствие 7.8. *Множество крайних точек $M_S^n$* для $n \geqslant 4$ в точности совпадает с $C_S^n$.

Доказательство. для $n \geqslant 4$ существует неотрицательный многочлен $P$, имеющий нулями $t$ и $1-t$. Так как $P$ порождает опорную плоскость к $M_S^n$, которая касается только в точке $\mathbf{c}_S(t)$, то эта точка должна быть крайней.

## § 8. Механические квадратуры и главные представления

В этом параграфе мы выведем некоторые классические формулы, используемые в численном интегрировании, опираясь на геометрические положения рассмотренной выше теории.

В общей постановке задача механических квадратур состоит в аппроксимации величины интеграла

$$\int_a^b f(t)\,dt \tag{8.1}$$

конечной суммой вида

$$\sum_{i=1}^{p} \lambda_i f(t_i). \tag{8.2}$$

Значения $t_1, \ldots, t_p$ и $\lambda_1, \ldots, \lambda_p$ характеризуют конкретную формулу численного интегрирования, т. е. с использованием этих чисел величина суммы (8.2) предлагается в качестве аппроксимации интеграла (8.1). Кроме того, для некоторого множества функций выполняется равенство квадратурной формулы и интеграла. В частности, если $\{u_i\}_0^n$ есть $T$-система на интервале $[\alpha, b]$, а $\{\lambda_i\}_1^p$ и $\{t_i\}_1^p$ обозначают веса и корни, соответствующие некоторому представлению моментной точки

$$c_i = \int_a^b u_i(t)\,dt, \qquad i = 0, 1, \ldots, n, \tag{8.3}$$

то

$$\int_a^b f(t)\,dt = \sum_{i=1}^{p} \lambda_i f(t_i) \tag{8.4}$$

для любого многочлена $f(t) = \sum_{i=0}^{n} a_i u_i(t)$. В частности, если $u_i(t) = t^i$, $i = 0, 1, \ldots, n$, а $n = 2m+1$ и мы выбрали нижнее главное представление моментной точки (8.3) при $[a, b] = [-1, 1]$, то равенство

$$\int_{-1}^{1} P(t)\,dt = \sum_{j=1}^{m+1} \lambda_p^* P(t_j^*) \tag{8.5}$$

справедливо для любого многочлена $P(t) = \sum_{i=0}^{2m+1} a_i t^i$ степени $2m+1$. Точки $t_1^*, t_2^*, \ldots, t_{m+1}^*$ удовлетворяют условию

$$-1 < t_1^* < t_2^* < \ldots < t_{m+1}^* < 1 \tag{8.6}$$

и на основании леммы 2.2 являются нулями многочлена Лежандра $P_{m+1}(t)$. Формула (8.5) известна как гауссова квадратурная формула и обладает характеристической двойной точностью по отношению ко многим квадратурным формулам, т. е. для $m+1$ определенных точек $t_1^*, t_2^*, \ldots, t_{m+1}^*$ соотношение (8.5) выполняется для $2m+2$ функций $1, t, \ldots, t^{2m+1}$.

Следующая квадратурная формула, включающая сплайн-многочлен, представляет значительный интерес в численном анализе (см., например, Шенберг [1958]).

Рассмотрим $n+1=l+r+1$ $(l \geqslant 1)$ функций

$$\begin{aligned} u_i(t) &= t^i, & i &= 0, 1, \ldots, l, \\ u_{l+j}(t) &= (t-x_j)_+^l, & j &= 1, \ldots, r, \end{aligned} \tag{8.7}$$

где

$$t_+^l = \begin{cases} t^l, & t \geqslant 0, \\ 0, & t < 0 \end{cases} \tag{8.8}$$

и

$$-1 < x_1 < x_2 < \ldots < x_r < 1 \tag{8.9}$$

— фиксированные точки. В примере 11 § 3 гл. I, было показано, что система (8.7) является $WT$-системой на $[-1, 1]$.

Пусть $\sigma_0$ обозначает конечную меру на $[-1, 1]$ такую, что моментная точка с координатами

$$c_i^0 = \int_{-1}^{1} u_i(t)\, d\sigma_0(t), \quad c = 0, 1, \ldots, n, \tag{8.10}$$

содержится внутри моментного пространства $\mathcal{M}_{n+1}^{(u)}$, образованного специальными функциями $u_0, u_1, \ldots, u_n$, определенными выше.

Теорема 8.1. *Если мера $\sigma_0$ такова, что $\mathbf{c}^0 = (c_0^0, c_1^0, \ldots, c_n^0) \in \in \operatorname{Int} \mathcal{M}_{n+1}^{(u)}$, то $\mathbf{c}^0$ имеет верхнее и нижнее главные представления индекса $(n+1)/2$.*

Замечание 8.1. Для дальнейшего заметим, что теорема 8.1 выполняется для $WT$-системы $\{(t-x_i)_+^m\}_{i=0}^n$, определенной на любом интервале $[a, b]$ и такой, что $a \leqslant x_0 < x_n \leqslant b$ (см. пример 10 § 3 гл. I).

После рассмотрения доказательства теоремы 8.1 мы опишем ее применение к квадратурным формулам. Применяя нижнее главное представление точки $\mathbf{c}^0$, получим следующее следствие.

Следствие 8.1а. *Если $n = 2m+1$, то существует квадратурная формула*

$$\sum_{\nu=1}^{m+1} \lambda_\nu^* f(t_\nu^*) \sim \int_{-1}^{1} f(t)\, d\sigma_0(t), \tag{8.11}$$

где

$$-1 < t_1^* < t_2^* < \ldots < t_{m+1}^* < 1, \quad \lambda_1^* > 0, \ldots, \lambda_{m+1}^* > 0, \tag{8.12}$$

*такая, что соотношение* (8.11) *является точным равенством для любой сплайн-функции вида*

$$f(t) = \sum_{i=0}^{l} a_i t^i + \sum_{i=1}^{r} b_i (t - x_i)_+^l. \tag{8.13}$$

Моментная точка (8.10), индуцированная мерой Лебега $d\sigma_0(t) = dt$, лежит внутри $\mathcal{M}_{n+1}^{(u)}$. Это есть очевидное следствие того факта, что функции $u_0, u_1, \ldots, u_n$ линейно независимы. Формальное доказательство этого утверждения здесь опущено. В специальном случае, когда $d\sigma_0(t) = dt$ и $r = 0$, соотношения (8.11) сводятся к гауссовой квадратурной формуле (8.5). Кроме того, (8.11) представляет собой всегда квадратурную формулу с $m + 1$ узлами, обладающую характеристической двойной точностью по отношению к заранее заданному множеству $l + r + 1 = n + 1 = 2(m + 1)$ функций

$$1, t, \ldots, t^l, (t - x_1)_+^l, (t - x_2)_+^l, \ldots, (t - x_r)_+^l.$$

Соответствующий результат, использующий верхнее главное представление, состоит в следующем.

Следствие 8.1 b. *Если* $n = 2m + 1$, *то существует квадратурная формула*

$$\sum_{\nu=1}^{m+2} \alpha_\nu^* f(s_\nu^*) \sim \int_{-1}^{1} f(t)\, d\sigma(t), \tag{8.14}$$

*где*

$$-1 = s_1^* < s_2^* < \ldots < s_{m+1}^* < s_{m+2}^* = 1, \ \alpha_\nu^* > 0, \ \nu = 1, \ldots m + 2, \tag{8.15}$$

*такая, что соотношение* (8.14) *является точным для любой сплайн-функции вида* (8.13).

В случае $d\sigma(t) = dt$ и $r = 0$ (8.14) сводится к формуле Радау (см. Радау [1880]). Квадратурная формула Радау приводится у Шенберга [1958]. Тесные связи между различными квадратурными формулами и каноническими представлениями подчеркиваются Крейном [1951] и Карлином и Шепли [1953].

Две другие квадратурные формулы также применимы, когда $n = 2m$. При этом узлы в (8.12) и (8.15) заменяются соответственно на

$$-1 = t_1^* < t_2^* < \ldots < t_{m+1}^* \qquad (t_{m+1}^* < 1) \tag{8.16}$$

и

$$s_1^* < s_2^* < \ldots < s_{m+1}^* = 1 \qquad (s_1^* > -1).$$

Доказательство теоремы 8.1. В примере 11 § 3 гл. I было показано, что система (8.7) является $WT$-системой на интервале $[-1, 1]$. Анализ доказательства показывает, что система (8.7)—$WT$-система на любом интервале $[a, b]$ при условии $a < x_1$ и $x_r < b$. В этом случае модифицированная система $\{v_i(t;\delta)\}_0^n$, определенная посредством

$$v_i(t;\delta)=\int\limits_{-1-\varepsilon}^{1+\varepsilon} G_\delta(x,t)\,u_i(x)\,dx, \tag{8.17}$$

где

$$G_\delta(x,t)=\frac{1}{\sqrt{2\pi\delta}}\exp\left\{-\frac{1}{2}\left(\frac{x-t}{\delta}\right)^2\right\},$$

образует $T$-систему на $[-1, 1]$ (ср. (3.16) гл. I). Мы расширили пределы в (8.17), чтобы гарантировать предельное соотношение

$$\lim_{\delta\to 0} v_i(t;\delta)=u_i(t),\qquad t\in[-1,1], \tag{8.18}$$

причем сходимость — равномерная на *замкнутом* интервале $[-1, 1]$.

Теперь докажем, что любая мера $\mu$, моментная точка которой является внутренней по отношению к $\mathcal{M}_{n+1}$, с необходимостью имеет спектр индекса по крайней мере $(n+1)/2$. Мы будем действовать методом от противного. Предположим, что $\mu$ удовлетворяет условиям, указанным выше, и имеет индекс, меньший чем $(n+1)/2$. По теореме 5.1 гл. I существует многочлен $v(t;\delta)=\sum\limits_{i=0}^{n} a_i(\delta)\,v_i(t;\delta)$, который обращается в нуль на спектре меры $\mu$. Можно предположить, что многочлен $v(t;\delta)$ нормирован в соответствии с условием $\sum\limits_{i=0}^{n}[a_i(\delta)]^2=1$. Выбрав подпоследовательность $\delta_k\to 0$ такую, что $\lim\limits_{k\to\infty} a_i(\delta_k)=a_i$, $i=0, 1, \ldots, n$, и принимая во внимание (8.18), делаем вывод о существовании неотрицательного многочлена $u(t)=\sum\limits_{i=0}^{n} a_i u_i(t)$, $\sum\limits_{i=0}^{n} a_i^2=1$, который обращается в нуль на спектре меры $\mu$. В этом случае мы видим, что $\sum\limits_{i=0}^{n} a_i c_i^0=0$ и $\sum\limits_{i=0}^{n} a_i c_i \geqslant 0$ для всех $c\in\mathcal{M}_{n+1}^{(u)}$. Следовательно, $(c_0^0, c_1^0, \ldots, c_n^0)$ лежит на границе $\mathcal{M}_{n+1}^{(u)}$ в нарушении предположения, что $c^0\in\operatorname{Int}\mathcal{M}_{n+1}^{(u)}$. Полученное противоречие означает, что $I(\mu)\geqslant(n+1)/2$. Таким образом, любая мера $\mu$, которая имеет скачок во внутренней точке моментного пространства, с необходимостью имеет индекс по крайней мере $(n+1)/2$. В этом случае мера $\sigma_0$ порождает моментную точку

$$c_i^0(\delta)=\int\limits_{-1}^{1} v_i(t;\delta)\,d\sigma_0(t),\qquad i=0, 1, \ldots n, \tag{8.19}$$

содержащуюся внутри моментного пространства $\mathcal{M}_{n+1}^{(v)}$, образованного $T$-системой $\{v_i(t;\delta)\}$. На основании теории, изложенной в гл. II, точке (8.19) доставляют верхнее и нижнее главные представления меры $\bar{\sigma}(t;\delta)$ и $\underline{\sigma}(t;\delta)$. Из (8.18) следует, что $v_0(t;\delta)\to 1$, когда $\delta\to 0$, так что

$$c_0^0(\delta)=\int_{-1}^{1} v_0(t;\delta)\,d\underline{\sigma}(t;\delta)\geqslant(1-\varepsilon)\int_{-1}^{1} d\underline{\sigma}(t;\delta).$$

Это неравенство в сочетании со сходимостью $c_0^0(\delta)$ к $c_0^0$ показывает, что интегралы $\int_{-1}^{1} d\sigma(t;\delta)$ имеют равномерно ограниченную вариацию при $\delta$, близком к нулю. Подобные аргументы используются и в случае мер $\bar{\sigma}(t;\delta)$.

Теперь можно выбрать подпоследовательность $\delta_k\to 0$ такую, что $\bar{\sigma}(t;\delta_k)$ и $\underline{\sigma}(t;\delta_k)$ слабо сходятся к мерам $\bar{\sigma}$ и $\underline{\sigma}$ соответственно. Эти меры, очевидно, имеют спектр индекса не выше чем $(n+1)/2$. Кроме того,

$$c_i^0=\int_{-1}^{1} u_i(t)\,d\sigma_0(t)=\lim_{k\to\infty}\int_{-1}^{1} v_i(t;\delta_k)\,d\sigma_0(t)=$$

$$=\lim_{k\to\infty}\int_{-1}^{1} v_i(t;\delta_k)\,d\underline{\sigma}(t;\delta_k)=\int_{-1}^{1} u_i(t)\,d\underline{\sigma}(t)$$

ввиду равномерной сходимости в (8.18).

Мера $\underline{\sigma}$, следовательно, удовлетворяет (8.10) и имеет индекс $(n+1)/2$. Аналогичные утверждения справедливы для $\bar{\sigma}$. Мы, таким образом, представили по крайней мере одно верхнее и одно нижнее главные представления точки $(c_0^0, c_1^0, \ldots, c_0^n)$, и доказательство завершено.

## § 9. Сплайн-функции с заданными нулями

В этом параграфе демонстрируется использование главного представления некоторых моментных точек для того, чтобы установить существование сплайн-многочлена с заданным множеством нулей.

Функция вида

$$P_{n,k}(t)=x^n+P_{n-1}(t)+\sum_{\nu=1}^{k} c_\nu(t-\xi_\nu)_+^{n-1}, \tag{9.1}$$

где $P_{n-1}(t)$ — многочлен степени $n-1$, называется *моносплайном* степени $n$ с узлами $\xi_1<\ldots<\xi_k$. Это понятие было введено Шенбергом [1946] и играет важную роль в задаче подгонки кривых и теории аппроксимации (см. также Сард [1963] и Джонсон [1960]).

Для наших последующих целей установим аналог основной теоремы алгебры в случае моносплайнов. С этой целью необходимо прежде всего точно определить кратность нуля в приложении к моносплайнам Под кратностью нуля $\xi$ моносплайна понимается то же самое, что и в случае обычных многочленов, если $\xi$ не является одним из узлов $\xi_1, \ldots, \xi_k$. Так как $(t-\xi)_+^{n-1}$ принадлежит классу непрерывности $C^{n-2}$, то обычное понятие кратности нуля также применяется, когда кратность не превышает $n-1$ для узла $\xi$. Остается рассмотреть возможность, что $P_{n,k}^{(j)}(\xi)=0$, $j=0, \ldots, n-2$, и $\xi$ является одним из узлов $\xi_1, \ldots, \xi_k$. Пусть $P_{n,k}^{(n-1)}(t)$ представляет собой $(n-1)$-ю производную справа от моносплайна $P_{n,k}(t)$. Тогда

$$P_{n,k}^{(n-1)}(t) = at + b + \sum_{\nu=1}^{k} b_\nu (t-\xi_\nu)_+^0, \qquad a>0, \tag{9.2}$$

где $t_+^0 = \begin{cases} 1, & t \geqslant 0. \\ 0, & t<0. \end{cases}$

Говорят, что моносплайн $P_{n,k}(t)$ имеет нуль порядка $n$ при $t=\xi_\nu$, если $P_{n,k}^{(n-1)}(t)$ изменяет знак в $\xi_\nu$. Если $P_{n,k}^{(n-1)}(t)$ не изменяет знак в $\xi_\nu$, то нуль имеет порядок $n+1$, если выполняется $P_{n,k}^{(n-1)}(\xi_\nu+)= =0$ или $P_{n,k}^{(n-1)}(\xi_\nu-)=0$, и порядок $n-1$ в случае, когда производная $P_{n,k}^{(n-1)}(t)$ отделена от нуля в узле $\xi_\nu$.

Легко устанавливаются следующие свойства.

(1) Максимальный порядок любого нуля моносплайна $P_{n,k}(t)$ равен $n+1$.

(2) $P_{n,k}(t)$ изменяет знак в $\xi$ тогда и только тогда, когда порядок нуля нечетный.

(3) Нуль порядка $m$ моносплайна $P_{n,k}(t)$ является нулем $P'_{n,k}(t)$ порядка $m-1$.

Так как производная справа $P_{n,k}^{(n-1)}(t)$ имеет вид (9.2), то она обладает, самое большее, $2k+1$ нулями. Применяя соответствующим образом теорему Ролля, приходим к следующему результату.

(4) Моносплайн (9.1) может иметь, самое большее, $n+2k$ нулей с учетом кратности.

Теперь обратим внимание на задачу построения моносплайна, имеющего заданное множество нулей. В качестве аналога основной теоремы алгебры может рассматриваться следующая теорема существования.

Теорема 9.1. *Пусть заданы $n+2k$ точек $\{x_i\}_{i=1}^{n+2k}$, удовлетворяющих условию $x_1<x_2<\ldots<x_{n+2k}$. Тогда существует единственный моносплайн $P_{n,k}(t)$, обращающийся в нуль точно на множестве $\{x_i\}_{i=1}^{n+2k}$.*

Эта теорема была опубликована в работе Шенберга [1958], но ее доказательство приведено не было. Метод, предложенный в его статье, оказался неработоспособным в общем случае. Карлин и Шумейкер [1966] предложили доказательство теоремы 9.1, которое также включает ряд усовершенствований и обобщение на общие

$T$-системы. Оказалось, что анализ должен быть исключительно тонким.

Джонсон [1960] использовал теорему 9.1 при рассмотрении наилучшей аппроксимации нулевой функции моносплайнами.

Результат теоремы 9.1 может быть обычным образом распространен на тот случай, когда некоторые из точек $x$ совпадают, при условии, что совокупность равных значений $x$ включает не больше чем $n+1$ членов. Таким образом, существует единственный моносплайн $P_{n,k}(t)$ с заданным множеством нулей $\{x_i\}_{i=1}^{n+2k}$, где совокупность равных значений $x$ интерпретируется как нуль соответствующей кратности моносплайна $P_{n,k}(t)$.

Мы приводим доказательство теоремы 9.1 для специального случая, который может быть сведен к ситуации, когда применима геометрическая теория моментных пространств.

Пусть заданы $n+2k$ точек $\{t_\nu\}_1^{n+2k}$, удовлетворяющих условию $t_1 = t_2 = \ldots = t_n < t_{n+1} < \ldots < t_{n+2k}$. Если $t_1 < \xi_1 < \ldots < \xi_k$, то моносплайн

$$(t-t_1)^n - \sum_{\nu=1}^{k} c_\nu (t-\xi_\nu)_+^{n-1}, \tag{9.3}$$

очевидно, имеет нуль порядка $n$ в точке $t_1$. Наша цель заключается в том, чтобы установить существование узлов $\{\xi_\nu\}_1^k$ и соответствующих коэффициентов $\{c_\nu\}_1^k$ таких, что моносплайн (9.3) обращается в нуль в точках $t_{n+1}, \ldots, t_{n+2k}$.

Рассмотрим моментное пространство, образованное $WT$-системой

$$u_i(\xi) = (t_{n+i}-\xi)_+^{n-1}, \qquad i = 1, \ldots, 2k, \tag{9.4}$$

определенной на $[a, b]$, где $a = t_1$ и $b = t_{n+2k}$. Из представления

$$\int_a^b (t_{n+i}-\xi)_+^{n-1}\, d\xi = \frac{(t_{n+i}-t_1)^n}{n}, \qquad i = 1, \ldots, 2k, \tag{9.5}$$

легко видно, что $2k$-мерная точка с координатами (9.5) является внутренней по отношению к моментному пространству, образованному функциями (9.4). Обращаясь к теореме 8.1 (см. замечание 8.1), обозначим через $\underline{\sigma}$ нижнюю главную меру точки с координатами (9.5). Если $\{\underline{\xi}_\nu, \underline{c}_\nu\}_1^k$ представляют собой корни и веса меры $\underline{\sigma}$, то

$$\sum_{\nu=1}^{k} \underline{c}_\nu (t_{n+i}-\underline{\xi}_\nu)_+^{n-1} = \frac{(t_{n+i}-t_1)^n}{n}, \qquad i = 1, \ldots, 2k.$$

Следовательно, многочлен

$$(t-t_1)^n - \sum_{\nu=1}^{k} n\underline{c}_\nu (t-\underline{\xi}_\nu)_+^{n-1} \tag{9.6}$$

имеет простые нули в точках $t_{n+1}, \ldots, t_{n+2k}$, а также нуль порядка $n$ в точке $t_1$. По свойству (4), приведенному перед формулировкой теоремы 9.1, не может быть других нулей, отличных от тех, которые указаны.

В случае, когда некоторые из точек $t_{n+1}, \ldots, t_{n+2k}$ совпадают, используются аналогичные аргументы, за исключением того, что некоторые из функций (9.4) заменяются соответствующими производными от $(t-\xi)_+^{n-1}$.

Пример 9.1. Предположим, мы хотим построить моносплайн на $[-1, 1]$ степени $n = 2k$, включающий $k$ узлов $\xi_1, \ldots, \xi_k$ и имеющий нули кратности $2k$ в каждой из точек $t=-1$ и $t=+1$. Аналогично, если мы хотим определить узлы $\{\xi_\nu\}_1^k$ и соответствующие коэффициенты $\{c_\nu\}_1^k$, для которых моносплайн

$$(t+1)^{2k} + \sum_{\nu=1}^{k} c_\nu (t-\xi_\nu)_+^{2k-1}$$

имеет нуль порядка $2k$ в точке $t=+1$. Рассмотрим моментное пространство, образованное функциями

$$1, (1-\xi), (1-\xi)^2, \ldots, (1-\xi)^{2k-1}, \qquad -1 \leq \xi \leq 1.$$

Так как эта система является $T$-системой и

$$\int_{-1}^{1} (1-\xi)^j \, d\xi = \frac{2^{j+1}}{j+1}, \qquad j = 0, 1, \ldots, 2k-1, \tag{9.7}$$

воспользуемся нижней главной мерой $\underline{\sigma}$, соответствующей моментной точке $(2, 2^2/2, 2^3/3, \ldots, 2^{2k}/2k)$. В частности, существуют точки $\{\underline{\xi}_\nu\}_1^k$, $-1 < \underline{\xi}_1 < \ldots < \underline{\xi}_k < 1$, и положительные константы $\{c_\nu\}_1^k$ такие, что

$$\sum_{\nu=1}^{k} \underline{c}_\nu (1-\underline{\xi}_\nu)^j = \frac{2^{j+1}}{j+1}, \qquad j = 0, 1, \ldots, 2k-1.$$

Тогда легко видеть, что многочлен

$$(t+1)^{2k} - 2k \sum_{\nu=1}^{k} \underline{c}_\nu (t-\underline{\xi}_\nu)_+^{2k-1} \tag{9.8}$$

имеет нуль порядка $2k$ в каждой из точек $t=+1$ и $t=-1$, как и требовалось. Узлы $\underline{\xi}_1 < \ldots < \underline{\xi}_k$ могут быть идентифицированы в соответствии с леммой 2.2 как нули $k$-го многочлена Лежандра. Моносплайн (9.8) известен как ядро Пеано и играет роль в формуле гауссовых механических квадратур.

Взамен нижней главной меры $\underline{\sigma}$ моментной точки (9.7) мы могли бы так же выбрать верхнюю главную меру $\bar{\sigma}$. Соответствующая сплайн-функция принимает вид

$$(t+1)^{2k} - 2k \sum_{\nu=1}^{k+1} \bar{c}_\nu (t - \bar{\xi}_\nu)_+^{2k-1},$$

где $-1 = \bar{\xi}_1 < \bar{\xi}_2 < \ldots < \bar{\xi}_k < \bar{\xi}_{k+1} = 1$. Рассмотрение производной порядка $(2k-1)$ от функции, определяемой соотношением (9.9), обнаруживает существование нуля порядка $2k+1$ в каждой из точек $t = \pm 1$. Узлы $\{\bar{\xi}_\nu\}_2^k$ совпадают с множеством нулей производной от $k$-го многочлена Лежандра, а (9.9) есть ядро Пеано, которое встречается в формуле механических квадратур Радау.

Приведенное выше рассмотрение частично дополняет работу Шенберга [1958]; исчерпывающие подробности можно найти у Карлина и Шумейкера [1966].

Глава V

# ЧЕБЫШЕВСКИЕ СИСТЕМЫ И ПРОСТРАНСТВА МОМЕНТОВ НА ИНТЕРВАЛЕ $[0, \infty)$

## § 1. Введение

Система из $n+1$ функций $u_0, u_1, \ldots, u_n$ называется *T-системой* на $[0, \infty)$, если $\{u_i\}_{i=0}^n$ — $T$-система на каждом конечном интервале $[0, A]$, где $A > 0$; т. е. если функции $u_0, u_1, \ldots, u_n$ непрерывны на $[0, \infty)$ и

$$\det [\, \| u_i (t_j) \|_{i,j=0}^n ] > 0$$

при любом выборе $0 \leq t_0 < t_1 < \ldots < t_n < \infty$. Понятия $CT$-систем и $ET$-систем переносятся на $[0, \infty)$ аналогично.

В этой главе мы развиваем соответствующую теорию глав I—IV для случая $T$-систем, определенных на полубесконечном интервале. В случае полубесконечного интервала точка $+\infty$ будет играть такую же роль, как и конец интервала в конечном случае. В связи с этим будет необходимо наложить соответствующие ограничения на поведение функций $u_i$ на бесконечности.

При соответствующих предположениях большая часть результатов будет получена при помощи перехода к вспомогательной $T$-системе, определенной на компактном интервале $[0, \infty]$ и естественным образом подобной нашей первоначальной системе $\{u_i\}_{i=0}^n$. Затем, используя уже полученные теоремы, мы сформулируем соответствующие результаты для исходной системы.

## § 2. Общая теорема о конической оболочке

В теореме 1.2 гл. II было показано, что если $\{u_i\}_{i=0}^n$ — $T$-система на конечном интервале $[a, b]$, то $\mathcal{M}_{n+1}$ является выпуклой конической оболочкой кривой $C_{n+1}$, образованной точками $(u_0 (t), u_1 (t), \ldots, u_n(t))$ при $t$, пробегающем интервал $[a, b]$. В этом случае моментное пространство $\mathcal{M}_{n+1}$ является замкнутым множеством. Теперь мы докажем более общую теорему, в которой моментное пространство не обязательно замкнуто. Так как доказательство теоремы не опирается ни на свойство функций $u_0, u_1, \ldots, u_n$ образовывать $T$-систему, ни на тот факт, что рассматриваемый интервал одномерен, то мы сформулируем теорему в большей общности, чем требуется для наших целей.

Пусть $z_1(t), z_2(t), \ldots, z_m(t)$ — непрерывные вещественнозначные функции, определенные для $t \in T$, где $T$ — замкнутое (не обязательно ограниченное) подмножество $R^r$ евклидова $r$-мерного пространства, снабженное стандартной топологией, индуцируемой открытыми множествами $R^r$. Мы образуем моментное пространство

$$\mathcal{M}_m^{(z)} = \mathcal{M}_m(T; z_1, z_2, \ldots, z_m) =$$
$$= \{\mathbf{c} = (c_1, \ldots, c_m) | c_i = \int_T z_i(t)\, d\sigma(t), \quad i = 1, 2, \ldots, m\}, \tag{2.1}$$

где $\sigma$ пробегает множество неотрицательных регулярных мер, ограниченной вариации, определенных на $T$ так, чтобы функции $z_1, z_2, \ldots$ $\ldots, z_m$ были абсолютно инте рируемы по отношению к $\sigma$. Рассмотрим подмножество $\mathcal{M}_m^{(z)}$, образованное точечными мерами, т. е. пусть

$$C_m = \{\mathbf{z}(t) = (z_1(t), z_2(t), \ldots, z_m(t)) | t \in T\} \tag{2.2}$$

Мы докажем следующую общую теорему.

Теорема 2.1. *Выпуклый конус, натянутый на $C_m$, есть $\mathcal{M}_m^{(z)}$. Символически: $\mathcal{C}(C_m) = \mathcal{M}_m^{(z)}$. (Необходимо подчеркнуть, что* $\mathcal{M}_m^{(z)}$ не обязано быть замкнутым.)

Доказательство. Доказывать будем индукцией по $m$. Если $m = 1$, то результат получается просто. Предположим, что теорема верна для случая $m - 1$. Нам требуется доказать, что

$$\mathcal{C}(C_m(T)) \supset \mathcal{M}_m(T, z_1(t), \ldots, z_m(t)),$$

так как противоположное включение очевидно. Пусть $\mathbf{c}$ — внутренняя точка (относительно $R^m$) $\mathcal{M}_m^{(z)}$. Мы можем определить ограниченное замкнутое подмножество $\widetilde{T} \subset T$ такое, что $\mathbf{c}$ является также внутренней точкой $\mathcal{M}_m^{(z)}(\widetilde{T})$.

Методом теоремы 1.2 гл. II можно показать в этом случае, что $\mathcal{C}(C_m(\widetilde{T})) = \mathcal{M}_m^{(z)}(\widetilde{T})$. Остается рассмотреть $\mathbf{c} \in \operatorname{Bd} \mathcal{M}_m^{(z)}$. В этом случае мы можем найти опорную гиперплоскость для $\mathcal{M}_m^{(z)}$ в $\mathbf{c}$, т. е. существуют постоянные $\{a_i\}_{i=1}^m$, не все равные нулю, такие, что

$$\sum_{i=1}^m a_i z_i(t) \geqslant 0, \quad t \in T \text{ и } \sum_{i=1}^m a_i c_i = 0. \tag{2.3}$$

Но $c_i = \int z_i(t)\, d\sigma^*(t)$ $(i = 1, 2, \ldots, m)$ для некоторого $\sigma^*$. Принимая во внимание первое из этих соотношений, получаем, что спектр $\sigma^*(t)$

заключен в замкнутом подмножестве $\hat{T} \subset T$, на котором $\sum_{i=1}^{m} a_i z_i(t)$ обращается в нуль; так как $\{a_i\}_{i=1}^{m}$ — не все нули, то мы можем представить, например, $z_m(t)$ как линейную комбинацию $z_1(t), z_2(t), \ldots$ $\ldots, z_{m-1}(t)$ на $\hat{T}$. Теперь мы построим моментное пространство $\mathcal{M}_{m-1}(\hat{T}, z_1, z_2, \ldots, z_{m-1})$ относительно $z_1(t), z_2(t), \ldots, z_{m-1}(t)$, где $t$ изменяется на $\hat{T}$,

Из индукционного предположения следует, что

$$\mathcal{M}_{m-1}(\hat{T}; z_1, z_2, \ldots, z_{m-1}) = \mathcal{C}(C_{m-1}(\hat{T})). \tag{2.4}$$

Ясно, что $(c_1, \ldots, c_{m-1})$ принадлежит $\mathcal{M}_{m-1}(\hat{T}, z_1, z_2, \ldots, z_{m-1})$. Фактически, выбрав $\sigma^*$, мы видим, что

$$c_i = \int_T z_i(t)\, d\sigma^*(t) = \int_{\hat{T}} z_i(t)\, d\sigma^*(t), \qquad i = 1, 2, \ldots, m.$$

Из (2.4) следует, что существует мера $\tilde{\sigma}$ с конечным спектром, содержащимся в $\hat{T}$, для которой

$$c_i = \int_T z_i(t)\, d\tilde{\sigma}(t), \qquad i = 1, 2, \ldots, m-1. \tag{2.5}$$

Обращаясь ко второму из соотношений в (2.3), мы также убеждаемся в том, что

$$c_m = \int_T z_m(t)\, d\tilde{\sigma}(t). \tag{2.6}$$

Представления (2.5) и (2.6) показывают, что $c \in \mathcal{C}(C_m(T))$. Теорема доказана.

Более общую форму теоремы о конической оболочке (теорема 2.1) можно найти у Рогозинского [1958]. Частные случаи этой теоремы встречаются у Карлина и Шепли [1953] и у Крейна [1951].

## § 3. Общая структура $\mathcal{M}_{n+1}$

Мы теперь рассмотрим моментные пространства на полубесконечном интервале $[a, \infty)$ (для удобства пусть $a = 0$). В этом параграфе будем предполагать, что $u_0(t), u_1(t), \ldots, u_n(t)$ образуют $T$-систему действительных непрерывных функций на $[0, \infty)$. Мы постулируем существование по крайней мере одного строго положительного многочлена $u(t) = \sum_{i=0}^{n} a_i u_i(t)$, удовлетворяюще о соотношениям $u(t) > 0$, $0 \leq t < \infty$ и $\liminf_{t \to \infty} u(t) > 0$.

Введем кривую

$$C_{n+1} = \{u(t) = (u_0(t), u_1(t), \ldots, u_n(t)) \mid 0 \le t < \infty\}$$

и моментное пространство

$$\mathcal{M}_{n+1} = \{\mathbf{c} = (c_0, c_1, \ldots, c_n) \mid c_i = \int_0^\infty u_i(t)\, d\sigma(t),\ i = 0, \ldots, n,\ \sigma \in \mathcal{D}\}, \tag{3.1}$$

где $\mathcal{D}$ — множество неубывающих функций ограниченной вариации на $[0, \infty)$, для которых все интегралы существуют абсолютно. Ясно, что пространство $\mathcal{M}_{n+1}$ является выпуклым конусом в $E^{n+1}$.

Переход с конечных на полубесконечные интервалы вносит существенные изменения в строение соответствующих моментных пространств. Например, вообще говоря, $\mathcal{M}_{n+1}$ не будет больше замкнутым, замкнутость будет достигаться только посредством введения несобственных распределений, допускающих «массу на бесконечности». Конечно, чтобы построить содержательную теорию, потребуется наложить дополнительные ограничения на $u_i$.

Мы рассмотрим три типа часто встречающихся ситуаций.

I. Предположим, что $\lim\limits_{t\to\infty} u_i(t) = u_i(\infty)$ существует и конечен для $i = 0, 1, \ldots, n$ и

$$U\begin{pmatrix} 0, & 1, & \ldots, & n-1, & n \\ t_0, & t_1, & \ldots, & t_{n-1}, & \infty \end{pmatrix} > 0$$

для всех $0 \le t_0 < t_1 < \ldots < t_{n-1} < \infty$. Если мы теперь будем обращаться с бесконечной точкой, как с обычными точками, допуская, что меры $\bar{\sigma}$ имеют скачки на бесконечности, то анализ конуса $\mathcal{M}_{n+1}$ проводится так же, как и в случае конечного интервала, изученного в гл. I — IV. Точнее, пусть

$$\bar{u}_i(s) = u_i(\operatorname{tg} s), \qquad 0 \le s < \pi/2,$$
$$\bar{u}_i\left(\frac{\pi}{2}\right) = u_i(\infty), \qquad i = 0, 1, \ldots, n.$$

Ясно, что функции $\bar{u}_i$ образуют $T$-систему непрерывных функций на $[0, \pi/2]$. Ее моментное пространство является замкнутым конусом, который мы обозначим $\overline{\mathcal{M}}_{n+1}$. Каждая мера на $[0, \pi/2]$ может быть представлена в форме $\bar{\sigma}(s) = \sigma(s) + M(s)$, где $\sigma$ не имеет массы в $\pi/2$ и $M(s)$ является мерой с массой $M$ в $\pi/2$. Из того, что

$$\int_0^{\pi/2} \bar{u}_i(s)\, d\bar{\sigma}(s) = \int_0^\infty u_i(t)\, d\sigma(\operatorname{arc\,tg} t) + M u_i(\infty)$$

следует, что [1]) $\overline{\mathcal{M}}_{n+1}=\mathcal{M}_{n+1}+\mathcal{R}_{\infty}^{(u)}$, где $\mathcal{R}_{\infty}^{(u)}$ — положительный луч, порожденный $\mathbf{u}(\infty)=(u_0(\infty), u_1(\infty), \ldots, u_n(\infty))$. При $t\to\infty$ $(u_0(t), u_1(t), \ldots, u_n(t))\to(u_0(\infty), u_1(\infty), \ldots, u_n(\infty))$, так что $\overline{\mathcal{M}}_{n+1}=$ $=\mathcal{M}_{n+1}+\mathcal{R}_{\infty}^{(u)}$ содержится в замыкании $\mathcal{M}_{n+1}$. Так как $\overline{\mathcal{M}}_{n+1}$ замкнуто, то оно в действительности является замыканием $\mathcal{M}_{n+1}$, что оправдывает черту в его обозначении. Наконец, так как замыкание $\overline{C}_{n+1}$ множества $C_{n+1}$ равно $C_{n+1}+\mathbf{u}(\infty)$, то мы получаем $\mathcal{C}(\overline{C}_{n+1})=\mathcal{C}(C_{n+1})+\mathcal{R}_{\infty}^{(u)}$. Но по теореме 2.1 $\mathcal{C}(C_{n+1})=\mathcal{M}_{n+1}$, так что окончательно

$$\mathcal{C}(\overline{C}_{n+1})=\mathcal{M}_{n+1}+\mathcal{R}_{\infty}^{(u)}=\overline{\mathcal{M}}_{n+1}.$$

В ряде случаев, когда будет встречаться система типа I, мы будем применять результаты гл. I — IV прямо к системе $\{u_i\}_{i=0}^{n}$, определенной на $[0, \infty]$.

II. ($i$) Предположим, что существует такое $A>0$, что для $t\geqslant A$

$$\lim_{t\to\infty}\frac{u_i(t)}{u_n(t)}=0 \text{ для } i<n.$$

(ii) В дополнение к тому, что полная система $u_0(t),\ldots,u_n(t)$ является системой Чебышева, мы постулируем, что $u_0(t), \ldots, u_{n-1}(t)$ также является $T$-системой на $[0, \infty)$. Это предположение сделано для того, чтобы система (3.3), определенная ниже, была $T$-системой на $[0, \infty]$. Необходимо отдельно проверить неравенство для определителя, когда последний аргумент $t_n=\infty$ (см. (3.4) ниже).

Мы теперь покажем, что $T$-система типа II может быть преобразована в $T$-систему типа I, сохранив все существенные особенности. В самом деле, положим

$$w(t)=\begin{cases}u_n(A), & 0\leq t\leq A,\\ u_n(t), & A\leq t<\infty,\end{cases}\tag{3.2}$$

и определим

$$v_i(t)=\frac{u_i(t)}{w(t)},\qquad i=0, 1, \ldots, n.\tag{3.3}$$

Тогда

$$v_i(\infty)=\lim_{t\to\infty}v_i(t)=\begin{cases}0, & i<n,\\ 1, & i=n,\end{cases}$$

$$V\begin{pmatrix}0, \ldots, n\\ t_0, \ldots, t_n\end{pmatrix}=\frac{1}{\prod\limits_{0}^{n}w(t_i)}U\begin{pmatrix}0, \ldots, n\\ t_0, \ldots, t_n\end{pmatrix}>0$$

[1]) $A+B$ обозначает множество векторов вида $a+b$, $a\in A$, $b\in B$.

для $0 \leq t_0 < \ldots < t_n < \infty$ и

$$V\begin{pmatrix} 0, \ldots, n-1, & n \\ t_0, \ldots, t_{n-1}, & \infty \end{pmatrix} = \frac{1}{\prod\limits_{i=0}^{n-1} w(t_i)} U\begin{pmatrix} 0, \ldots, n-1 \\ t_0, \ldots, t_{n-1} \end{pmatrix} > 0. \qquad (3.4)$$

Поэтому, применяя результаты, установленные для $T$-системы типа I к системе $\{v_i\}$, мы получим $\overline{\mathcal{M}}_{n+1}^{(v)} = \mathcal{M}_{n+1}^{(v)} + \mathcal{R}_\infty^{(v)}$, где $\mathcal{R}_\infty^{(v)}$ — положительный луч, порожденный $(v_0(\infty), v_1(\infty), \ldots, v_n(\infty)) = (0, 0, \ldots, 1)$.

Лемма 3.1

$$\mathcal{M}_{n+1}^{(u)} = \mathcal{M}_{n+1}^{(v)}.$$

Доказательство. Для $\sigma \in \mathcal{D}^{(u)}$ положим $\tilde{\sigma}(t) = \int\limits_0^t w(\tau) d\sigma(\tau)$ [1]. Тогда

$$\int\limits_0^\infty d\tilde{\sigma}(t) = \int\limits_0^A u_n(A)\, d\sigma(t) + \int\limits_A^\infty u_n(t)\, d\sigma(t) < \infty,$$

$$\int\limits_0^\infty |v_i(t)|\, d\tilde{\sigma}(t) = \int\limits_0^\infty |u_i(t)|\, d\sigma(t) < \infty, \qquad i = 0, 1, \ldots, n,$$

так что $\tilde{\sigma} \in \mathcal{D}^{(v)}$. Более того, $\int\limits_0^\infty v_i(t)\, d\tilde{\sigma}(t) = \int\limits_0^\infty u_i(t)\, d\sigma(t)$, $i=0,1,\ldots,n$, откуда следует, что $\mathcal{M}_{n+1}^{(u)} \subset \mathcal{M}_{n+1}^{(v)}$. С другой стороны, по теореме 2.1 $\mathcal{M}_{n+1}^{(v)} = \mathcal{C}(C_{n+1}^{(v)})$, где

$$C_{n+1}^{(v)} = \{\mathbf{v}(t) = (v_0(t), v_1(t), \ldots, v_n(t)) \mid 0 \leq t < \infty\}.$$

Из включения $C_{n+1}^{(v)} \subset \mathcal{M}_{n+1}^{(u)}$ следует, что $\overline{\mathcal{M}}_{n+1}^{(v)} \subset \mathcal{M}_{n+1}^{(u)}$.

Теорема 3.1 *Если $\overline{\mathcal{M}}_{n+1}^{(u)}$ обозначает замыкание $\mathcal{M}_{n+1}^{u}$, то $\overline{\mathcal{M}}_{n+1}^{(u)} = \mathcal{M}_{n+1}^{(u)} + \mathcal{R}_\infty^{(v)}$, где $\mathcal{R}_\infty^{(v)} = \{(0, 0, \ldots, 0, \lambda) \mid \lambda \geqslant 0\}$.*

Доказательство. Утверждение теоремы немедленно следует из леммы 3.1.

Замыкание выпуклого конуса $\overline{\mathcal{M}}_{n+1}$ может быть также охарактеризовано в терминах $\mathcal{P}_{n+1}$, конуса, дуального к $\mathcal{M}_{n+1}$ ($\mathcal{P}_{n+1}$ обозначает так же, как и в компактном случае, пространство коэффициентов множества всех неотрицательных многочленов, определенных на $[0, \infty)$). Из леммы 9. 1 гл. II следует, что если $\mathcal{C}$ — выпуклый конус, то $\overline{\mathcal{C}} = \mathcal{C}^{++}$, где $\mathcal{C}^{+}$ обозначает конус, дуальный к $\mathcal{C}$. В

---

[1]) $\mathcal{D}^{(u)}$ обозначает множество всех конечных мер $\sigma$, для которых $\int\limits_{[0,\infty)} |u_i|\, d\sigma < \infty$, $i = 0, 1, \ldots, n$.

этом случае из того, что $\mathcal{M}_{n+1}^{+}=\mathcal{P}_{n+1}$, вытекает соотношение $\overline{\mathcal{M}}_{n+1}=$ $=\mathcal{M}_{n+1}^{++}=\mathcal{P}_{n+1}^{+}$, так что $\overline{\mathcal{M}}_{n+1}$ является дуальным конусом к $\mathcal{P}_{n+1}$. Расписывая определение дуальности, мы получим следующую характеризацию: $\mathbf{c}\in\mathcal{M}_{n+1}$ тогда и только тогда, когда $\sum_{i=0}^{n} c_i a_i \geqslant 0$ для любого ненулевого $\mathbf{a}=(a_0, \ldots, a_n)$ из $\mathcal{P}_{n+1}$. Более того, если имеет место строгое неравенство для каждого ненулевого $\mathbf{a}$, то в действительности $\mathbf{c}\in \operatorname{Int} \overline{\mathcal{M}}_{n+1}=\operatorname{Int}\mathcal{M}_{n+1}$.

По определению $\mathbf{a}\in\mathcal{P}_{n+1}$, если $\sum_{i=0}^{n} c_i a_i \geqslant 0$ для любого $\mathbf{c}=(c_0, c_1, \ldots, c_n)\in\mathcal{M}_{n+1}$ и, следовательно, $\sum_{i=0}^{n} a_i c_i \geqslant 0$ для каждого $\mathbf{c}\in\overline{\mathcal{M}}_{n+1}$.

Мы также знаем, что $\sum_{i=0}^{n} c_i a_i > 0$ для любого $\mathbf{c}\in\overline{\mathcal{M}}_{n+1}$ $(c\neq 0)$ тогда и только тогда, когда $\mathbf{a}\in\operatorname{Int}\mathcal{P}_{n+1}$.

III. Предположим, что существует положительная непрерывная функция $w(t)$, определенная на $[0, \infty)$, такая, что

(i) $w(t)$ мажорируется на бесконечности некоторым многочленом $\sum_{i=0}^{n} a_i u_i(t)$,

(ii) $\lim_{t\to\infty} \frac{u_i(t)}{w(t)} = v_i(\infty)$ существует и конечен для $i=0, 1, \ldots, n$,

(iii) функции $v_i(t)=u_i(t)/w(t)$, $i=0, 1, \ldots, n$, удовлетворяют условию

$$V\begin{pmatrix} 0, \ldots, n-1, n \\ t_0, \ldots, t_{n-1}, \infty \end{pmatrix} > 0$$

при любом выборе $0\leq t_0 < \ldots < t_{n-1} < \infty$.

(Условие (i) является техническим требованием и накладывается для того, чтобы $w(t)$ была интегрируемой относительно любой меры $\sigma$, для которой $\int u_i(t)\, d\sigma(t)$, $i=0, 1, \ldots, n$, определен.)

Условия, налагаемые в определениях $T$-систем типов I и II, в действительности являются частными случаями условий из определения $T$-системы типа III. Для $T$-системы типа I мы можем положить $w(t)\equiv 1$, так как в начале этого раздела мы предположили, что существует многочлен $u(t)$, для которого $\lim_{t\to\infty} \inf u(t) \geqslant 1$. Для $T$-системы типа II мы можем выбрать $w(t)$, как в (3.2). Мы считаем полезным выделить первые два типа, так как эти частные случаи часто встречаются в приложениях.

Продолжая наше обсуждение $T$-систем типа III, мы заметим, что система $\{v_i\}_{i=0}^n$ имеет тип I. Поэтому для системы $\{v_i\}$

$$\overline{\mathcal{M}}_{n+1}^{(v)} = \mathcal{M}_{n+1}^{(v)} + \mathcal{R}_\infty^{(v)} = \mathcal{C}\,(\overline{C}_{n+1}^{(v)}), \tag{3.5}$$

где $\mathcal{R}_\infty^{(v)}$ — положительный луч, порожденный вектором $(v_0(\infty), v_1(\infty), \ldots, v_n(\infty))$.

Далее можно показать, следуя доказательству леммы 3.1, что $\mathcal{M}_{n+1}^{(u)} = \mathcal{M}_{n+1}^{(v)}$, так что из (3.5) следует, что $\overline{\mathcal{M}}_{n+1}^{(u)} = \mathcal{M}_{n+1}^{(u)} + \mathcal{R}_\infty^{(v)}$. В этом месте полезно подытожить все существенное в формулах представления для $\mathcal{M}_{n+1}$ и $\overline{\mathcal{M}}_{n+1}$ в случае $T$-систем типов I, II и III.

Тип I.

(i) $\mathbf{c} = (c_0, c_1, \ldots, c_n) \in \mathcal{M}_{n+1}$ тогда и только тогда, когда

$$c_i = \int_0^\infty u_i(t)\, d\sigma(t), \qquad i = 0, 1, \ldots, n, \tag{3.6}$$

и $\sigma$ — неотрицательная конечная мера на $[0, \infty)$.

(ii) $\mathbf{c} = (c_0, \ldots, c_n) \in \overline{\mathcal{M}}_{n+1}$ тогда и только тогда, когда

$$c_i = \int_0^\infty u_i(t)\, d\sigma(t) + \lambda u_i(\infty), \qquad i = 0, 1, \ldots, n \tag{3.7}$$

($\lambda \geqslant 0$ и $\sigma$ — неотрицательная конечная мера на $[0, \infty)$).

Тип II.

(i) $\mathbf{c} \in \mathcal{M}_{n+1}$ тогда и только тогда, когда справедливо (3.6).

(ii) $\mathbf{c} \in \overline{\mathcal{M}}_{n+1}$ тогда и только тогда, когда

$$c_i = \int_0^\infty u_i(t)\, d\sigma(t), \qquad i = 0, 1, \ldots, n-1,$$

$$c_n = \int_0^\infty u_n(t)\, d\sigma(t) + \lambda$$

($\lambda \geqslant 0$ и $\sigma$ — неотрицательная конечная мера на $[0, \infty)$).

Тип III.

(i) $\mathbf{c} \in \mathcal{M}_{n+1}$ тогда и только тогда, когда справедливо (3.6).

(*ii*) $\mathbf{c} \in \overline{\mathcal{M}}_{n+1}$ тогда и только тогда, когда

$$c_i = \int_0^\infty u_i(t)\, d\sigma(t) + \lambda v_i(\infty), \qquad i = 0, 1, \ldots, n$$

($\lambda \geqslant 0$ и $\sigma$ — неотрицательная конечная мера на $[0, \infty)$).

Описанная в этом разделе техника вывода, основанная на компактификации луча и подходящем преобразовании $T$-системы, определенной на $[0, \infty)$, в соответствующую $T$-систему, определенную на конечном замкнутом интервале, была использована впервые у Ахиезера и Крейна [1938] и более широко у Крейна [1951]. Некоторые варианты таких технических приемов в полиномиальных случаях также были использованы у Шохата и Тамаркина [1943], гл. I.

## § 4. Канонические представления, максимальная масса $\rho(t_0)$ и двумерные сечения

Пусть $\{u_i\}_0^n$ — $T$-система типа II. Это предполагает, в частности, что $\{u_i\}_{i=0}^n$ и $\{u_i\}_{i=0}^{n-1}$ являются $T$-системами на $[0, \infty)$. Предположение о том, что $\{u_i\}_0^{n-1}$ является $T$-системой, требуется для того, чтобы функции

$$v_i(t) = \begin{cases} \dfrac{u_i(t)}{w(t)}, & 0 \leq t < \infty, \\ \delta_{in}, & t = \infty \end{cases} \qquad (i = 0, 1, \ldots, n)$$

образовывали $T$-систему на замкнутом интервале $[0, \infty]$, где $w(t)$ предполагается строго положительной, непрерывной на $[0, \infty)$ и совпадающей с $u_n(t)$ для больших $t$. Однажды преобразование перехода от $u_i$ к $v_i$ уже было произведено, так что совершенно ясно, как интерпретировать для системы $\{u_i\}_{i=0}^n$ различные понятия, такие, как канонические и главные представления, максимальная масса и т. д., определенные для систем $\{v_i\}_{i=0}^n$. При помощи этих преобразований мы получаем следующие результаты.

Для каждого $c \in \operatorname{Int} \mathcal{M}_{n+1}$ существуют два «главных» представления индекса $(n+1)/2$. (При вычислении индекса точки $t = \infty$ и $t = 0$ берутся с весом $1/2$, а все остальные — с весом единица.) При верхнем главном представлении имеется масса на $\infty$, с другой стороны, нижнее главное представление дает нам настоящую конечную меру, которая для $n = 2m$ имеет массу в точках $\{t_i^*\}_{i=1}^{m+1}$, удовлетворяющих неравенствам $0 = t_1^* < t_2^* < \ldots < t_{m+1}^* < \infty$, а для $n = 2m+1$ имеет массу в $m+1$ точках, удовлетворяющих неравенствам $0 < t_1^* < t_2^* < \ldots < t_{m+1}^* < \infty$. Существует «каноническая мера», соответствующая $\xi_0$ $(0 < \xi_0 < \infty)$, индекса, не большего чем $(n+2)/2$. Когда имеется масса на бесконечности, на нее можно не обращать внимания и, таким образом, получить меру, представляющую только моменты $c_0, c_1, \ldots, c_{n-1}$ относительно системы $u_0, u_1, \ldots, u_{n-1}$. Каждое из упомянутых представлений можно вывести, если определить соответствующее представление для системы $\{v_i\}_{i=0}^n$ и затем выразить $v_i$ в терминах $u_i$. Например, если $c$ принадлежит $\operatorname{Int} \mathcal{M}_{n+1} = \operatorname{Int} \mathcal{M}_{n+1}^{(u)}$, то по лемме 3.1 $c \in \operatorname{Int} \mathcal{M}_{n+1}^{(v)}$, поэтому для каждого $\xi \in (0, \infty)$ мы имеем представление

$$c_i = \sum_{j=1}^{p} \alpha_j v_i(t_j), \qquad i = 0, 1, \ldots, n, \tag{4.1}$$

где $0 \leq t_1 < t_2 < \ldots < t_{p-1} < t_p \leq \infty$, $p \leq [(n+2)/2] + 1$ и $\xi \in \in \{t_1, \ldots, t_p\}$.

Если $t_p < \infty$, то

$$c_i = \sum_{j=1}^{p} \lambda_j u_i(t_j), \qquad i = 0, 1, \ldots, n,$$

где $\lambda_j = \alpha_j / w(t_j)$, $j = 1, \ldots, p$.

Если $t_p = \infty$, то (так как $v_i(\infty) = 0$ для $i = 0, 1, \ldots, n-1$) уравнение (4.1) переходит в

$$c_i = \sum_{j=1}^{p-1} \lambda_j u_i(t_j), \qquad i = 0, 1, \ldots, n-1$$

$$c_n = \sum_{j=1}^{p-1} \lambda_j u_n(\lambda_j) + \alpha_p \geq \sum_{j=1}^{p-1} \lambda_j u_n(t_j).$$

Таким же образом читатель может быстро вывести и другие результаты о представлениях.

Масса в точке $t_0$ канонического представления для $c \in \operatorname{Int} \mathcal{M}_{n+1}$ может быть связана с экстремальной задачей, как и в случае компактного интервала (см. § 4 гл. II и § 3 гл. IV).

Для фиксированной меры $\sigma$, моментная точка которой $(c_0, c_1, \ldots \ldots, c_n)$ принадлежит $\operatorname{Int} \mathcal{M}_{n+1}$, рассмотрим величину

$$\rho(t_0) = \min \int_0^{\infty} u(t)\, d\sigma(t), \tag{4.2}$$

где минимум берется по всем неотрицательным многочленам $u \in \mathcal{P}_{n+1}$ таким, что $u(t_0) = 1$.

Равенство (4.2) может быть переписано в виде

$$\rho(t_0) = \min_u \int_0^{\infty} \frac{u(t)}{w(t)} w(t)\, d\sigma(t) = \min_v \int_0^{\infty} v(t)\, d\mu,$$

где $\mu$ $(d\mu(t) = w(t)\, d\sigma(t))$ имеет моменты $c_0, c_1, \ldots, c_n$ по отношению к системе $\{v_i\}_{i=0}^{n}$, определенной в (3.3) и минимум берется по всем неотрицательным $v$-многочленам, удовлетворяющим условию $v(t_0) = w^{-1}(t_0)$.

Методом, подобным методу § 4 гл. II, мы получаем

$$\rho(t_0) = \frac{\rho_v(t_0)}{w(t_0)},$$

где $\rho_v(t_0)$ — максимальная масса канонического представления $\mu_{t_0}$ (включающего $t_0$) точки $(c_0, c_1, \ldots, c_n)$ относительно системы $\{v_i\}_{i=0}^{n}$.

Минимизирующий многочлен $v$ обращается в нуль во всех корнях этого представления, кроме $t_0$.

В терминах системы $\{v_i\}_0^n$ точка с имеет два главных представления. Мы обозначим корни верхнего и нижнего представлений через $\{s_i^*\}$ и $\{t_i^*\}$ соответственно. Заметим, что наибольшая точка в верхнем представлении всегда $+\infty$. Эти корни разбивают $[0, \infty)$ на непересекающиеся интервалы:

$$[s_1^*, t_2^*] \qquad [s_2^*, t_3^*] \ldots [s_m^*, t_{m+1}^*], \tag{A}$$

$$[t_1^*, s_1^*] \qquad [t_2^*, s_2^*] \ldots [t_{m+1}^*, \infty), \tag{B}$$

если $n = 2m$, где $t_1^* = 0$ и

$$[s_1^*, t_1^*] \quad [s_2^*, t_2^*] \ldots [s_{m+1}^*, t_{m+1}^*] \tag{A'}$$

$$[t_1^*, s_2^*] \quad [t_2^*, s_3^*] \ldots [t_{m+1}^*, \infty), \tag{B'}$$

если $n = 2m + 1$, где теперь $s_1^* = 0$.

Внутренние интервалы (A) и (A′) соответствуют верхним каноническим интервалам $K_i$, а внутренние интервалы (B) и (B′)—нижним каноническим интервалам $J_i$ (см. §§ 5 и 6 гл. II). Пусть $\overline{K}$ обозначает объединение интервалов (A) (или (A′), если $n$ нечетно) и $\overline{J}$ обозначает объединение интервалов (B) (или (B′), если $n$ нечетно). Заметим, что точка $+\infty$, соответствующая концевой точке $b$, исключается из $\overline{J}$.

Если минимизирующий многочлен $v$ должен обращаться в нуль при $t = \infty$, то коэффициент при $v_n$ должен быть нулем, так как $v_k(\infty) = \delta_{k,n}$ для $k = 0, 1, \ldots, n$. В случае $u_i(t) = t^i, i = 0, 1, \ldots, n$, это означает, что многочлен $u(t) = w(t)\, v(t)$, на котором достигается минимум в (4.2), имеет степень, не большую чем $n - 1$. Тогда вид многочлена $u$ может быть легко установлен; например, если $n = 2m$ и $t_0$ является внутренней точкой одного из интервалов (A), то $u(t)$ имеет вид $\alpha t \prod_{j=1}^{m-1} (t - t_j)^2$, где $\alpha$ — постоянная. Во всех случаях величина $\rho(t_0)$ может быть вычислена при помощи одной из систем $\{\underline{P}_k(t)\}$ и $\{\underline{Q}_k(t)\}$, которые ортонормированы по отношению к $d\sigma(t)$ и $t\, d\sigma(t)$ соответственно. Это приводит к следующей теореме.

Теорема 4.1. *Величина* $\rho(t_0$ *для* $T$*-системы* $\{t^i\}_{i=0}^n$ *на интервале* $[0, \infty)$ *определяется равенством*:

(i) *если* $n = 2m$,

$$\rho^{-1}(t_0) = \begin{cases} t_0 \sum\limits_{k=0}^{m-1} \underline{Q}_k^2(t_0), & t_0 \in \overline{K}, \\ \sum\limits_{k=0}^{m} \underline{P}_k^2(t_0), & t_0 \in \overline{J}, \end{cases}$$

(ii) *если* $n = 2m + 1$,

$$\rho^{-1}(t_0) = \begin{cases} \sum_{k=0}^{m} P_{-k}^2(t_0), & t_0 \in \overline{K}, \\ t_0 \sum_{k=0}^{m} Q_{-k}^2(t_0), & t_0 \in \overline{J}. \end{cases}$$

Явная формула для максимальной массы $\rho(t_0)$, приведенная в этой теореме, была первоначально найдена Шенбергом и Сеге [1960]. Исторические комментарии на эту тему можно найти у Сеге [1959].

Мы теперь перейдем к краткому обсуждению двумерных сечений и распространим на наш случай результаты § 7 гл. II. Мы предполагаем, что как $\{u_i\}_{i=0}^{n}$, так и $\{u_i\}_{i=0}^{n+2}$ являются $T$-системами типа II, т. е. системы $\{u_i\}_{i=0}^{k}$ для $k = n-1,\ n,\ n+1,\ n+2$ являются $T$-системами, $u_n(t)$ и $u_{n+2}(t)$ — строго положительны для больших $t$ и

$$\lim_{t\to\infty} \frac{u_i(t)}{u_n(t)} = 0, \qquad 0 \leq i < n,$$

$$\lim_{t\to\infty} \frac{u_i(t)}{u_{n+2}(t)} = 0, \qquad 0 \leq i < n+2.$$

Точка $\mathbf{c} = (c_0, \ldots, c_n) \in \operatorname{Int} \mathcal{M}_{n+1}^{(u)}$ также является внутренней для $\mathcal{M}_{n+1}^{(\bar{v})}$ (см. § 3), порожденного функциями $\bar{v}_i(x)$, $0 \leq i \leq n$, определенными на интервале $[0, \pi/2]$ равенством

$$\bar{v}_i(x) = \begin{cases} \dfrac{u_i(\operatorname{tg} x)}{w(\operatorname{tg} x)}, & 0 \leq x < \dfrac{\pi}{2}, \\ \delta_{in}, & x = \dfrac{\pi}{2}, \end{cases} \qquad i = 0, 1, \ldots, n.$$

Здесь $w(t) > 0$ для $t \in [0, \infty)$ и $w(t) = u_n(t)$ для $t$ достаточно больших. Поэтому точка $\mathbf{c}$ обладает нижним главным представлением по отношению к $T$-системе $\{v_i\}_{i=0}^{n}$ на $[0, \pi/2]$. Пусть $(x_1^0, x_2^0)$ — интервал между двумя соседними корнями этого представления, и рассмотрим точку $x' \in (x_1^0, x_2^0)$. Если $\rho_{\bar{v}}(x')$ обозначает максимальную массу в $x'$ и $0 < \alpha < 1$, то

$$(c_0^*, \ldots, c_n^*) = (c_0, \ldots, c_n) - (1-\alpha)\,\rho_{\bar{v}}(x')\,(\bar{v}_0(x'), \ldots, \bar{v}_n(x'))$$

содержится в $\operatorname{Int} \mathcal{M}_{n+1}^{(\bar{v})}$ и, следовательно, $\mathbf{c}^*$ обладает нижним главным представлением. Тепер› ясно, что для каждого $t' \in (s_1^0, s_2^0)$, где $s_i^0 = \operatorname{tg} x_i^0$, $i = 1, 2$, существует представление $\sigma(t, t', \lambda)$ точки $(c_0, \ldots, c_n)$ индекса $(n+3)/2$ с произвольным весом $\lambda \in (0, \rho(t'))$, где $\rho(t') = \rho_{\bar{v}}(x')\,w^{-1}(x')$. Как и в § 7 гл. II, мы положим

$$A = \left\{(c_{n+1}, c_{n+2}) \mid c_i = \int_0^\infty u_i d\sigma, i = n+1, n+2, \sigma \in V(\mathrm{c})\right\}$$

и рассмотрим отображение области

$$B = \{(t', \lambda) \mid t' \in (s_1^0, s_2^0),\ \lambda \in (0, \rho(t'))\}$$

в $A$, определенное равенством $\varphi(t', \lambda) = (c_{n+1}, c_{n+2})$, где $c_i = \int u_i(t) \times \times d\sigma(t, t', \lambda)$, $i = n+1$, $n+2$. Отображение $\varphi$ взаимно однозначно, непрерывно и отображает $B$ на $\operatorname{Int} A$ так же, как и в случае компактного интервала. Для системы $\{\bar{v}_i\}_{i=0}^{n+2}$ проходят рассуждения, которые проводились в § 7 гл. II.

## § 5. Теорема Маркова — Крейна

Пусть $\{u_i\}_0^n$ и $\{u_i\}_0^{n+1}$ — $T$-системы на $[0, \infty)$; рассмотрим

$$I_{\min} = \inf_{\sigma \in V(\mathrm{c})} \int_0^\infty u_{n+1}(t)\, d\sigma(t),$$

где $\mathrm{c} \in \operatorname{Int} \mathcal{M}_{n+1}$ и $V(\mathrm{c})$ — множество всех конечных мер $\sigma$, определенных на $[0, \infty)$, таких, что $c_i = \int_0^\infty u_i\, d\sigma$, $i = 0, 1, \ldots, n$, и $\int_0^\infty |u_i|\, d\sigma < \infty$, $i = 0, 1, \ldots, n+1$.

Ясно, что для достаточно больших $b$ (скажем, $b > b_0$) $\mathrm{c} \in \operatorname{Int} \mathcal{M}_{n+1}(b)$, где $\mathcal{M}_{n+1}(b)$ — моментное пространство, порожденное системой $\{u_i\}_{i=0}^n$ на $[0, b]$. Поэтому моментная точка $\mathrm{c}$ допускает нижнее главное представление $\underline{\sigma}$ индекса $(n+1)/2$ относительно интервала $[0, b]$, и это представление не должно зависеть от $b$ $(b > b_0)$.

Рассмотрим произвольное $\sigma \in V(\mathrm{c})$. Так как моментная точка $\left(c_0, c_1, \ldots, c_n, \int_0^\infty u_{n+1} d\sigma\right)$ принадлежит $\mathcal{M}_{n+2}$, то из теоремы 2.1 следует, что существует мера $\tilde{\sigma} \in V(\mathrm{c})$ с конечным спектром, удовлетворяющая $\int_0^\infty u_{n+1} d\sigma = \int_0^\infty u_{n+1} d\tilde{\sigma}$. Очевидно, что $\tilde{\sigma}$ концентрирует свою массу на конечном интервале $[0, b]$ для достаточно большого $b$. Применяя теорему 1.1 гл. III, мы заключаем, что $\int_0^\infty u_{n+1} d\tilde{\sigma} \geqslant$

$\geqslant \int\limits_0^\infty u_{n+1} d\underline{\sigma}$. Из этого следует, что $I_{\min} = \int\limits_0^\infty u_{n+1} d\underline{\sigma}$. Далее, так как точка $(c_0, c_1, \ldots, c_n, I_{\min})$ — граничная точка $\mathcal{M}_{n+2}$, то мера $\sigma \in V(\mathbf{c})$, для которой $\int\limits_0^\infty u_{n+1} d\sigma = I_{\min}$ единственна. Таким образом, мы доказали следующую теорему.

Теорема 5.1. *Если $\{u_i\}_0^n$ и $\{u_i\}_0^{n+1}$ — $T$-системы на $[0, \infty)$ и $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$, то наименьшее значение*

$$I_{\min} = \min_{\sigma \in V(\mathbf{c})} \int\limits_0^\infty u_{n+1}(t)\, d\sigma(t) \tag{5.1}$$

*достигается только для $\sigma = \underline{\sigma}$ — нижнего главного представления* $\mathbf{c}$.

Верхняя граница интеграла в (5.1) может не достигаться и, вообще говоря, может быть даже бесконечной. Однако если предполагать, что $\{u_i\}_0^n$ является системой типа II и

$$\lim_{t \to \infty} \frac{u_{n+1}(t)}{u_n(t)} = \gamma \qquad (\gamma \text{ — конечно}),$$

то значение $I_{\max}$ существует. В действительности,

$$I_{\max} = \sum_{j=1}^{q-1} \alpha_j^* u_{n+1}(s_j^*) + \gamma \rho_q,$$

где $\alpha_j^* = \rho_j / w(s_j^*)$, $j = 1, 2, \ldots, q-1$ и $\{\rho_j\}_1^q$ — веса верхнего главного представления $\mathbf{c}$, выраженные из системы $\{v_i\}_0^n$, определенной в (3.3).

Для того чтобы доказать аналог теоремы Маркова — Крейна (§ 2 гл. III), мы предположим, что система $\{u_i\}_0^n$ и $\Omega = u_{n+1}$ удовлетворяют следующим условиям:

(i) $\{u_i\}_0^n$ — система типа II,

(ii) $u_0, u_1, \ldots, u_k, \Omega$ и $u_0, \ldots, u_k$ — $T$-системы на $[0, \infty)$, для каждого $k = 0, 1, \ldots, n$ и $\Omega(t) > 0$ для $t \geqslant 0$.

Пусть $\sigma$ — фиксированная мера, представляющая точку $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$. Рассмотрим моментную точку $\mathbf{c}(b) = (c_0(b), c_1(b), \ldots, c_n(b))$, где $c_i(b) = \int\limits_0^b u_i d\sigma$, $i = 0, 1, \ldots, n$. Мы утверждаем, что если $\mathcal{M}_{n+1}(b)$ — моментное пространство, порожденное системой $\{u_i\}_0^n$ на интервале $[0, b]$, то $\mathbf{c}(b) \in \operatorname{Int} \mathcal{M}_{n+1}(b)$ для достаточно больших $b$. Действительно, мера $\sigma$ на $[0, \infty)$ имеет индекс по крайней мере $(n+1)/2$ и, следовательно, ее индекс относительно $[0, b)$ по крайней мере $(n+1)/2$ для достаточно большого $b$. Поэтому $\mathbf{c}(b) \in \operatorname{Int} \mathcal{M}_{n+1}(b)$.

В силу предположения (ii) можно применить теорему 2.1 гл. III. Тогда получим

$$\int_0^{\xi-} \Omega(t)\, d\sigma_\xi(t, b) \leq \int_0^{\xi-} \Omega(t)\, d\sigma(t) \leq \int_0^{\xi\pm} \Omega(t)\, d\sigma(t) \leq$$
$$\leq \int_0^{\xi+} \Omega(t)\, d\sigma_\xi(t; b), \qquad (0 < \xi < b), \qquad (5.2)$$

где $\sigma_\xi(t, b)$ — каноническая мера для $c(b)$, включающая $\xi$ относительно системы $\{u_i\}_0^n$ на интервале $[0, b]$. Чаще всего в неравенстве (5.2) имеет место строгое неравенство.

Для системы $\{v_i\}_0^n$, определенной в (3.3), мера $w(t)\, d\sigma_\xi(t; b)$ имеет индекс, не превосходящий $(n+2)/2$, и представляет $c(b)$ относительно системы $\{v_i\}_0^n$, определенной на $[0, \infty)$. Так как $\{v_i\}_0^n$ является $T$-системой на замкнутом интервале $[0, \infty]$, то существует строго положительный многочлен $v = \sum_{i=0}^{n} a_i v_i$ на $[0, \infty]$. При таком выборе $v$

$$\{\min_{\tau\in[0,\infty]} v(\tau)\} \int_0^{\infty} w(t)\, d\sigma_\xi(t, b) \leq \int_0^{\infty} v(t)\, w(t)\, d\sigma_\xi(t, b) = \sum_{i=0}^{n} a_i c_i(b).$$

Правая часть, очевидно, ограничена по $b$, так что меры $w(t) d\sigma_\xi(t; b)$ имеют равномерно ограниченную вариацию на $[0, \infty]$. Тогда существует мера $\rho_\xi$ на $[0, \infty]$ такая, что

$$c_i = \lim_{b\to\infty} \int_{[0,\infty]} v_i(t)\, w(t)\, d\sigma_\xi(t; b) = \int_{[0,\infty]} v_i(t)\, d\rho_\xi(t), \quad i = 0, 1, \ldots, n.$$

Мера $\rho_\xi$ имеет индекс не более чем $(n+2)/2$, возможно, с массой на бесконечности. Она представляет $c = (c_0, c_1, \ldots, c_n)$ относительно системы $\{v_i\}_0^n$ на интервале $[0, \infty]$. Кроме того, $\rho_\xi$ имеет массу в $\xi$, так как в противном случае $\rho_\xi$ имело бы индекс не более чем $n/2$, откуда следовало бы, что $c$ является граничной точкой $\mathcal{M}_{n+1}$. Поэтому $\rho_\xi$ совпадает с каноническим представлением $c$, включающим $\xi$, относительно системы $\{v_i\}_0^n$ на интервале $[0, \infty]$. Далее точки, являющиеся носителями меры $w(t)\, d\sigma_\xi(t; b)$ (исключая массу в $\xi$) не могут сходиться к $\xi$, так что

$$\lim_{b\to\infty} \int_0^{\xi+} \Omega(t)\, d\sigma_\xi(t; b) = \int_0^{\xi+} \frac{\Omega(t)}{w(t)}\, d\rho_\xi(t),$$

$$\lim_{b\to\infty} \int_0^{\xi-} \Omega(t)\, d\sigma_\xi(t; b) = \int_0^{\xi-} \frac{\Omega(t)}{w(t)}\, d\rho_\xi(t).$$

Проведенный анализ приводит нас к теореме.

Теорема 5.2. *Пусть* $u_0, u_1, \ldots, u_n$, $\Omega$ *удовлетворяют условиям* (i) *и* (ii), *и пусть* $\mathbf{c} = (c_0, c_1, \ldots, c_n) \in \operatorname{Int} \mathcal{M}_{n+1}$.

*Тогда, если* $\sigma \in V(\mathbf{c})$ *и* $0 < \xi < \infty$, *то*

$$\int_0^{\xi-} \Omega(t)\, d\sigma_\xi(t) \leq \int_0^{\xi-} \Omega(t)\, d\sigma(t) \leq \int_0^{\xi+} \Omega(t)\, d\sigma(t) \leq \int_0^{\xi+} \Omega(t)\, d\sigma_\xi(t),$$

*где* $d\sigma_\xi(t) = w^{-1}(t)\, d\rho_\xi(t)$ *и* $\rho_\xi$ — *каноническая мера точки* $\mathbf{c}$, *включающая* $\xi$, *относительно системы* $\{v_i\}_0^n$ *на интервале* $[0, \infty]$ ($v_i$ *явно определяются в* (3.3)).

Рассуждения этого параграфа главным образом следуют рассуждениям Крейна [1934].

## § 6. Другой класс экстремальных задач, решениями которых являются канонические меры

В §4 мы отметили, что если для $T$-системы $\{u_i\}_0^n$ типа II рассмотреть соответствующую систему $\{v_i\}_0^n$, определяемую (3.3), то мы могли бы перенести большую часть результатов гл. II и III со случая компактного интервала $[a, b]$ на полупрямую $[0, \infty)$. Для того чтобы сделать некоторые из этих утверждений (например, теорему Маркова — Крейна) строгими, требуется дополнительное ограничение. Это также относится к перенесению результатов § 5 гл. III, в котором мы исследовали экстремальные значения определенных интегралов в том случае, когда множество мер, по которым берется экстремум, состоит из мер, соответствующих некоторому подмножеству $\mathcal{M}_{n+1}$, а не единственной точке $\mathbf{c}$ из $\mathcal{M}_{n+1}$, как в теореме Маркова — Крейна. Главным результатом этого параграфа является теорема 6.2. Мы применим эту теорему в § 7 для исследования аппроксимации преобразований Лапласа — Стилтьеса экспоненциальными многочленами (см. теорему 7.1). Доказательства, приведенные ниже, подобны доказательствам § 5 гл. III.

Пусть $u_0, \ldots, u_n, u_{n+1}$ — функции, определенные на $[0, \infty)$, удовлетворяющие следующим условиям.

Условие 6.1.

(a) $\{u_i\}_0^n$ и $\{u_i\}_0^{n+1}$ — $T$-системы,

(b) для $i = 0, 1, \ldots, n+1$ определители порядка $n+1$, образованные из $u_0, u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_{n+1}$ все имеют один знак $\varepsilon_i$ ($\varepsilon_i = \pm 1$) (см. условие 5.1 гл. III),

(c) $u_{n+1}(t) > 0, \quad t \geqslant A > 0$,

(d) $\lim\limits_{t \to \infty} \dfrac{u_i(t)}{u_{n+1}(t)} = 0, \quad i = 0, 1, \ldots, n$.

Ясно, что из этих предположений следует, что $\{u_i\}_0^{n+1}$ $T$-система типа II.

Ниже мы ограничим наше внимание мерами $\sigma$, сконцентрированными на интервале $[0, \infty)$ и такими, что $\int u_{n+1} d\sigma < \infty$. Если не оговорено противное, то все встречающиеся ниже интегралы рассматриваются по области $[0, \infty)$.

Теорема 6.1. *Пусть $A$, $B$ и $C$ — разбиение множества чисел $0, 1, \ldots, n$. Если $\{u_i\}_0^{n+1}$ удовлетворяет условию 6.1 и класс мер, удовлетворяющих условиям*

$$\int u_i d\sigma \begin{cases} = d_i, & i \in A, \\ \leq d_i, & i \in B, \\ \geq d_i, & i \in C, \end{cases}$$

*не пуст, то*

$$\underline{\alpha} = \min_{\sigma \in V} \int u_{n+1} d\sigma$$

*достигается для единственной меры $\underline{\sigma}$ из $V$ индекса, не большего чем $(n+1)/2$.*

Доказательство. Пусть $v_i(t) = u_i(t)/w(t)$, $0 \leq i \leq n+1$, где $w$ — строго положительная непрерывная функция на $[0, \infty)$ такая, что $w(t) = u_{n+1}(t)$ для $t > A$. Тогда

$$\underline{\alpha} = \min_{\mu \in U} \int v_{n+1} d\mu,$$

где $U$ — класс мер $\mu$, для которых $\int d\mu < \infty$ и

$$\int v_i d\mu \begin{cases} = d_i, & i \in A, \\ \leq d_i, & i \in B, \\ \geq d_i, & i \in C. \end{cases}$$

При помощи стандартных рассуждений получаем, что $\underline{\alpha}$ достигается на некоторой мере $\underline{\mu} \in U$. Положим теперь $c_i^0 = \int v_i d\underline{\mu}$, $i = 0, 1, \ldots, n$. Ясно, что точка $(c_0^0, \ldots, c_n^0, \underline{\alpha})$ — граничная точка моментного пространства, порожденного функциями $v_0, v_1, \ldots, v_{n+1}$ на $[0, \infty)$. Следовательно, мера $\underline{\mu}$ имеет индекс, не превосходящий $(n+1)/2$ на $[0, \infty)$. Чтобы показать, что мера $d\underline{\sigma}(t) = w^{-1}(t)\, d\underline{\mu}(t)$ удовлетворяет всем требованиям теоремы, остается доказать единственность.

Доказательство единственности для $\underline{\sigma}$ основывается на следующей лемме, которая потребуется в дальнейшем. Определим для каждой моментной точки $\mathbf{c} \in \mathcal{M}_{n+1}^{(u)}$ величину

$$\underline{c}_{n+1}(\mathbf{c}) = \inf_{\sigma \in V(\mathbf{c})} \int u_{n+1} d\sigma.$$

Лемма 6.1. *Если $u_0, \ldots, u_{n+1}$ удовлетворяет условиям 6.1 (a), (c), (d) и условию 6.1 (b) для фиксированного $i$, то $(-1)^{n-i} \times \varepsilon_i \underline{c}_{n+1}(\mathbf{c})$ — строго возрастает по $c_i$.*

Доказательство получается, если повторить дословно рассуждения, проводимые при доказательстве леммы 5.1 гл. III и поэтому опускается.

Приступим к доказательству единственности в теореме 6.1. Предположим, что $\sigma$ ($\sigma \neq \underline{\sigma}$) удовлетворяет условиям теоремы и на $\sigma$ достигается минимум $\underline{\alpha}$.

Для описанного множества моментов относительно $u_0, \ldots, u_n$ величина $\underline{\alpha}$ достигается на единственной мере индекса, не большего чем $(n+1)/2$, так что $\int u_j d\sigma \neq \int u_j d\underline{\sigma}$ для некоторого $j$ $(0 \leq j \leq n)$. Мера $\sigma' = (\sigma + \underline{\sigma})/2$ (которая, очевидно, удовлетворяет условиям и на которой достигается минимум $\underline{\alpha}$) такова, что $c'_j = \int u_j d\sigma'$ лежит между $\int u_j d\sigma$ и $\int u_j d\underline{\sigma}$, в то время как другие моменты остаются неизменными.

Но тогда из леммы 6.1 следует, что мы можем уменьшить $\underline{\alpha}$, что противоречит определению.

Следствие 6.1a. *Сохраняя обозначения теоремы* 6.1, *определим* $\underline{D} = \{i \mid \int u_i d\underline{\sigma} = d_i,\ 0 \leq i \leq n\}$. *Тогда* $\underline{\sigma}$ *является единственной мерой в* $V_1$, *на которой достигается*

$$\underline{\beta} = \min_{\sigma \in V_1} \int u_{n+1} d\sigma,$$

*где* $V_1$ — *класс мер* $\sigma$, *удовлетворяющих условию*

$$\int u_i d\sigma \begin{cases} = d_i, & i \in A \cap \underline{D}, \\ \leq d_i, & i \in B \cap \underline{D}, \\ \geq d_i, & i \in C \cap \underline{D}. \end{cases}$$

Доказательство. См. следствие 5.1a гл. III.

Пусть $v_0, \ldots, v_n$ — другое множество функций, определенных на $[0, \infty)$, таких, что для $v_0, \ldots, v_n, u_{n+1}$ выполнено условие 6.1. Применяя рассуждения следствия 5.1b гл. III, мы получим

Следствие 6.1b. *Если* $\underline{\sigma}$, $A, B, C$ *и* $\underline{D}$ — *те же, что и в следствии* 6.1a *и* $e_i = \int v_i d\underline{\sigma}$, $0 \leq i \leq n$, *то* $\underline{\sigma}$ *является единственной мерой в* $W_1$, *на которой достигается*

$$\underline{\gamma} = \min_{\sigma \in W_1} \int u_{n+1} d\sigma,$$

*где* $W_1$ *состоит из всех* $\sigma$, *удовлетворяющих условию*

$$\int v_i d\sigma \begin{cases} = e_i, & i \in A \cap \underline{D}, \\ \leq e_i, & i \in B \cap \underline{D}, \\ \geq e_i, & i \in C \cap \underline{D}. \end{cases}$$

Доказательство опускаем.

Теперь предположим, что мы можем пополнить систему $\{u_i\}_0^{n+1}$ одной функцией $v_i$ для любого фиксированного $i = 0, 1, \ldots, n+1$ так, что для получившегося множества функций справедливо

Условие 6.2.

(a) $\{u_0, \ldots, u_{i-1}, v_i, u_i, \ldots, u_n\}$ и $\{u_0, \ldots, u_{i-1}, v_i, u_i, \ldots, u_{n+1}\}$, являются $T$-системами;

(b) определители $(n+1)$-го порядка, образованные из любого упорядоченного множества из $n+1$ функций, выбранных из $\{u_0, \ldots$ $\ldots, u_{i-1}, v_i, u_i, \ldots, u_n\}$, имеют все строго один знак (см. условие 5.1 (b) гл. III),

(c) $\lim\limits_{t\to\infty} (v_i(t))/(u_{n+1}(t)) = 0$. Следуя рассуждениям теоремы 5.2 гл. III, мы получаем теорему.

Теорема 6.2. *Если функции* $\{u_0, \ldots, u_{i-1}, v_i, u_i, \ldots, u_{n+1}\}$ *удовлетворяют условию* 6.2 *и условиям* 6.1 (a), (c), (d) и $V$ *и* $\underline{\sigma}$ *определены в теореме* 6.1, *то*

$$\min_{\sigma\in V} (-1)^{n-i-1} \int v_i d\sigma = \underline{\delta}_i$$

*достигается для единственной меры* $\underline{\sigma}$.

## § 7. Применения теории канонических мер к аппроксимации преобразований Лапласа и Стилтьеса

В этом параграфе мы применяем развитую выше теорию к некоторым важным примерам интерполяции и аппроксимации. Основновное содержание этих примеров состоит в следующем. Рассмотрим функцию $K(x, t)$ двух переменных, определенную на интервалах $X$ и $T$ действительной прямой и рассмотрим преобразование

$$f(x) = \int_T K(x, t)\, d\mu(t), \tag{7.1}$$

которое переводит меры $\mu$ в функции $f(x)$.

В некоторых случаях мы можем сопоставить $f(x)$ конечную сумму $\sum\limits_j K(x, t_j)\lambda_j$, которая интерполирует $f(x)$ в выбранных точках и аппроксимирует функцию $f(x)$ во всем $X$ нужным образом. Общая процедура работает так. Если ядро $K(x, t)$ везде строго положительно (см. определение 2.2 гл. I), то величины $f(x_0), f(x_1), \ldots$ $\ldots, f(x_n)$ могут рассматриваться как моментные точки, порожденные $T$-системой $u_i(t) = K(x_i, t)$, $i = 0, 1, \ldots, n$. Мы можем теперь заменить $\mu$ на каноническую или главную меру, сосредоточенную на конечном числе точек и порождающую те же самые моменты, т. е.

$$f(x_i) = \sum_{j=1}^{p} K(x_i, t_j)\lambda_j \qquad i = 0, 1, \ldots, n.$$

В этом случае функция

$$g(x)=\sum_{j=1}^{p}K(x,t_j)\lambda_j,$$

определенная на $X$, совпадает с $f$ в вы ранных точках $x_0, x_1, \ldots, x_n$ и будет мало отклоняться от $f(x)$.

В первом примере, рассматриваемом ниже, мы будем иметь дело с ядром $e^{xt}$, для которого преобразование 7.1 совпадает с преобразованием Лапласа — Стилтьеса $\int\limits_0^\infty e^{xt}d\sigma(t)$.

Конечная аппроксимация приводит нас к экспоненциальным многочленам с положительными коэффициентами.

Второй пример относится к преобразованию Стилтьеса $\int\limits_0^\infty d\sigma(t)/(x+t)$ рациональных функций.

Пример 7.1. Мы говорим, что функция $f(x)$ положительно определена на интервале $(a, b)$, если она может быть представлена в виде преобразования Лапласа — Стилтьеса, т. е.

$$f(x)=\int_{-\infty}^{\infty}e^{xt}d\mu(t), \qquad a<x<b, \tag{7.2}$$

где $\mu$ — неотрицательная регулярная сигма-конечная мера.

Определение может быть также сформулировано при помощи квадратичных форм (см. Виддер [1941]) Кроме того, говорят, что функция $f(x)$ абсолютно монотонна на интервале $(a, b]$, если $f$ имеет производные всех порядков на $(a, b]$ и $f^{(k)}(x)\geqslant 0$ $(a<x<b)$, $k=0, 1, 2 \ldots$ Бернштейн установил, что необходимое и достаточное условие для того, чтобы функция $f(x)$ была абсолютно монотонной на интервале $(-\infty, 0]$, состоит в том, что она допускает представления в виде

$$f(x)=\int_0^\infty e^{xt}d\sigma(t), \qquad -\infty<x\leqq 0, \tag{7.3}$$

где $\sigma$ — ограниченная неубывающая функция на $[0, \infty)$. В частном случае, когда $\sigma$ обладает только конечным числом точек роста, (7.3) приводится к экспоненциальному многочлену

$$P(x)=\sum_{i=1}^{m}\rho_i e^{x\xi_i} \tag{7.4}$$

$(0\leqq\xi_1<\xi_2<\ldots<\xi_m<\infty)$ с положительными коэффициентами. Говорят, что степень многочлена равна $2m$, если $\xi_1>0$, и $2m-1$, если $\xi_1 = 0$. Заметим, что степень является просто удвоенным индексом множества $\{\xi_1, \xi_2, \ldots, \xi_m\}$.

Мы формулируем следующий результат.

Теорема 7.1. *Пусть $f(x)$ — абсолютно монотонна на $(-\infty, 0]$, и предположим, что $f(x)$ не является многочленом степени, не большей чем $n-1$. Тогда для произвольных $x_1 < x_2 < \ldots < x_n < 0$ существует единственный экспоненциальный многочлен $P(x)$ степени $n$ с положительными коэффициентами такой, что*

$$P(x_k) = f(x_k), \qquad k = 1, 2, \ldots, n,$$

*и если $P(x) \neq f(x)$, то*

$$(-1)^{n-k}[f(x) - P(x)] > 0 \text{ для } x_k < x < x_{k+1}, \tag{7.5}$$

$(k = 0, 1, \ldots, n \; (x_0 = -\infty, \; x_{n+1} = 0)$.

Доказательство. Положим $u_i(t) = e^{x_i t}$ $(0 \leq t < \infty)$, $i = 1, 2, \ldots, n$ (обращаем внимание, что здесь нумерация функций производится от 1, а не от 0). Очевидно, что $u_i(t)/u_n(t) \to 0$ при $t \to \infty$.

В силу условий, наложенных на $f(x)$, равенства

$$c_i^0 = f(x_i) = \int_0^\infty u_i(t)\, d\sigma(t), \qquad i = 1, 2, \ldots, n, \tag{7.6}$$

определяют внутреннюю точку $\mathbf{c}^0$ моментного пространства, порожденного системой $\{u_i\}_{i=1}^n$. Следовательно, существует нижнее главное представление $\underline{\sigma}$ индекса $n/2$ такое, что

$$c_i^0 = \int_0^\infty u_i(t)\, d\underline{\sigma}(t), \qquad i = 1, 2, \ldots, n. \tag{7.7}$$

Искомый экспоненциальный многочлен $P(x)$ определяется равенством

$$P(x) = \int_0^\infty e^{xt} d\underline{\sigma}(t).$$

Для заданной точки $x$ такой, что $x_k < x < x_{k+1}$, положим $v_k(t) = e^{xt}$. Определив $u_{n+1}(t) = e^{x_{n+1}t} \equiv 1$ $(x_{n+1} = 0)$, мы можем применить теорему 6.2, в которой в качестве $B$ и $C$ берутся пустые множества. Тогда получим

$$(-1)^{n-k}\int_0^\infty e^{xt} d\sigma > (-1)^{n-k}\int_0^\infty e^{xt} d\underline{\sigma} \qquad (\sigma \neq \underline{\sigma}).$$

Заметим, что, применяя теорему 6.2, мы сдвинули все индексы на один, так как функции нумеровались от единицы. Поэтому $(-1)^{n-k}[f(x) - P(x)] > 0$ для $x_k < x < x_{k+1}$.

Теорему 7.1 можно обобщить на случай, когда на интерполирующий многочлен степени $n$ накладываются условия вида

$$f(x_k) = P(x_k), \quad f'(x_k) = P'(x_k), \ \ldots, \ f^{(n_k-1)}(x_k) = P^{(n_k-1)}(x_k)$$
$$(k = 1, \ldots, p),$$

где $x_1 < x_2 < \ldots < x_p < 0$, $n_1 + n_2 + \ldots + n_p = n$. Это обобщение получается, если применить приведенные выше рассуждения к $T$-системе

$$u_0(t) = e^{x_1 t}, \quad u_1(t) = te^{x_1 t}, \ \ldots, \ u_{n_1}(t) = t^{n_1} e^{x_1 t},$$
$$u_{n_1+1}(t) = e^{x_2 t}, \ldots \text{ и т. д.}$$

Пример 7.2. Второе преобразование, которое мы рассматриваем, имеет вид

$$f(x) = \int_0^\infty \frac{d\sigma(t)}{x+t}, \qquad 0 < x < \infty, \tag{7.8}$$

если $\sigma$ — неотрицательная мера.

Если мера $\sigma$ в (7.8) обладает только конечным числом точек роста $0 \leq \xi_1 < \xi_2 < \ldots < \xi_m$ с соответствующими весами $\rho_j$ ($j = 1, 2, \ldots, m$), то (7.8) представляет собой рациональную функцию

$$R(x) = \sum_{j=1}^{m} \frac{\rho_j}{\xi_j + x} \qquad (\rho_j > 0, \ j = 1, \ldots, m). \tag{7.9}$$

Мы назовем $2m$ степенью $R$, если $\xi_1 > 0$; если $\xi_1 = 0$, то степенью $R$ назовем $2m - 1$.

Класс функций, допускающих представление (7.8), может быть описан аксиоматически (см. Ахиезер [1965]), как класс аналитических функций, удовлетворяющих следующим условиям:

(i) $f(z)$ — аналитическая в комплексной плоскости, разрезанной вдоль полуоси $(-\infty, 0)$;

(ii) $\operatorname{Im} f(z)/\operatorname{Im} z < 0$ при $\operatorname{Im} z \neq 0$;

(iii) существует конечный верхний предел $\overline{\lim\limits_{|y|\to\infty}} \, yf(iy)$ для вещественных $y$. Отсюда непосредственно получается, что

$$S_0(f) = \int_0^\infty d\sigma(t) = i \overline{\lim_{|y|\to\infty}} \, yf(iy).$$

Следующая теорема связывает преобразования Стилтьеса с рациональными функциями.

Теорема 7.2. *Пусть $f(x)$ — функция вида (7.8), не совпадающая с рациональной функцией степени, не большей чем $n$. Тогда для произвольных $0 < x_1 < x_2 < \ldots < x_n$ существует* *единственная рациональная функция* $R^*(x)$ *степени* $n+1$ *с положительными коэффициентами, удовлетворяющая равенствам*

$$R^*(x_k) = f(x_k), \quad k = 1, 2, \ldots, n, \qquad S_0(R^*) = S_0(f), \tag{7.10}$$

$$(-1)^{n-k}[f(x) - R^*(x)] > 0, \quad x_k < x < x_{k+1} \tag{7.11}$$

($k = 0, 1, \ldots, n$; $x_0 = 0$, $x_{n+1} = \infty$). *Более того, существует единственная рациональная функция* $R^{**}(x)$ *степени* $n$, *удовлетворяющая условиям*

$$R^{**}(x_k) = f(x_k), \quad k = 1, 2, \ldots, n, \quad S_0(R^{**}) < S_0(f),$$

$$(-1)^{n-k}[f(x) - R^{**}(x)] < 0, \quad x_k < x < x_{k+1} \tag{7.12}$$

($k = 0, 1, \ldots, n$; $x_0 = 0$, $x_{n+1} = \infty$).

Доказательство. Положим $u_k(t) = 1/(x_{k+1} + t)$ для $k = 0, 1, \ldots, n-1$, и пусть $u_n(t) \equiv 1$. Заметим, что $u_k(\infty) = \delta_{kn}$, поэтому $\{u_i\}_0^n$ — $T$-система на $[0, \infty]$. Последнее можно проверить непосредственно, если принять во внимание значение определителя Коши, приведенное в примере 4 § 3 гл. I. Точка $\mathbf{c}$ с координатами $c_k = f(x_{k+1})$, $k = 0, 1, \ldots, n-1$, и $c_n = S_0(f)$ лежит внутри $\mathcal{M}_{n+1}$ и, следовательно, допускает верхнее и нижнее главные представления $\bar{\sigma}$ и $\underline{\sigma}$ индекса $(n+1)/2$. Заметим, что $\bar{\sigma}$ имеет положительную массу на бесконечности. Многочлены $R^*$ и $R^{**}$ определяются равенствами

$$R^*(x) = \int\limits_{[0,\infty)} \frac{d\underline{\sigma}}{x+t}, \quad R^{**}(x) = \int\limits_{[0,\infty)} \frac{d\bar{\sigma}}{x+t}, \qquad 0 < x < \infty.$$

Ясно, что $R^*(x_k) = R^{**}(x_k) = f(x_k)$ для $k = 1, 2, \ldots, n$. Многочлен $R^*$ имеет степень $n+1$ и $S_0(R^*) = S_0(f)$. Следовательно, масса в бесконечности отсутствует, $R^{**}$ имеет степень $n$ и $S_0(R^{**}) < S_0(f)$.

Если $x_k < x < x_{k+1}$, то положим $\Omega(t) = (-1)^{n-k}/(x+t)$. Тогда функции $u_0, u_1, \ldots, u_n, \Omega$ образуют $T$-систему на $[0, \infty]$. Поэтому по теореме 5.1

$$\int \Omega(t)\, d\sigma = (-1)^{n-k} \int \frac{d\sigma}{x+t} > (-1)^{n-k} \int \frac{d\underline{\sigma}}{x+t},$$

$$\int \Omega(t)\, d\sigma < (-1)^{n-k} \int \frac{d\bar{\sigma}}{x+t},$$

откуда следуют соотношения (7.11) и (7.12).

Первый результат, относящийся к интерполяции (пример 7.1), получен Бернштейном. Крейн [1951] и Розенблюм [1951] независимо уточнили его. В примере 7.2 речь идет о преобразованиях Стилтьеса и интерполировании аналитической функции, отображающей верхнюю полуплоскость в себя рациональными функциями. Исторически Пик, Левнер, Крейн и другие внесли свой вклад в развитие этих вопросов. Теорема 7.2 в приведенной формулировке принадлежит Крейну [1951], но здесь приведено другое ее доказательство. Преобразование Стилтьеса играет решающую роль в исследовании задачи об определенности и единственности в классической проблеме моментов (см. Шохат и Тамаркин [1943] и Ахиезер [1965]).

## § 8. Теоремы о представлении положительных многочленов

В этом параграфе мы перенесем теоремы о представлении § 10 гл. II на случай $T$-систем, определенных на полупрямой $[0, \infty)$. Конкретные примеры представлений приведены в следствии 8.1.

Материал этого параграфа до леммы 8.1 взят из работы Карлина [1963].

Мы предполагаем, что система $\{u_i\}_0^n$ имеет тип II, т. е.:

(i) $u_n(t) > 0$ для больших $t$;

(ii) $\lim\limits_{t\to\infty} u_i(t)/u_n(t) = 0, \qquad i = 0, 1, \ldots, n-1,$

(iii) $\{u_i(t)\}_0^{n-1}$ и $\{u_i(t)\}_0^n$ — $T$-системы.

Как и раньше, выберем непрерывную функцию $w(t) > 0$ для $t \in [0, \infty)$ так, чтобы

$$\lim_{t\to\infty} \frac{u_n(t)}{w(t)} = 1$$

и положим

$$v_k(t) = \begin{cases} \dfrac{u_k(t)}{w(t)} & \text{для } t \in [0, \infty), \\ \delta_{k,n} & \text{для } t = \infty \end{cases}$$

$(k = 0, 1, \ldots, n)$.

Система $\{v_k(t)\}_0^n$ является $T$-системой на замкнутом интервале $[0, \infty]$, так что $\{g_k(x)\}_0^n$, определенная равенствами

$$g_k(x) = v_k(\operatorname{tg} \pi x/2), \qquad k = 0, 1, \ldots, n,$$

становится $T$-системой на интервале $[0, 1]$. К системе $\{g_k(x)\}_0^n$ можно теперь применить рассуждения § 10 гл. II. Тогда получим следующую теорему.

Теорема 8.1 *Если* $u(t) = \sum\limits_{i=0}^{n} a_i u_i(t)$ *положителен на* $[0, \infty)$ *и* $a_n > 0$, *то существует единственное представление*

$$u(t = \underline{u}(t) + \overline{u}\ t),$$

*в которой* $\underline{u}(t)$ *и* $\overline{u}(t)$ — *неотрицательные многочлены, удовлетворяющие следующим свойствам:*

(i) *Если* $n = 2m$, *то многочлены* $\underline{u}(t)$ *и* $\bar{u}(t)$ *обладают каждый* $m$ *различными нулями* $\{\underline{t}_i\}_1^m$ *и* $\{\bar{t}_i\}_1^m$, *причем* $0 = \bar{t}_1 < \underline{t} < \bar{t}_2 < \ldots$ $\ldots < \bar{t}_m < \underline{t}_m < \infty$.

(ii) *Если* $n = 2m + 1$, *то* $\underline{u}(t)$ *и* $\bar{u}(t)$ *имеют нули* $\{\underline{t}_i\}_1^{m+1}$ *и* $\{\bar{t}_i\}_1^m$ *соответственно, причем* $0 = \underline{t}_1 < \bar{t}_1 < \underline{t}_2 < \ldots < \bar{t}_m < \underline{t}_{m+1} < \infty$.

(iii) *Коэффициент при* $u_n(t)$ *в* $\underline{u}(t)$ *равен* $a_n$.

Замечание 8.1. Сравнивая эту теорему со следствием 10.1 (а) гл. II, можно заметить, что заключения в них почти одинаковы, только отбрасывается требование, чтобы на конце интервала $b$ был корень. Это требование учитывается в $\bar{u}(t)$, имеющем корень на $+\infty$ и выражается в том, что старший коэффициент $a_n$ в многочлене $\bar{u}$ равен нулю. В частности, если $u_i(t) = t^i$, $i = 0, 1, \ldots, n$, то многочлен $\bar{u}$ имеет степень $n - 1$ (см. замечание 8.2 ниже).

Следствие 8.1. *Если* $\{u_i(t)\}_0^n$ — *T-система, определенная на* $[0, \infty)$ *равенствами* $u_i(t) = t^i$, $i = 0, 1, \ldots, n$, *то*:

(i) *для любого многочлена* $P_{2m}(t) = \sum_{i=0}^{2m} a_i t^i$, *положительного на* $[0, \infty)$, *с* $a_{2m} > 0$ *существует единственное представление*

$$P_{2m}(t) = a_{2m} \prod_{j=1}^{m} (t - \underline{t}_j)^2 + \beta t \prod_{j=2}^{m} (t - \bar{t}_j)^2, \qquad \beta > 0,$$

*где* $0 = \bar{t}_1 < \underline{t}_1 < \bar{t}_2 < \ldots < \bar{t}_m < \underline{t}_m < \infty$;

(ii) *для каждого многочлена* $P_{2m+1}(t) = \sum_{i=0}^{2m+1} a_i t^i$, *положительного на* $[0, \infty)$, *с* $a_{2m+1} > 0$ *существует единственное представление*

$$P_{2m+1}(t) = a_{2m+1} t \prod_{j=2}^{m+1} (t - \underline{t}_j)^2 + \beta \prod_{j=1}^{m} (t - \bar{t}_j)^2,$$

где $\beta > 0$ и $0 = t_1 < \bar{t}_1 < \underline{t}_2 < \bar{t}_2 < \ldots < \bar{t}_m < \underline{t}_{m+1} < \infty$.

Замечание 8.2. Если в предыдущем следствии положить $a_n = a_{n-1} = \ldots = a_{n-r+1} = 0$ и $a_{n-r} > 0$, то говорят, что $P_n(t)$ имеет нуль на бесконечности кратности $r$. Как и в следствии 10.3 теоремы 10.3 гл. II, каждый из многочленов в этом представлении должен иметь нуль кратности $r$ на бесконечности, так что, по существу, нам следовало бы иметь дело со случаем, когда $n$ заменяется на $n - r$.

Прежде чем переходить к обсуждению структуры граничных и крайних точек пространства многочленов $\mathcal{P}_{n+1}$, мы приведем более прямую конструкцию представлений для специального случая следствия 8.1. Эти конструкции обладают рядом интересных дополнительных свойств непрерывности нулей и скалярного множителя, которые используются в рассуждениях о неподвижной точке при доказательстве теоремы 10.1 гл. II.

Предполагается, что все рассматриваемые ниже многочлены $P_n(t)$ имеют положительный старший коэффициент, который нормирован и, следовательно, равен единице.

Лемма 8.1. *Если $P_{2m}(t)>0$ на $[0, \infty)$ и его старший коэффициент равен единице, то для любого числа $t_m>0$ существует единственный многочлен*

$$P_m^2(t)=\prod_{j=1}^{m}(t-t_j)^2, \qquad 0<t_1<\dots<t_m,$$

*и число $\alpha>0$, для которых $P_{2m}(t)\geqslant\alpha^{-1}P_m^2(t)$ $(0\leqslant t\leqslant t_m)$, причем равенство имеет место в каждом из интервалов $(t_{i-1}, t_i)$, $i=1,\dots,m$,* где $t_0=0$. Если $\alpha\leqslant 1$, *то равенство в интервале $[0, t_1]$ имеет место при $t=0$.*

Доказательство. Пусть

$$\alpha_j=\max_{t_j\leqslant t\leqslant t_{j+1}}\frac{P_m^2(t)}{P_{2m}(t)}, \qquad j=0, 1, \dots, m-1; \quad t_0=0.$$

Так же, как и в теореме 10.1 гл. II, применяя рассуждение о неподвижной точке, можно показать, что существует единственное множество $(t_1, \dots, t_{m-1})$, для которого $\alpha_j=\alpha>0$ $(j=0, 1, \dots, m-1)$ и для которого выполняются условия леммы. Если $\alpha\leqslant s$, то разность $P_{2m}(t)-\alpha^{-1}P_m^2(t)$ будет иметь количество нулей, большее своей степени, если только нуль в $[0, t_1]$ не находится в $t=0$.

Лемма 8.2. *В условиях леммы 8.1 множитель $\alpha=\alpha(t_m)$ является возрастающей непрерывной функцией относительно $t_m$, такой, что $\alpha(t_m)\to\infty$ при $t_m\to\infty$. Более того, $t_j(t_m)$, $j=1, 2, \dots, m-1$, являются возрастающими непрерывными функциями от $t_m$, причем, если $t_m\neq t'_m$, то не существует $i$ и $j$ таких, что*

$$t_j(t'_m)\leqslant t_i(t_m)<t_{i+1}(t_m)\leqslant t_{j+1}(t'_m). \tag{8.1}$$

Доказательство. Сначала мы докажем, что если $t_m\leqslant T$, то можно найти постоянную $C$ такую, что $\alpha(t_m)\leqslant C$. Пусть $p=\min\limits_{0\leqslant t\leqslant T}P_{2m}(t)$. Если $t_m\leqslant T$, то очевидно, что

$$\alpha(t_m)\leqslant C=\frac{1}{p}\min_{(t_j)}\max_{0\leqslant t\leqslant T}P_m^2(t),$$

где $t_j$ удовлетворяют условиям $0<t_1<t_2<\dots<t_m$. Поэтому функции $t_j(t_m)$, $j=1, \dots, m-1$, и $\alpha(t_m)$ ограничены, если ограничены $t_m$. Непрерывность $t_j(t_m)$ и $\alpha(t_m)$ можно вывести из единственности в лемме 8.1.

Если бы (8.1) было справедливо для некоторых $t_m\neq t'_m$, то разность между построенными многочленами имела большее чем возможно количество нулей. Из неравенств (8.1) следует, что каждая из $t_j$ возрастает, как функция от $t_m$.

Чтобы показать возрастание $\alpha$ по $t_m$, заметим, что $\alpha\to 0$ при $t_m\to 0$, и если $\alpha(t_m)=\alpha(\bar{t}_m)$ для $t_m\neq\bar{t}_m$, то $\alpha^{-1}P_m^2$ и $\alpha^{-1}\bar{P}_m^2$ имеют $2m$ общих нулей, в то время как их разность имеет степень $2m-1$. Наконец, чтобы показать, что $\alpha\to\infty$ при $t_m\to\infty$, выберем $C>0$ так, чтобы существовала последовательность

$\{t_m^{(k)}\} \to \infty$ такая, что $\alpha(t_m^{(k)}) < C$. Тогда бы мы имели равенство

$$CP_{2m}(t) > \prod_{j=1}^{m} (t - t_j^{(k)})^2$$

для $0 \leqslant t \leqslant t_m^{(k)}$, которое явно нарушается при $t = 0$ и достаточно больших $k$.

Лемма 8.3. *Существует единственное представление многочлена $P_{2m}(t) > 0$ $(0 \leqslant t < \infty)$ в виде суммы*

$$P_{2m}(t) = \prod_{j=1}^{m} (t - \underline{t}_j)^2 + \beta t \prod_{j=2}^{m} (t - \bar{t}_j)^2, \qquad \beta > 0,$$

$(0 = \bar{t}_1 < \underline{t}_1 < \bar{t}_2 < \underline{t}_2 < \ldots < \bar{t}_m < \underline{t}_m < \infty)$.

Доказательство. Так как по лемме 8.2 $\alpha \to \infty$ при $t_m \to \infty$, то найдется $\underline{t}_m$ такое, что $\alpha(\underline{t}_m) = 1$. Найдем соответствующий единственный многочлен $P_m^2(t) = \prod_{j=1}^{m} (t - \underline{t}_j)^2$ такой, что разность $P_{2m} - P_m^2$ является многочленом степени $2m - 1$, неотрицательна при $0 \leqslant t \leqslant t_m$ и обращается в нуль при $t = 0$ и в точках $\bar{t}_j$ $(\underline{t}_j < \bar{t}_{j+1} < \underline{t}_{j+1})$. Поэтому эта разность представима в виде

$$\beta t \prod_{j=2}^{m} (t - \bar{t}_j)^2, \quad \beta > 0.$$

Для многочленов нечетной степени результат получается аналогично.

## § 9. Крайние точки и граница $\mathcal{P}_{n+1}[0, \infty)$

Последние четыре параграфа этой части будут посвящены различным уточнениям наших ранее полученных результатов в частном случае $T$-системы $\{t^i\}_0^n$ на $[0, \infty)$. В этом параграфе мы исследуем структуру границы и крайних точек выпуклого конуса $\mathcal{P}_{n+1}[0, \infty)$, состоящего из всех неотрицательных многочленов $\sum_0^n a_i t^i$ на $[0, \infty)$.

Говорят, что многочлен $P(t) = \sum_{i=0}^{n} a_i t^i$ имеет $k$ $(0 \leq k \leq n)$ нулей на бесконечности, если $a_n = a_{n-1} = \ldots = a_{n-k+1} = 0$ и $a_{n-k} \neq 0$.

Теорема 9.1 *Точка $P(t)$ пространства $\mathcal{P}_{n+1}[0, \infty)$ является:*

(a) *внутренней точкой тогда и только тогда, когда она не имеет нулей в* $[0, \infty]$;

(b) *граничной точкой тогда и только тогда, когда она имеет по крайней мере один нуль в* $[0, \infty]$;

(c) *точкой крайнего луча тогда и только тогда, когда она имеет $n$ нулей в* $[0, \infty]$.

Доказательство. Чтобы доказать (а), мы заметим, что если $P(t)$ не имеет нулей в $[0, \infty]$, то $a_n > 0$ и $P(t) > 0$ для $t \in [0, \infty)$. В этом случае все его нули являются комплексными или действительными отрицательными и останутся таковыми при малом изменении коэффициентов $a_i$. Из этого следует, что $P(t)$ — внутренняя точка.

Если у $P(t)$ есть нули в $[0, \infty]$ и один из них конечен, например $t_0$, то $P(t) = (t - t_0) P_1(t)$ становится отрицательным для некоторого $t$, если $P(t)$ заменить на $(t - t_0 - \varepsilon) P_1(t)$ для достаточно малого $\varepsilon > 0$. Если $P(t)$ имеет нуль в $\infty$, то $a_n = 0$, так что $P(t)$ станет отрицательным для достаточно больших $t$ при добавлении члена $-\varepsilon t^n$. Доказательство (а) закончено. Утверждение (b) следует непосредственно из (а).

Чтобы доказать (с), мы предположим сначала, что многочлен $P(t) \in \mathcal{P}_{n+1}$ не имеет $n$ нулей на $[0, \infty]$. В этом случае у $P(t)$ имеются комплексные или действительные отрицательные нули. Если у $P(t)$ имеется комплексный нуль $a + bi$ $(b > 0)$, то $P(t) = [(t-a)^2 + b^2] \cdot P_1(t) = (t-a)^2 P_1(t) + b^2 P_1(t)$. Очевидно, что $P_1(t) \geqslant 0$ для $t \in [0, \infty)$, поэтому как $(t-a)^2 P_1(t)$, так и $b^2 P_1(t)$ принадлежат $\mathcal{P}_{n+1}$. Верхнее представление показывает, что $P(t)$ не лежит на крайнем луче $\mathcal{P}_{n+1}$.

Таким же образом можно доказать, что если $P(t)$ имеет отрицательный нуль $t_0$, то $P(t)$ не лежит ни на каком крайнем луче $\mathcal{P}_{n+1}$.

Если $P(t) \in \mathcal{P}_{n+1}$ имеет $n$ нулей в $[0, \infty]$, то $P(t)$ порождает экстремальный луч в $\mathcal{P}_{n+1}$. Доказательство этого факта просто, и мы его опускаем.

## § 10. Классический критерий для моментных точек

Пусть $u_i(t) = t^i$, $i = 0, 1, \ldots, n$, на интервале $[0, \infty)$. Воспользуемся результатами § 8 для вывода критерия принадлежности точки $(c_0, \ldots, c_n)$ к $\mathcal{M}_{n+1}$. Из следствия 8.1 нам известно, что каждый неотрицательный многочлен $\sum_{i=0}^{n} a_i t^i = P(t)$ на $[0, \infty)$ можно записать в виде

$$P(t) = P_m^2(t) + tQ_{m-1}^2(t), \text{ если } n = 2m,$$

$$P(t) = tP_m^2(t) + Q_m^2(t), \quad \text{если } n = 2m + 1,$$

где $P_m$ и $Q_m$ — многочлены степени $m$.

Необходимое и достаточное условие для того, чтобы $\mathbf{c}$ принадлежала $\overline{\mathcal{M}}_{n+1}$, состоит в том, что для всех $\mathbf{a} \in \mathcal{P}_{n+1}$

$$(\mathbf{c}, \mathbf{a}) \geqslant 0. \tag{10.1}$$

Для проверки (10.1) достаточно рассмотреть только точки $\mathbf{a}$, соответствующие многочленам вида $P_m^2(t)$ и $tQ_{m-1}^2(t)$, когда $n$ —

четное и соответствующие $tP_m^2(t)$ и $Q_m^2(t)$, когда $n$ — нечетное. Расписывая (10.1) в этих случаях, мы выводим следующую теорему (см. теорему 1.1 гл. IV).

Теорема 10.1. *Точка* $\mathbf{c} = (c_0, c_1, \ldots, c_n)$ *принадлежит* $\overline{\mathcal{M}}_{n+1}$ *тогда и только тогда, когда матрицы*

$$\| c_{i+j} \|_{i,j=0}^{m}, \quad \| c_{i+j+1} \|_{i,j=0}^{m-1} \quad \text{для} \quad n = 2m,$$

$$\| c_{i+j} \|_{i,j=0}^{m}, \quad \| c_{i+j+1} \|_{i,j=0}^{m} \quad \text{для} \quad n = 2m+1$$

*неотрицательно определены. Точка* $\mathbf{c} = (c_0, \ldots, c_n) \in \operatorname{Int} \overline{\mathcal{M}}_{n+1} = \operatorname{Int} \mathcal{M}_{n+1}$ *тогда и только тогда, когда вышеуказанные матрицы положительно определены.*

В том случае, когда квадратичные формы неотрицательно определены, мы получим представление в виде

$$c_i = \int_0^\infty t^i d\sigma(t), \quad i = 0, 1, \ldots, n-1, \quad c_n = \int_0^\infty t^n d\sigma(t) + \lambda \tag{10.2}$$

($\lambda \geqslant 0$, $\sigma$ — ограниченная неотрицательная мера). Если обе формы положительно определены, то в представлении (10.2) $\lambda = 0$.

Приведенный выше результат и обсуждения с разных точек зрения имеются у Шохата и Тамаркина [1943].

## § 11. Соотношение между $\mathcal{P}_{n+1}$ и $\mathcal{M}_{n+1}$

В этом параграфе мы изучим связь между моментным пространством $\mathcal{M}_{n+1}$ и дуальным ему пространством $\mathcal{P}_{n+1}$.

Пусть $\underline{P}_k(t)$, $k = 0, 1, \ldots, m$ $(m = [n/2])$ — ортогональные многочлены по отношению к некоторой конечной мере на $[0, \infty)$, и пусть $\underline{Q}_k(t)$, $k = 0, 1, \ldots, m$, — многочлены, ортогональные по отношению к мере $t\, d\sigma(t)$. Заметим, что если $c_k = \int_0^\infty t^k d\sigma$, $k = 0, 1, \ldots, n$, то многочлены $\underline{P}_k(t)$ и $\underline{Q}_k(t)$ пропорциональны определителям

$$\underline{\Delta}_{2k}(t) = \begin{vmatrix} c_0 \ldots c_{k-1} & 1 \\ c_1 \ldots c_k & t \\ \ldots\ldots & \ldots \\ c_k \ldots c_{2k-1} & t^k \end{vmatrix}, \tag{11.1}$$

$$\underline{\Delta}_{2k+1}(t) = \begin{vmatrix} c_1 \ \ldots c_k & 1 \\ c_2 \ \ldots c_{k+1} & t \\ \ldots\ldots & \ldots \\ c_{k+1} \ldots c_{2k} & t^k \end{vmatrix}, \tag{11.2}$$

соответственно.

Далее, пусть $M^n$ и $P^n$ являются сечениями $\mathcal{M}_{n+1}$ и $\mathcal{P}_{n+1}$ соответственно, полученными наложением нормирующих условий

$$c_0 = 1 \quad \text{и} \quad \sum_{k=0}^{n} k! a_k = 1,$$

где

$$\sum_{i=0}^{n} a_i t^i \in \mathcal{P}_{n+1}.$$

Теорема 11.1. *Если $n = 2m$, то существует взаимно однозначное соответствие между $\operatorname{Int} M^{2m}$, упорядоченными парами многочленов $(\underline{P}_m, \underline{Q}_m)$ и открытым симплексом строго перемежающихся корней*

$$0 = s_0 < t_1 < s_1 < \ldots < t_m < s_m. \tag{11.3}$$

*Если $n = 2m-1$, то существуют взаимно однозначные соответствия между $\operatorname{Int} \mathcal{M}^{2m-1}$, упорядоченными парами многочленов $(\underline{P}_m, \underline{Q}_{m-1})$ и открытым симплексом строго перемежающихся корней*

$$0 = s_0 < t_1 < s_1 < \ldots < s_{m-1} < t_m. \tag{11.4}$$

Замечание 11.1. Для $n = 2m$ желаемое соответствие достигается при помощи специального выбора корней $0 = s_0 < s_1 < \ldots$ $\ldots < s_m$ и $t_1 < \ldots < t_m$, как корней нижних главных представлений точек $(c_0, \ldots, c_{2m})$ и $(c_0, \ldots, c_{2m-1})$ соответственно. Далее эти пары множеств корней являются нулями многочленов $t\underline{Q}_m(t)$ и $\underline{\mathcal{P}}_m(t)$ соответственно. Это соответствие можно показать схематично следующим образом:

корни нижнего главного представления

$$(c_0, \ldots, c_{2m}) \begin{array}{c}\leftarrow\\ \rightarrow\end{array} \left\{\begin{array}{l}(c_0, \ldots, c_{2m-1}, c_{2m})\\ (c_0, \ldots, c_{2m-1})\end{array}\right\} \begin{array}{c}\leftarrow\\ \rightarrow\end{array} \left\{\begin{array}{l}0 < s_1 < \ldots < s_m\\ t_1 < \ldots < t_m\end{array}\right\} \begin{array}{c}\rightarrow\\ \leftarrow\end{array}$$

корни многочленов

$$\begin{array}{c}\rightarrow\\ \leftarrow\end{array} \left\{\begin{array}{l}t\underline{Q}_m(t)\\ \underline{P}_m(t)\end{array}\right\}.$$

Следует заметить, что это соответствие отличается от соответствия, установленного в гл. II (см. замечание 5.1 гл. II), в котором единственной моментной точке $(c_0, \ldots, c_{2m})$ ставились в соответствие корни верхнего и нижнего главных представлений.

Доказательство теоремы 11.1 ($n = 2m$). Точка $\mathbf{c} = (c_0, \ldots, c_{2m})$ определяет множества перемежающихся точек $\{s_j\}_0^m$ и

$\{t_j\}_1^m$ как корни нижних главных представлений $(c_0, \ldots, c_{2m})$ и $(c_0, \ldots, c_{2m-1})$ соответственно. Следовательно, множества $\{t_j\}_1^m$ и $\{s_j\}_0^m$ удовлетворяют (11.3). Мы сначала определим $\{\lambda_j\}_1^m$ и $\{\lambda'_j\}_0^m$ так, чтобы

$$\sum_{j=1}^{m} \lambda_j t_j^i - \sum_{j=0}^{m} \lambda'_j s_j^i = 0, \quad i = 0, 1, \ldots, 2m-1.$$

Это линейная система $2m$ уравнений с $2m+1$ неизвестными $\lambda_1, \ldots, \lambda_m, -\lambda'_0, \ldots, -\lambda'_m$ с матрицей коэффициентов ранга $2m$. Известно, что соответствующие компоненты ее решения (с точностью до числового множителя) пропорциональны (с чередующимися знаками) минорам порядка $2m$, полученным последовательным вычеркиванием одного столбца из матрицы коэффициентов. Легко видеть, что каждый такой минор порядка $2m$ должен быть определителем Вандермонда. Принимая во внимание предположение о том, что корни $\{t_j\}_1^m$ и $\{s_j\}_0^m$ чередуются, как показано в (11.4), мы выводим, что величины $\lambda$ и $\lambda'$ имеют неизменный знак. Теперь числовой множитель определяется так, чтобы $\sum_{j=1}^{m} \lambda_j = 1$. Далее мы полагаем

$$\begin{aligned} c_i &= \sum_{j=1}^{m} \lambda_j t_j^i = \sum_{j=0}^{m} \lambda'_j s_j^i, \quad i = 0, 1, \ldots, 2m-1, \\ c_{2m} &= \sum_{j=0}^{m} \lambda'_j s_j^{2m}. \end{aligned} \tag{11.5}$$

Очевидно, что $(c_0, \ldots, c_{2m-1})$ и $(c_0, c_1, \ldots, c_{2m})$ — требуемые моментные точки.

Непрерывность соответствия между $\operatorname{Int} M^{2m}$ и множеством всех пар перемежающихся корней очевидна. Если $\mathbf{c}^{(r)} \to \mathbf{c} \in \operatorname{Int} M^{2m}$, то $\mathbf{c}^{(r)}$ и $\mathbf{c}$ лежат внутри соответствующего сечения, взятого из моментного пространства на $[0, b]$ для достаточно большого $b$. Теперь непрерывность доказывается, как при доказательстве теоремы 8.1 гл. II. Непрерывность отображения пар перемежающихся корней в сечение $M^{2m}$, очевидно, следует из (11.5), так как $\lambda_j$ и $\lambda'_j$ являются непрерывными функциями этих корней.

Теперь мы покажем, что соответствие между парами $(\underline{P}_m, \underline{Q}_m)$ и $\operatorname{Int} M^{2m}$ взаимно однозначно и непрерывно. Взаимная однозначность и непрерывность отображения $\operatorname{Int} M^{2m}$ на пары многочленов $(\underline{P}_m, \underline{Q}_m)$ получается из (11.1) и (11.2). Обратно, если $\underline{P}_m$ и $\underline{Q}_m$ многочлены из систем $\{\underline{P}_k\}_0^m$ и $\{\underline{Q}_k\}_0^m$, которые ортонормированы относительно $d\sigma$ и $t\,d\sigma$, то (см. лемму 2.2 гл IV) $\underline{P}_m(t)$ и $t\underline{Q}_m(t)$ обращаются в нуль в корнях нижних главных представлений точек $(c_0, \ldots, c_{2m-1})$

и $(c_0, \ldots, c_{2m})$, где $c_k = \int\limits_0^\infty t^k d\sigma(t)$, $k = 0, 1, \ldots, 2m$. Более того, эти корни однозначно определяют моментную точку $(c_0, \ldots, c_{2m})$, и соответствие между ними непрерывно. Доказательство закончено.

Теорема 11.2, приводимая ниже, показывает, что внутренняя часть сечения моментного пространства $\mathcal{M}_{n+1}$ и симплекс пар перемежающихся корней естественным образом проективно эквивалентны.

Теорема 11.2. *Пусть $p_m^2(t)$ и $tq_{m-1}(t)$ — многочлены, полученные нормировкой $P_m^2(t)$ и $tQ_{m-1}^2(t)$ соответственно, так, что они лежат в $P^{2m}$. Тогда существует гомеоморфизм между* $\operatorname{Int} P^{2m}$ *и* $\operatorname{Int} \mathcal{M}^{2m}$ *такой, что многочлены вида*

$$\lambda p_m^2 + (1-\lambda)\, tq_{m-1}^2 \quad (0 < \lambda < 1),$$

*где $p_m$ и $q_{m-1}$ фиксированы, а $\lambda$ меняется в единичном интервале, соответствуют моментным точкам, у которых $c_0, c_1, \ldots, c_{2m-1}$ постоянны, а $c_{2m}$ меняется в интервале $(\bar{c}_{2m}, \infty)$. Такое же соответствие имеется и между* $\operatorname{Int} P^{2m+1}$ *и* $\operatorname{Int} M^{2m+1}$.

Доказательство. В силу следствия 8.1 каждый многочлен $P(t) \in \operatorname{Int} P^{2m}$ может быть единственным образом представлен в виде

$$P(t) = \alpha \prod_{j=1}^{m} (t - t_j)^2 + \beta t \prod_{j=1}^{m-1} (t - s_j)^2,$$

где $\alpha > 0$, $\beta > 0$, а $\{t_j\}_1^m$ и $\{s_j\}_0^{m-1}$ удовлетворяют (11.4). Теорема 11.1 утверждает, что существует единственная моментная точка $(c_0, \ldots, c_{2m-1})$ в $\operatorname{Int} M^{2m-1}$, соответствующая корням $\{t_j\}_1^m$ и $\{s_j\}_0^{m-1}$. Положим

$$p_m^2(t) = k \prod_{j=1}^{m} (t - t_j)^2, \quad tq_{m-1}^2(t) = k' t \prod_{j=1}^{m-1} (t - s_j)^2,$$

где $k$ и $k'$ — постоянные, определенные так, чтобы $p_m^2(t)$ и $tq_{m-1}^2(t)$ лежали в $\operatorname{Int} P^{2m}$.

Теперь рассмотрим семейство многочленов

$$P(t; \lambda) = \lambda p_m^2(t) + (1-\lambda)\, tq_{m-1}^2(t), \qquad 0 < \lambda < 1. \tag{11.6}$$

Очевидно, что $P(t; \lambda)$ при каждом $\lambda\ (0 < \lambda < 1)$ принадлежит $\operatorname{Int} P^{2m}$, а при некотором фиксированном $\lambda$ $P(t; \lambda)$ совпадает с заданным многочленом $P(t)$.

Требуемый гомеоморфизм будет полностью построен, если мы сопоставим линейный сегмент $l_1$ в $P^{2m}$, определенный в (11.6) при фиксированных $p_m$ и $q_{m-1}$, линейному сегменту $l_2$ в $M^{2m}$, определяемому при фиксированных $c_0, \ldots, c_{2m-1}$ и $c_{2m}$, изменяющемуся в

$(\underline{c}_{2m}, \infty)$. Этого можно достичь, отображая интервал $(0,1)$ на интервал $(\underline{c}_{2m}, \infty)$ при помощи преобразования

$$c_{2m} = \underline{c}_{2m} + \operatorname{tg} \pi\lambda/2.$$

Это отображение $\operatorname{Int} M^{2m}$ в $\operatorname{Int} P^{2m}$, как легко видеть, взаимно однозначно и взаимно непрерывно.

## § 12. Соотношение между $M^n[0, a]$ и $M^n[0, \infty)$

В этом параграфе мы выведем теорему непрерывности, связывающую моментные последовательности из моментного пространства на конечном интервале с моментными последовательностями из моментного пространства на полупрямой.

Можно образовать сечения $M^{2m}[0, \infty)$ и $M^{2m-1}[0, \infty)$, переходя к пределу от $M^{2m}[0, a]$ и $M^{2m-1}[0, a]$ соответственно, при $a \to \infty$ Этот предельный процесс сохраняет определенные соотношения в том смысле, который будет объяснен ниже. Результаты будут сформулированы и доказаны только для случая $M^{2m-1}[0, a]$. Аналогичные результаты справедливы также и в четном случае.

Если $(c_0, \ldots, c_{2m-1}) \in \operatorname{Int} M^{2m-1}[0, 1]$, то многочлены, ортонормированные относительно $d\sigma$ и $t(1-t)\,d\sigma$, обозначаются через $\{\underline{P}_k\}_0^m$ и $\{\overline{P}_k\}_0^{m-1}$ (см. § 2 гл. IV).

Следуя ходу рассуждений теоремы 11.1, мы можем построить взаимно однозначные непрерывные соответствия между точками $\operatorname{Int} M^{2m-1}[0, 1]$, упорядоченными парами многочленов $(\underline{P}_m, \overline{P}_{m-1})$ и открытым симплексом

$$0 < t_1 < w_1 < t_2 < \ldots < w_{m-1} < t_m < 1, \tag{12.1}$$

где $\{t_j\}_1^m$ — корни $\underline{P}_m$, а $\{w_j\}_1^{m-1}$ — корни $\overline{P}_{m-1}$. (Это отображение отличается от отображения теоремы 11.1 тем, что мы теперь соотносим множество (12.1) с верхним и нижним главными представлениями). Далее единственными опорными многочленами (см. следствие 2.2 (а) гл. IV) в точках $\underline{c} = (c_0, \ldots, c_{2m-1}, \underline{c}_{2m})$ и $\overline{c} = (c_0, \ldots, c_{2m-1}, \overline{c}_{2m})$ являются многочлены $\underline{P}_m^2(t)$ и $t(1-t)\overline{P}_{m-1}^2(t)$ соответственно.

Указанное соответствие, очевидно, распространяется на интервал $[0, a]$ при $a > 0$. Точнее, положим

$$\bar{q}_{m-1}(t; a) = \begin{vmatrix} c_1 - \dfrac{c_2}{a}, \ldots, c_{m-1} - \dfrac{c_m}{a}, 1 \\ c_2 - \dfrac{c_3}{a}, \ldots, c_m - \dfrac{c_{m+1}}{a}, t \\ \cdots\cdots\cdots\cdots \\ c_m - \dfrac{c_{m+1}}{a}, \ldots, c_{2m-2} - \dfrac{c_{2m-1}}{a}, t^{m-1} \end{vmatrix}.$$

Пусть $\sigma_a(t)$ — мера такая, что $\sigma_a(0-) = 0$, $\sigma_a(a) = 1$ и соответствующая ей моментная точка $(c_0, \ldots, c_{2m-1})$ содержится в $\operatorname{Int} M^{2m-1}[0, a]$. Тогда $\sigma_a(at) \equiv \sigma_1(t)$ — мера на $[0, 1]$ и $(c_0, c_1/a, \ldots, c_{2m-1}/a^{2m-1})$ принадлежит $M^{2m-1}[0, 1]$. Единственными опорными многочленами (см. страницу 54) к $M^{2m}[0, 1]$ в $(c_0, c_1 a^{-1}, \ldots, \underline{c}_{2m} a^{-2m})$ и $(c_0, c_1 a^{-1}, \ldots, \bar{c}_{2m} a^{-2m})$ с точностью до постоянного множителя являются

$$\begin{vmatrix} c_0 & \dots & c_{m-1}a^{-m+1} & 1 \\ c_1a^{-1} & \dots & c_m a^{-m} & t \\ \dots & \dots & \dots & \dots \\ c_m a^{-m} & \dots & c_{2m-1}a^{2m+1} & t^m \end{vmatrix}^2$$

и

$$t(1-t)\begin{vmatrix} c_1a^{-1} - c_2a^{-2} & \dots & c_{m-1}a^{-m+1} - c_m a^{-m} & 1 \\ c_2a^{-2} - c_3a^{-3} & \dots & c_m a^{-m} - c_{m+1}a^{-m-1} & t \\ \dots & \dots & \dots & \dots \\ c_m a^{-m} - c_{m+1}a^{-m-1} & \dots & c_{2m-2}a^{-2m+2} - c_{2m-1}a^{-2m+1} & t^{m-1} \end{vmatrix}$$

соответственно. Следовательно, опорные плоскости к $M^{2m}[0, a]$ в $\underline{c}$ и $\bar{c}$, полученные подстановкой $\tau = at$ и домножением на степень $a$, пропорциональны $\underline{P}_m^2(\tau)$ и $\tau(1-\tau/a)\,\bar{q}_{m-1}(\tau, a)$ соответственно. Конечно, использованная переменная несущественна для положения плоскостей. Постоянные множители существенны только для вычисления нормирующих постоянных.

Если $a \to \infty$, то опорные плоскости, индуцируемые $\underline{P}_m^2(t)$ и $t(1-t/a)\,\bar{q}_{m-1}^2(t; a)$ приближаются к плоскостям, соответствующим $\underline{P}_m^2(t)$ и $t\underline{Q}_{m-1}^2(t)$ (первая фактически не меняется). Следовательно, гомеоморфное отображение между $\operatorname{Int} M^{2m-1}[0, a]$, парами $(\underline{P}_m(t), \bar{q}_{m-1}(t; a))$ и открытым симплексом $0 < t_1 < < w_1 < \dots < w_{m-1} < t_m < a$ дает гомеоморфное отображение между $\operatorname{Int} M^{2m-1}[0, \infty)$, $(\underline{P}_m, \underline{Q}_{m-1})$ и открытым симплексом $0 < t_1 < s_1 < \dots < s_{m-1} < t_m < \infty$.

Полученные результаты можно сформулировать в виде следующего утверждения

Теорема 12.1. *Если дана точка* $\mathbf{c} = (c_0, \dots, c_{2m-1})$, *принадлежащая* $\operatorname{Int} M^{2m-1}[0, \infty)$, *то существует* $A > 0$ *такое, что при* $a > A$, $\mathbf{c} \in \operatorname{Int} M^{2m-1}[0, a]$; *причем точка* $\mathbf{c}$ *однозначно определяет*:

(i) *корни*

$$0 < t_1 < s_1 < \dots < s_{m-1} < t_m \leqslant A$$

*опорных плоскостей к* $M^{2m}[0, \infty)$ *в* $(c_0, \dots, c_{2m-1}, \underline{c}_{2m})$ *и* $(c_0, \dots c_{2m-2}, \underline{c}_{2m-1})$;

(ii) *для любого* $a > A$ *корни*

$$0 < t_1 < w_1(a) < \dots < w_{m-1}(a) < t_m \leqslant A$$

*опорных плоскостей к* $M^{2m}[0, a]$ *в* $(c_0, \dots, c_{2m-1}, \underline{c}_{2m})$ *и* $(c_0, \dots, c_{2m-1}, \bar{c}_{2m})$.

*Дополнительно имеем, что*

$$\lim_{a \to \infty} w_j(a) = s_j \qquad (j = 1, \dots, m-1).$$

## Глава VI

# МОМЕНТНЫЕ ПРОСТРАНСТВА ПЕРИОДИЧЕСКИХ ФУНКЦИЙ И $T$-СИСТЕМЫ НА $(-\infty, \infty)$

Понятие $T$-системы (см. 1.1 гл.I), согласно ее определению, связано с действительной прямой, вернее, с упорядоченным множеством. Мы довольно подробно рассмотрели случай интервала. Другим общим примером существенно упорядоченного множества является окружность, если зафиксирована некоторая точка в качестве начала отсчета. В последнем случае окружность эквивалентна прямой линии. Можно ожидать, что в этом случае большая часть теории допускает формулировки с необходимыми изменениями.

В этой главе мы рассмотрим два типа систем: $T$-системы периодических функций, которые на самом деле определены на окружности, и $T$-системы на двустороннем бесконечном интервале $(-\infty, \infty)$.

§§ 1—6 посвящены случаю периодических функций. Так как большая часть анализа близка к анализу, данному для непериодических функций на замкнутом интервале $[a, b]$, то мы ограничимся кратким изложением, а в некоторых случаях мы просто сформулируем результаты, не подкрепляя их формальными доказательствами.

В §§ 7—12 рассматривается случай $T$-систем на $(-\infty, \infty)$. Для этих типов систем мы наложим ограничения на поведение функции на $\pm\infty$, которые позволяют отождествить точки $+\infty$ и $-\infty$ и свести рассмотрение этих систем к периодическим функциям.

## § 1. Определения

Пусть $u_0(t)$, $u_1(t)$, ..., $u_n(t)$ — система непрерывных функций на конечном интервале $[a, b)$, удовлетворяющих условию $u_i(a) = u_i(b)$. Если мы распространим эти функции на всю прямую как периодические, то любой полуоткрытый интервал длины $b-a$ может служить нам областью определения. Таким образом, функции можно рассматривать как определенные на окружности длины $b-a$. Это рассуждение подсказывает нам, что все дальнейшие рассмотрения должны быть инвариантны относительно группы вращений, т. е. во всех рассмотрениях совокупность $t_0 < t_1 < \ldots$ $\ldots t_n < t_0 + b - a$ не должна отличаться от совокупности $t_k <$ $< t_{k+1} < \ldots < t_n < t_1 + b - a < \ldots < t_k + b - a$ $(k = 2, \ldots$ $\ldots, n)$. Другими словами, любой набор $\{t_i\}$ рассматривается на окружности, и никакое специальное расположение не считается начальным. Важно подчеркнуть, что когда $n$ нечетно, то нужно требовать сохранения знака определителя, который характеризует чебышевское свойство. Фактически, если $n = 2m + 1$ и $t_0 < t_1 < \ldots$ $\ldots < t_{2m+1} < t_0 + b - a$, то определители

$$U\begin{pmatrix} 0, 1, \ldots, 2m, 2m+1 \\ t_0, t_1, \ldots, t_{2m}, t_{2m+1} \end{pmatrix} \quad \text{и} \quad U\begin{pmatrix} 0, 1, \ldots, 2m, 2m+1 \\ t_1, t_2, \ldots, t_{2m+1}, t_0 \end{pmatrix}$$

имеют противоположные знаки. По этой причине мы предполагаем в дальнейшем, что $n = 2m$. Тогда то, что $\{u_i\}_0^{2m}$ является $T$-системой, можно записать так:

$$U\begin{pmatrix} 0, 1, \ldots, 2m \\ t_0, t_1, \ldots, t_{2m} \end{pmatrix} = \begin{vmatrix} u_0(t_0), & \ldots, u_0(t_{2m}) \\ u_1(t_0), & \ldots, u_1(t_{2m}) \\ \ldots & \ldots \\ u_{2m}(t_0), & \ldots, u_{2m}(t_{2m}) \end{vmatrix} > 0, \tag{1.1}$$

если только $t_0 < t_1 < \ldots < t_{2m} < t_0 + b - a$. Понятие $ET$-системы периодических функций порядка $k$ определяется обычным путем (см. § 2 гл. I). При циклической перестановке столбцов неравенство (1.1) остается неизменным.

Простым типичным примером $T$-системы периодических функций является система

$$1, \cos t, \sin t, \ldots, \cos mt, \sin mt \qquad (0 \leq t < 2\pi).$$

Известно, что в этом случае определитель (1.1) равен

$$U\begin{pmatrix} 0, 1, \ldots, 2m \\ t_0, t_1, \ldots, t_{2m} \end{pmatrix} = 4^m \prod_{j<i} \sin\left(\frac{t_i - t_j}{2}\right). \tag{1.2}$$

Это соотношение можно вывести, выразив $\cos kt$ и $\sin kt$ через экспоненциальные функции $e^{ikt}$ и $e^{-ikt}$ и используя значение определителя Вандермонда, приведенное в § 1 гл. I.

## § 2. Основные свойства периодических $T$-систем

Отметим несколько свойств периодических $T$-систем, доказательства которых получаются простым применением ранее проделанного анализа (см. гл. I, II).

(1) Число нулей любого $u$-многочлена четно и не более, чем $n = 2m$ при условии, что мы считаем узловые нули один раз, а неузловые дважды.

(2) Если даны $k+2l$ $(k+l \leq m)$ точек $t_1, t_2, \ldots, t_k, t_{k+1}, \ldots$ $\ldots, t_{k+2l}$ в $[a, b)$, то можно построить многочлен $u(t) = \sum_{i=0}^{2m} a_i u_i(t)$, имеющий неузловые нули в точках $t_1, \ldots, t_k$, узловые нули—в точках $t_{k+1}, \ldots, t_{k+2l}$ и не имеющий других нулей.

Мы построим моментное пространство $\mathcal{M}_{n+1} = \Big\{ \mathbf{c} = (c_0, c_1, \ldots$ $\ldots, c_{2m}) \mid c_i = \int\limits_{[a,b)} u_i(t)\, d\sigma(t),\ i = 0, 1, \ldots, 2m,\ \sigma$ — неотрицательная ограниченная мера$\Big\}$. Легко поверить, что $\mathcal{M}_{n+1}$ — замкнутый выпуклый конус. Пусть

$$c_k = \sum_{j=1}^{p} \rho_j u_k(t_j), \qquad k = 0, 1, \ldots, 2m, \tag{2.1}$$

где $\rho_j > 0$, $j = 1, 2, \ldots, p$, и $a \leq t_1 < t_2 < \ldots < t_p < t_1 + b - a$.

Определение 2.1. Точки $\{t_j\}_{j=1}^{p}$ называются *корнями представления* (2.1) *точки* $\mathbf{c}$, *а* $\rho_j$ — *их весами. Индексом* $I(\mathbf{c})$ *моментной точки* $\mathbf{c}$ называется минимальное число корней, которое требуется для представления $\mathbf{c}$ в виде (2.1).

(3) Точка $\mathbf{c}$ является граничной точкой $\mathcal{M}_{n+1}$ тогда и только тогда, когда $I(\mathbf{c}) \leq m$. Каждая граничная точка допускает единственное представление.

(4) Пусть $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$. Для каждого $t_0$ существует единственное каноническое представление индекса $m+1$, включающее $t_0$ как корень.

(5) Корни двух различных канонических представлений $\mathbf{c}$ строго перемежаются. Корни канонического представления являются непрерывными функциями любого из них. Вес, соответствующий любому корню $t_0$ канонического представления, является максимальной массой $\rho(t_0)$, доставляемой мерами, представляющими моментную точку $\mathbf{c}$. Более того, $\rho(t_0)$ — непрерывная функция $t_0$.

Значение максимальной массы $\rho(t_0)$ можно выразить в виде

$$\rho(t_0) = \min \int\limits_{[a,b)} u(t)\, d\sigma,$$

где $\sigma \in V(\mathbf{c})$, и минимум берется по всем неотрицательным многочленам $u$, нормированным так, что $u(t_0) = 1$.

## § 3. Двумерные сечения

Мы подчеркивали в начале § 1, что дополненная одной периодической функцией система $\{u_i\}_0^n$ не может сохранять чебышевское свойство (инвариантное при циклических перестановках), так как образованная система будет состоять уже из нечетного числа функций. Поэтому естественно увеличивать систему $\{u_i\}_0^n$ на пары периодических функций $v(t) = u_{n+1}(t)$ и $w(t) = u_{n+2}(t)$ так, чтобы система

$$u_0(t), u_1(t), \ldots, u_n(t), u_{n+1}(t), u_{n+2}(t) \qquad (3.1)$$

была тоже периодической $T$-системой на $[a, b)$.

Рассмотрим множество $A$ всех пар чисел $(x, y)$, допускающих представление

$$\begin{aligned} x &= \int v(t)\, d\sigma(t), \\ y &= \int w(t)\, d\sigma(t), \end{aligned} \qquad (\sigma \in V(c^0), c^0 \in \mathcal{M}_{n+1}). \qquad (3.2)$$

(В некотором смысле двумерная область в случае периодической $T$-системы является аналогом линейного сегмента $R$ $(c^0)$ § 6 гл. II).

Ясно, что $A$ — выпуклое множество, которое вырождается в единственную точку, когда $c^0 \in \mathcal{D}\mathcal{M}_{n+1}$. Мы сосредоточим теперь внимание на случае $c_0 \in \operatorname{Int} \mathcal{M}_{n+1}$.

Теорема 3.1 (Крейн [1951]). *Множество $A$ имеет внутренние точки. Пусть $c^0 \in \operatorname{Int} \mathcal{M}_{n+1}$. Точка $(x_0, y_0) \in A$ является граничной точкой $A$ тогда и только тогда, когда существует единственная мера $\sigma_0 \in V(c^0)$, образом которой при отображении* (3.2) *является $(x_0, y_0)$. Если $(x_0, y_0)$ — граничная точка $A$, то соответствующая мера $\sigma_0$ является канонической.*

Замечание 3.1. Обсуждаемые в этом параграфе вопросы, относящиеся к двумерным сечениям, были впервые рассмотрены Крейном [1951], им же было найдено соответствие между границей и множеством канонических представлений, описанным выше. Теорема 3.2, приведенная ниже, в формулировке которой каждой внутренней точке двумерного сечения ставится в соответствие вес и корень, является новой.

Доказательство. Рассмотрим моментное пространство $\mathcal{M}_{n+3}$, порожденное системой (3.1). Пусть $\sigma_0$ — каноническое представление $c^0 \in \mathcal{M}_{n+1}$. По определению оно имеет индекс, не больший чем $m + 1$. Обращаясь к свойству (3) § 2, мы приходим к выводу, что $\sigma_0$ порождает граничную точку $\mathcal{M}_{n+3}$. В частности, $(x_0, y_0)$ — граничная точка сечения $A$. Обратно, граничная точка $(x_0, y_0)$ множества $A$ является, очевидно, граничной точкой $\mathcal{M}_{n+3}$. Поэтому $(c_0^0, c_1^0, \ldots, c_{2m}^0, x_0, y_0)$ допускает единственное представление $\sigma_0$ индекса, не большего чем $m + 1$. Индекс не может быть меньшим или равным $m$, так как в противном случае нарушилось бы условие $c^0 \in \operatorname{Int} \mathcal{M}_{n+1}$. Поэтому $\sigma_0$ — каноническое представление.

Следствие 3.1. *На границе множества $A$ не существует линейных сегментов.*

Теорема 3.1 показывает, что канонические представления $c^0$ находятся с границей $A$ во взаимно однозначном соответствии. Пусть

$\sigma_{t^*}$ — каноническая мера точки $\mathbf{c}^0$, включающая $t^*$, и пусть $t^*$ и $t^{**}$ — два последовательных корня $\sigma_{t^*}$. Каждое $t' \in [t^*, t^{**})$ индуцирует единственное каноническое представление $\sigma_{t'}$. Поэтому мы имеем взаимно однозначное и взаимно непрерывное отображение интервала $[t^*, t^{**})$ на границу $A$.

Чтобы получить представление внутренней части $A$, мы можем поступить, как в § 7 гл. II. Пусть $\rho(t')$ — масса в точке $t' \in (t^*, t^{**})$. Если положить $\alpha = 0$ в равенстве

$$(\tilde{c}_0, \ldots, \tilde{c}_{2m}) = (c_0^0, \ldots, c_{2m}^0) - (1 - \alpha)\,\rho(t')\,(u_0(t'), \ldots, u_{2m}(t')),$$

то образованная точка является граничной точкой $\mathcal{M}_{2n+1}$, так как $(\tilde{c}_0, \ldots, \tilde{c}_{2m})$ представляется мерой $\sigma_{t'}$ с массой, удаленной из точки $t'$. Точка $\tilde{\mathbf{c}} = (\tilde{c}_0, \ldots, \tilde{c}_{2m})$ содержится внутри $\mathcal{M}_{2n+1}$ для всех $\alpha \in (0, 1)$. Следовательно, существует каноническое представление $\tilde{\mathbf{c}}$, включающее $t^*$ и $m$ других различных точек. Ни одна из этих точек не может равняться $t'$, так как в этом случае получаемое представление для $\mathbf{c}^0$ имело бы индекс $m + 1$ и включало бы $t^*$ и $t'$. Однако это невозможно в силу выбора $t'$. Таким образом, мы показали, что для каждого $t' \in (t^*, t^{**})$ существует представление $\mathbf{c}$ индекса $m + 2$, включающее $t'$ и $t^*$ с произвольным весом $\lambda \in (0, \rho(t'))$ в $t'$.

Мы обозначим эту специальную меру через $\sigma(t, t', \lambda)$ и рассмотрим отображение области

$$B = \{(t', \lambda) \mid t^* < t' < t^{**} \text{ и } 0 < \lambda < \rho(t')\}$$

во внутреннюю часть $A$, определенное равенством

$$\varphi(t', \lambda) = (c_{2m+1}, c_{2m+2}), \tag{3.3}$$

где

$$c_i = \int_a^b u_i(t)\, d\sigma(t; t', \lambda), \quad i = 2m + 1, 2m + 2.$$

Теорема 3.2. *Если* $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{2m+1}$, $t^*$ *и* $t^{**}$ — *два последовательных корня* $\sigma_{t^*}$ *(где* $\sigma_{t^*}$ — *каноническая мера для* $\mathbf{c}^0$*) и* $t' \in (t^*, t^{**})$, *то существует представление* $\mathbf{c}^0$ *индекса* $m + 2$, *включающее* $t^*$ *и* $t'$ *с весом* $\lambda \in (0, \rho(t'))$ *в точке* $t'$. *Это представление индуцирует отображение* $\varphi$ *области* $B$ *на внутреннюю часть* $A$, *которое является взаимно однозначным и непрерывным.*

Свойства отображения $\varphi$ можно установить, повторяя дословно доказательство теоремы 7.1 гл. II, детали опускаем.

Отображение $\varphi$ также является непрерывным на границе $B$ в том смысле, что при стремлении $\lambda$ к $\rho(t')$ мера $\sigma(t, t', \lambda)$ стремится к канонической мере $\sigma_{t'}$ точки $(c_0^0, \ldots, c_{2m}^0)$, следующие два момента которой определяют граничную точку $A$. Однако отображение границы не является взаимно однозначным. Фактически меры

$\sigma(t; t', \rho(t')) = \sigma_{t'}(t)$ при $t'$, меняющемся в $(t^*, t^{**})$, отображаются соотношением (3.2) в границу $A$ взаимно однозначно и непрерывно, исключая точку $(c^*_{2m+1}, c^*_{2m+2})$, которая соответствует $\sigma_{t^*}$. Если $t = t^*$ или $t = t^{**}$, либо $\lambda = 0$, то соответствующие меры $\sigma(t; t', \lambda)$ переходят в $\sigma_{t^*}(t)$.

Если корни $t', t'_1, \ldots, t'_m$ и $t^*, t^*_1, \ldots, t^*_m$ канонических мер $\sigma_{t'}$ и $\sigma_{t^*}$ расположены циклически, так, что

$$t^* < t' < t^*_1 < t'_1 < \ldots < t^*_m < t'_m < t^* + b - a,$$

то корни $t^*, t', t_1, \ldots, t_m$ меры $\sigma(t; t', \lambda)$ таковы, что $t^*_i < t_i < t'_i$, $i = 1, 2, \ldots, m$, и при возрастании $\lambda$ от 0 до $\rho(t')$ корень $t_i$ изменяется непрерывно от $t^*_i$ до $t'_i$.

## § 4. Тригонометрическая проблема моментов

В этом параграфе мы докажем следующий классический результат.

Теорема 4.1. *Для $T$-системы* $1, \cos t, \sin t, \ldots, \cos mt, \sin mt$ *на* $[0, 2\pi)$ *точка* $\mathbf{c} = (\alpha_0, \alpha_1, \beta_1, \ldots, \alpha_m, \beta_m)$ *является моментной точкой* $\mathcal{M}_{2m+1}$ *тогда и только тогда, когда*

$$\sum_{p,q=0}^{m} d_p \bar{d}_q \gamma_{p-q} \geqslant 0 \tag{4.1}$$

*для любого набора* $d_0, d_1, \ldots, d_m$, *где* $\gamma_0 = \alpha_0 > 0$, $\gamma_k = \alpha_k - i\beta_k$ $(k = 1, 2, \ldots, m)$ *и* $\gamma_{-k} = \bar{\gamma}_k$. *Более того,* $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$ *тогда и только тогда, когда квадратичная форма* (4.1) *положительно определена.*

Доказательство. Результат, приведенный в лемме 9.2 гл II, переносится на случаи периодических функций, так что $\mathbf{c} \in \mathcal{M}_{n+1}$ тогда и только тогда, когда

$$a_0 + \sum_{k=1}^{m} (a_k \cos kt + b_k \sin kt) \to a_0\alpha_0 + \sum_{k=1}^{m} (a_k\alpha_k + b_k\beta_k), \tag{4.2}$$

а обе части этого неравенства неотрицательны; и $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$ тогда и только тогда, когда справа стоит строгое неравенство при условии, что слева в (4.2) стоит нетривиальный многочлен. Теперь неравенство $a_0 + \sum\limits_{k=1}^{m} a_k \cos kt + b_k \sin kt \geqslant 0$ можно переписать в виде

$$\frac{1}{2} \sum_{k=0}^{m} (\xi_k e^{ikt} + \bar{\xi}_k e^{-ikt}) \geqslant 0, \tag{4.3}$$

где $\xi_k = a_k - ib_k$, $k = 0, 1, \ldots, m, b_0 = 0$.

Мы покажем, что класс многочленов, удовлетворяющих условию (4.3), совпадает с классом многочленов, допускающих представление

$$\left| \sum_{k=0}^{m} d_k e^{ikt} \right|^2,$$

где $d_0, d_1, \ldots, d_m$ — произвольные комплексные числа.

Очевидно, любой многочлен вида (4.4) удовлетворяет (4.3). Для того чтобы установить обратное, мы сначала запишем многочлен $\varphi(t)$, удовлетворяющий (4.3), в виде

$$\varphi(t) = |z^{-m} F(z)|, \quad z = e^{it},$$

где

$$F(z) = \frac{1}{2} \sum_{k=0}^{m} (\xi_k z^{m+k} + \bar{\xi}_k z^{m-k}).$$

Многочлен $F(z)$ имеет $2m$ комплексных корней. Для каждого корня $z_0$ $(0 < |z_0| < 1)$ $1/\bar{z}_0$ также является корнем и обратно, что видно из равенства $F(\bar{z}_0^{-1}) = \overline{z_0^{-2m} F(z_0)}$. К тому же каждый корень $e^{i\alpha}$ обладает четной кратностью, так как $\varphi(t) \geqslant 0$.

Поэтому

$$|F(z)| = c\,|z^p| \left| \prod_{j=1}^{l} (z - z_j)(z - 1/\bar{z}_j) \prod_{r=1}^{k} (z - e^{i\alpha_r})^2 \right|,$$

где $c > 0$ и $p + 2l + 2k = 2m$. Заменяя $z$ на $e^{it}$, мы видим, что многочлен $\varphi(t)$ очевидным образом можно выразить в виде (4.4).

Заключение теоремы для $\mathbf{c} \in \mathcal{M}_{n+1}$ и $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$ немедленно следует, если применить (4.2) к многочленам вида (4.4).

Во всех источниках теорема 4.1, которая выражает классический критерий для проверки того, является ли последовательность моментной последовательностью тригонометрической проблемы моментов, используется теорема Файера — Рисса (см. (4.3) и (4.4) о представлении положительных многочленов) (см. Шохат и Тамаркин [1943], гл. II).

## § 5. Примеры

Следующие ниже рассуждения принадлежат Крейну [1951] (см. также Геронимус [1958] и Ахиезер [1965], гл. 5). Для второго примера этого параграфа истоком послужили работы Каратеодори, Герглотца, Пика и др.

(1) Мы сосредоточим внимание на системе функций

$$1, \cos t, \sin t, \ldots, \cos mt, \sin mt \qquad (0 \leq t < 2\pi). \tag{5.1}$$

По теореме 4.1 множество постоянных

$$\alpha_0, \alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_m, \beta_m \tag{5.2}$$

определяет моментную точку относительно системы (5.1) тогда и только тогда, когда

$$\sum_{p,q=0}^{m} d_p \bar{d}_q \gamma_{q-p} \geqslant 0 \tag{5.3}$$

при всяком выборе комплексных чисел $d_0, d_1, \ldots, d_m$, где $\gamma_0 = \alpha_0 > 0$, $\gamma_k = \alpha_k - i\beta_k$ $(k = 1, \ldots, m)$ и $\gamma_{-k} = \bar{\gamma}_k$. Точка (5.2) является внутренней точкой моментного пространства тогда и только тогда, когда упомянутая квадратичная форма положительно определена. Известно, что последнее утверждение эквивалентно последовательности неравенств

$$\Delta_k = \begin{vmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{k-1} & \gamma_k \\ \gamma_{-1} & \gamma_0 & \cdots & \gamma_{k-2} & \gamma_{k-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_{-k} & \gamma_{-(k-1)} & \cdots & \gamma_{-1} & \gamma_0 \end{vmatrix} > 0; \qquad k = 0, 1, \ldots, m. \tag{5.4}$$

Пусть (5.2) определяет внутреннюю точку моментного пространства $\mathcal{M}_{2m+1}$, т. е. пусть имеет место (5.4). Мы исследуем область $A$, определенную в (3.2), соответствующую функциям

$$v(t) = \cos(m+1)t, \quad w(t) = \sin(m+1)t.$$

Точке $\eta = x - iy$ соответствует $(x, y) \in A$ тогда и только тогда, когда

$$\bar{\eta} = \int e^{i(m+1)t}\, d\sigma(t) \tag{5.5}$$

для некоторой $\sigma \in V(\alpha_0, \alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_m, \beta_m)$. Принимая во внимание критерий (5.4), мы видим, что $\eta$ является граничной точкой $A$ тогда и только тогда, когда имеет место (5.4) и

$$\begin{vmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_m & \eta \\ \gamma_{-1} & \gamma_0 & \cdots & \gamma_{m-1} & \gamma_m \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \bar{\eta} & \gamma_{-m} & \cdots & \gamma_{-1} & \gamma_0 \end{vmatrix} = 0. \tag{5.6}$$

Разлагая (5.6) с помощью формулы для определителя Сильвестра (см. сноску на стр. 43), где основной блок определяется вычеркиванием первой и последней строк и первого и последнего столбцов, мы получим

$$0=\begin{vmatrix} \Delta_m & R_m(\eta) \\ \overline{R_m(\eta)} & \Delta_m \end{vmatrix}, \tag{5.7}$$

где $R_m(\eta)=\begin{vmatrix} \gamma_1 & \dots & \gamma_m & \eta \\ \gamma_0 & \dots & \gamma_{m-1} & \gamma_m \\ \dots & \dots & \dots & \dots \\ \gamma_{-m+1} & \dots & \gamma_0 & \gamma_1 \end{vmatrix}$.

Теперь, очевидно, $R_m(\eta)=(-1)^m\eta\Delta_{m-1}+C_m$, где

$$C_m=\begin{vmatrix} \gamma_1 & \dots & \gamma_m & 0 \\ \gamma_0 & \dots & \gamma_{m-1} & \gamma_m \\ \dots & \dots & \dots & \dots \\ \gamma_{-m+1} & \dots & \gamma_0 & \gamma_1 \end{vmatrix}.$$

Поэтому уравнение границы имеет вид

$$|\Delta_m|^2=|(-1)^m\eta\Delta_{m-1}+C_m|^2,$$

или, что то же самое,

$$\left|\eta+(-1)^m\frac{C_m}{\Delta_{m-1}}\right|^2=\left|\frac{\Delta_m}{\Delta_{m-1}}\right|^2. \tag{5.8}$$

Область $A$ (смотри (3.2)) является, очевидно, кругом с центром в $[((-1)^{m+1}C_m)/\Delta_{m-1}]$ и радиусом $\Delta_m/\Delta_{m-1}$. Используя очевидные обозначения, рассмотрим граничную точку $(\gamma_0, \gamma_1, \dots, \gamma_m, \eta_0)$ в $\mathcal{M}_{n+3}$. Мы утверждаем, что корни единственного канонического представления, соответствующего моментной точке

$$(\gamma_0, \dots, \gamma_m, \eta_0) \tag{5.9}$$

совпадают с корнями многочлена

$$R(e^{it})=\begin{vmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_m & 1 \\ \gamma_{-1} & \gamma_0 & \dots & \gamma_{m-1} & e^{it} \\ \dots & \dots & \dots & \dots & \dots \\ \gamma_{-m} & \gamma_{-m+1} & \dots & \gamma_0 & e^{imt} \\ \bar{\eta}_0 & \gamma_{-m} & \dots & \gamma_{-1} & e^{i(m+1)t} \end{vmatrix}. \tag{5.10}$$

Этот многочлен нетривиален, так как $\Delta_m>0$.

Доказательство этого утверждения состоит в следующем. Если $(\gamma_0, \dots, \gamma_n)\in \operatorname{Int}\mathcal{M}_{n+1}$ и (5.9) — граничная точка, то справедливы соотношения (5.4) и (5.6). При помощи интегрирования убеждаемся, что $R(e^{it})$ ортогональна функциям

$$1,\ e^{-it}, \dots, e^{-i(m+1)t} \tag{5.11}$$

относительно единственного канонического представления $\sigma$ точки $(\gamma_0, \ldots, \gamma_m, \eta_0)$. Ортогональность $R(e^{it})$ функции $e^{-i(m+1)t}$ следует из (5.6). Поэтому

$$\int_0^{2\pi} |R(e^{it})|^2 d\sigma(t) = 0, \tag{5 12}$$

так что $R(e^{it})$ должно обращаться в нуль как раз в корнях $\sigma$. Разлагая (5.10), мы получим, что

$$R(e^{it}) = P_1(e^{it}) + \bar{\eta}_0 P_2(e^{it}). \tag{5.13}$$

Заметим, что $R(e^{it})$ зависит линейно от $\bar{\eta}_0$.

(2) Исследуем область $A$ для системы функций (5.1) при другом выборе функций $v(t)$ и $w(t)$, а именно:

$$v(t) = \operatorname{Re}\frac{z_0 + e^{it}}{z_0 - e^{it}}, \quad w(t) = \operatorname{Im}\frac{z_0 + e^{it}}{z_0 - e^{it}}, \tag{5.14}$$

где $z_0$ фиксированная точка в единичном круге $0 < |z_0| < 1$. Мы утверждаем, что

$$1, \cos t, \sin t, \ldots, \cos mt, \sin mt, v(t), w(t)$$

является $T$-системой при подходящем выборе знака $w(t)$[1]. Чтобы установить этот факт, достаточно доказать, что уравнение

$$0 = \sum_{k=-m}^{m} a_k e^{kit} + a\frac{z_0 + e^{it}}{z_0 - e^{it}} + \bar{a}\frac{\bar{z}_0 + e^{-it}}{\bar{z}_0 - e^{-it}} \tag{5.15}$$

$(\bar{a}_k = a_{-k}, k = 0, 1, \ldots, m)$ имеет не более $2m + 2$ нулей. Этот результат получится, если записать многочлен в правой части (5.15) при помощи комплексной переменной $z = e^{it}$, Действительно, подставляя $e^{it} = z$ и упрощая (5.15), приходим к равенству

$$0 = \frac{P_{2m+2}(z)}{z^m (z_0 - z)(z\bar{z}_0 - 1)},$$

и желаемый результат теперь очевиден.

Пусть $\gamma = (\gamma_0, \ldots, \gamma_m)$ — внутренняя точка моментного пространства $\mathcal{M}_{n+1}$, порожденного $T$-системой (5.1). Множество всех канонических мер $\mathcal{K} \in V(\gamma)$ может быть поставлено в соответствие набору всех многочленов вида (5.10) ($\eta = \eta_0$ — единственный свободный параметр) в том смысле, что корни этого многочлена совпадают с корнями соответствующей канонической меры.

---

[1]) См. сноску на стр. 33.

Граница множества $A$ представляет собой множество точек

$$BdA: z = x + iy = \int_0^{2\pi} \frac{z_0 + e^{it}}{z_0 - e^{it}} d\sigma(t),$$

где $\sigma$ пробегает $\mathcal{K}$.

Положим

$$S_\eta(z) = \int \frac{z + e^{it}}{z - e^{it}} [R_\eta(z) - R_\eta(e^{it})] d\sigma(t), \tag{5.16}$$

где $R_\eta$ — многочлен, определенный в (5.10).

Рассмотрим каноническое представление $\sigma_\eta(t)$, соответствующее $\eta$. Многочлен $R_\eta$ обращается в нуль на спектре $\sigma_\eta(t)$. Тогда (5.16) принимает вид

$$\int \frac{z + e^{it}}{z - e^{it}} d\sigma_\eta(t) = \frac{S_\eta(z)}{R_\eta(z)}. \tag{5.17}$$

Из (5.13) видно, что $R_\eta(e^{it})$ — линейная функция $\eta$. Так как $S_\eta$ не зависит от $\sigma \in V(\gamma)$, т. е. $S_\eta$ зависит только от $\eta$ и $\gamma$ и не зависит от выбора $\sigma \in V(\gamma)$, то отсюда следует, что

$$S_\eta(z) = S_1(z) + \bar{\eta} S_2(z),$$

где $S_1(z)$ и $S_2(z)$ — фиксированные многочлены. Подстановка в (5.17) приводит к выражению

$$\int \frac{z_0 + e^{it}}{z_0 - e^{it}} d\sigma_\eta(t) = \frac{S_1(z_0) + \bar{\eta} S_2(z_0)}{P_1(z_0) + \bar{\eta} P_2(z_0)}. \tag{5.18}$$

Из того, что (5.18) является дробно-линейным преобразованием, мы заключаем, что когда $\eta$ пробегает границу круга (5.8), граница $A$ описывает круг $A_{z_0}(\gamma_0, \gamma_1, \ldots, \gamma_m)$.

Мы завершаем этот параграф применением полученных результатов к анализу предыдущих рассуждений (принадлежащих Пику).

Множество $\Gamma$ всех функций вида

$$F(z) = \int_0^{2\pi} \frac{e^{it} + z}{e^{it} - z} d\sigma(t)$$

($|z| < 1$, $\sigma$ — неотрицательная ограниченная мера) можно охарактеризовать как множество аналитических функций ($|z| < 1$), отображающих единичный круг в правую полуплоскость, таких, что $F(0)$ действительно. Если $\sigma \in V(\alpha_0, \alpha_1, \beta_1, \ldots, \alpha_m, \beta_m)$, то $F(z)$ принадлежит множеству функций, для которых числа $\gamma_0, 2\gamma_1, \ldots, 2\gamma_m$ являются первыми $m + 1$ коэффициентами их разложения в степенные ряды. Таким образом, в предыдущих рассуждениях мы установили следующее предложение.

*Множество значений в данной точке $z_0$ $(0 < |z_0| < 1)$ всех $F \in \Gamma$, первые $m$ коэффициентов которых $\gamma_0, 2\gamma_1, \ldots, 2\gamma_m$ $(\gamma_0 > 0)$ фиксированы, заполняет круг $A_{z_0}$ $(\gamma_0, \gamma_1, \ldots, \gamma_m)$.*

## § 6. Представление для положительных многочленов и крайние точки $\mathcal{M}_{2m+1}$

В этом параграфе мы распространим на случай периодической $T$-системы на $[a, b)$ соответствующие теоремы о представлении и конструкции § 10 гл. II. Напомним, что в случае непериодических $T$-систем $(n = 2m)$ мы определили для каждой положительной непрерывной функции $f(t)$ два неотрицательных многочлена $\bar{u}(t)$ и $\underline{u}(t)$, один из которых обращался в нуль в $m-1$ внутренних точках и на обоих концах, а другой — в $m$ внутренних точках. Эти многочлены изменялись от 0 до $f(t)$ определенным образом. В периодическом случае нет естественной концевой точки для области определения функций $\{u_i\}_0^{2m}$. Действительно, любая точка $t^* \in [a, b)$ может быть выбрана в качестве одного из концов интервала, и мы можем рассматривать функции, определенные на интервале $[t^*, t^* + b - a)$. С учетом этого вариант теоремы 10.1 гл. II для периодического случая допускает следующее выражение. Пусть $f(t)$ — непрерывная периодическая функция периода $b - a$. Для любого фиксированного $t^* \in [a, b)$ мы построим периодический многочлен $u^*(t; t^*)$, изменяющийся между нулем и $f(t)$ с корнем $t^*$ и имеющий колебательный характер, как у функции $\bar{u}(t)$ из теоремы 10.1 гл. II, где областью определения является $[t^*, t^* + b - a)$. Точная формулировка утверждения дается ниже в теореме 6.1.

В этом параграфе также будут рассмотрены свойства перемежаемости корней для семейства многочленов $u^*(t; t^*)$, для которых $t^*$ рассматривается как параметр. Конструкция $u^*(t; t^*)$ играет важную роль в решении экстремальных проблем Маркова— Бернштейна, которые будут рассмотрены в гл. IX.

Теорема 6.1 (Карлин [1963]). *Если $\{u_i\}_0^{2m}$ — $T$-система периодических функций на $[a, b)$ и $f(t)$ — периодическая непрерывная положительная функция, то для каждого $t^* \in [a, b)$ существует многочлен $u^*(t; t^*)$, который удовлетворяет условиям:*

(i) $f(t) \geq u^*(t; t^*) \geq 0$ *для* $t \in [a, b)$;

(ii) $u^*(t; t^*)$ *имеет $m$ различных нулей, в том числе и* $t^*$;

(iii) $f(t) - u^*(t; t^*)$ *обращается в нуль по крайней мере один раз между каждой парой (рассматриваемой в циклическом смысле) различных нулей* $u^*(t; t^*)$.

*Более того, многочлен $u^*(t; t^*)$ со свойствами* (i) — (iii) *единствен, если $\{u_i\}_0^{2m}$ — $ET$-система порядка* 2.

Многочлен $u^*(t; t^*)$ проиллюстрирован на рис. 9 для случая $n = 2m = 6$. Пунктирная линия обозначает многочлен $u^{**}(t; t^*) = f(t) - u^*(t; t^*)$ при условии, что $f(t)$ — многочлен (см. следствие 6.1).

Доказательство. Построение многочлена $u^*(t;\ t^*)$ следует построению теоремы 10.1 гл. II.

Сначала мы предположим, что $\{u_i(t)\}_0^n$ — $ET$-система и для каждого множества $\xi = \{t_1, \ldots, t_m\}$ из $m$ точек (не обязательно различных), которое всегда включает $t^*$, построим неотрицательный многочлен $u_\xi(t)$, имеющий эти точки нулями соответствующей кратности. На каждом из $m$ «интервалов» между парами точек мы определим $\alpha_i$, $i = 1, 2, \ldots, m$, как и ранее. Применение теоремы Брауэра о неподвижной точке доказывает, что для некоторого расположения

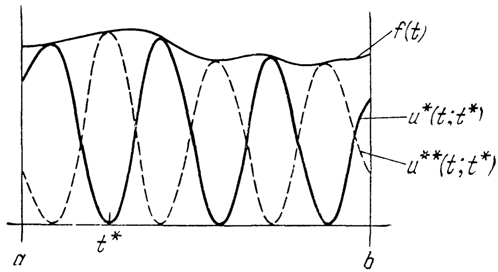


Рис. 9

$m$ различных точек, соответствующего, скажем, $\xi$, справедливо равенство $\alpha_i = \alpha > 0$ для $i = 1, 2, \ldots, m$. Требуемым многочленом является теперь $u^*(t;\ t^*) = \alpha^{-1}u_\xi(t)$.

Распространение результата в случае чебышевских предположений можно сделать так же, как и в теореме 10.1 гл. II. Доказательство единственности также повторяет доказательство теоремы 10.1 гл. II и будет опущено.

Следствие 6.1. *Если $u(t)$ — положительный многочлен, то для каждого $t^* \in [a, b)$ существует единственное представление*

$$u(t) = u^*(t;\ t^*) + u^{**}(t;\ t^*),$$

*где $t^*$ является нулем $u^*(t;\ t^*)$ и где как $u^*(t;\ t^*)$ так и $u^{**}(t;\ t^*)$ — неотрицательные многочлены, имеющие по $m$ корней, которые строго перемежаются. (Для единственности мы предполагаем, что $\{u_i(t)\}_0^n$ — $ET$-система порядка 2.)*

Элементарным примером к этому следствию для $T$-системы $1$, $\cos t$, $\sin t$, $\ldots$, $\cos mt$, $\sin mt$ на интервале $[0,\ 2\pi)$ при выборе $u(t) \equiv 1$ является следующий:

$$1 = \begin{cases} \sin^2 k(t - t^*) + \cos^2 k(t - t^*), & m = 2k, \\ \sin^2\left(k + \frac{1}{2}\right)(t - t^*) + \cos^2\left(k + \frac{1}{2}\right)(t - t^*), & m = 2k + 1. \end{cases}$$

Здесь $u^*(t;\ t^*)$ равно $\sin^2 k(t - t^*)$ или $\sin^2\left(k + \frac{1}{2}\right)(t - t^*)$. Мы теперь покажем, что нули двух различных многочленов $u^*(t;\ t^*)$ и

$u^*(t;\ t^0)$ строго перемежаются, как это отражено на рис. 10. Более того, каждый из нулей $u^*(t;\ t^*)$ является монотонной непрерывной функцией от $t^*$.

Пусть $t_1^0, t_2^0, \ldots, t_m^0$ — нули фиксированного многочлена, построенного, как в теореме 6.1, расположенные против часовой стрелки. Пусть $t_i(t^*)$, $i = 1, 2, \ldots, m$ $(t_1(t^*) = t^*)$, обозначает нули многочлена $u^*(t;\ t^*)$, обращающегося в нуль в $t^*$. Предположим, что эти нули также расположены в циклическом порядке. Тогда имеет место следующее свойство непрерывности нулей.

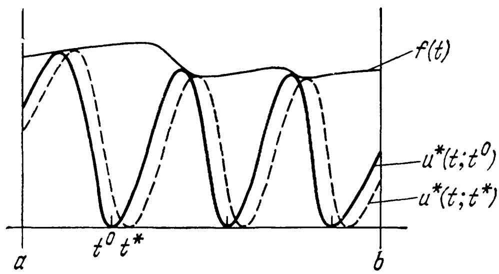



Рис 10.

Лемма 6.1. *Если многочлен $u^*(t;\ t^*)$ — единствен (например, если $\{u_i\}_0^{2m}$ — ET-система порядка 2), то $t_i(t^*)$, $i = 1, 2, \ldots, m$,— непрерывные монотонные функции $t^*$ на $[t_1^0, t_2^0]$, удовлетворяющие условиям $t_i(t_1^0) = t_i^0$ и $t_i(t_2^0) = t_{i+1}^0$, $i = 1, \ldots, m$ $(t_{m+1}^0 = t_1^0)$.*

Замечание 6.1. *Если функция $f(t)$ постоянная, то нули $t_i(t^*)$, $i = 1, 2, \ldots, m$, можно получить простым сдвигом, т. е. если $f(t) \equiv 1$, то $u^*(t;\ s^*) = u^*(t - s^* + t^*;\ t^*)$.*

Доказательство. Из неравенства можно заключить, что любые два различных многочлена $u^*(t;\ t^*)$ имеют различные множества нулей. Пусть $t_1, \ldots, t_m$ и $s_1, \ldots, s_m$ — нули двух различных многочленов $u^*(t;\ t_1)$ и $u^*(t;\ s_1)$; расположим каждое из этих множеств в циклическом порядке. Мы утверждаем, что никакие два различных нуля $s_i$ и $s_{i+1}$ одного из многочленов не могут лежать между двумя нулями $t_j$ и $t_{j+1}$ — другого. Чтобы проверить это, мы предположим противное, т. е. предположим, что для фиксированных $i$ и $j$ $t_j \leq s_i < s_{i+1} \leq t_{j+1}$. Если $s_i^* \in (s_i, s_{i+1})$ таково, что $f(s_i^*) - u^*(s_i^*, s_1) = 0$, то разность двух рассматриваемых многочленсв имеет по крайней мере два нуля в каждом из интервалов $[t_j, s_i^*]$, $[s_i^*, t_{j+1}]$ и $[t_k, t_{k+1}]$ для $k \neq j$, при этом общее число нулей становится равным $2(m-1) + 4 = 2m + 2$. Это возможно только тогда, когда многочлены совпадают. Теперь, если последовательность величин $t^*$, скажем $t_{(k)}^*$, приближается к $t^*$ против часовой стрелки, то $t_i(t_{(k)}^*)$ должно приближаться к $t_i(t^*)$ с той же стороны. В противном случае свойство строгой перемежаемости нулей было бы нарушено.

Теорема 6.2 (Карлин [1963]). *Пусть* $\{u_i\}_0^{2m}$ — *$T$-система периодических функций на* $[a, b)$, *и пусть* $f(t)$ *и* $g(t)$ — *две непрерывные периодические функции на* $[a, b)$ *такие, что существует многочлен* $v(t)$, *удовлетворяющий условию* $f(t) > v(t) > g(t)$; *тогда справедливы следующие утверждения*:

(a) *Для каждого* $t_0 \in [a, b)$ *существует единственный многочлен* $u(t; t_0)$, *удовлетворяющий условиям*:

(i) $f(t) \geqslant u(t; t_0) \geqslant g(t)$ и $v(t_0) - u(t_0; t_0) = 0$;

(ii) *существует* $2m$ *точек* $s_1, s_2, \ldots, s_{2m}$ *таких, что* $t_0 < s_1 < s_2 < \ldots < s_{2m} < t_0 + b - a$ *и*

$$u(s_i; t_0) = \begin{cases} f(s_i), & \text{если } i \text{ — нечетное}, \\ g(s_i), & \text{если } i \text{ — четное}. \end{cases} \tag{6.1}$$

(b) *Пусть условие* (ii) *заменено на условие* (ii)$'$, *которое получается из* 6.1, *если поменять местами функции* $f$ *и* $g$. *Тогда существует единственный многочлен, удовлетворяющий* (i) *и* (ii)$'$.

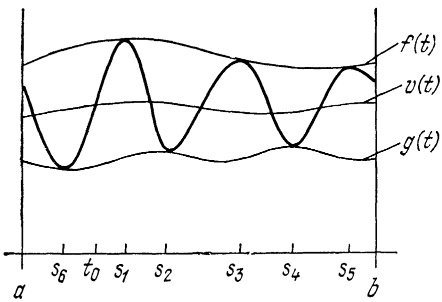


Рис 11.

Доказательство. Построение требуемых многочленов проводится как в теореме 10.2 гл. II, и поэтому мы его опускаем. На рис. 11 мы приводим для иллюстрации в случае $n = 6$ многочлен $u(t; t_0)$ из теоремы 6.2.

Следующая теорема уточняет теорему 6.1. Она может быть получена как видоизменение теоремы 10.3 гл. II для случая периодических $T$-систем. Доказательство проводится так же, как и в теореме 10.3 гл. II, и мы его тоже опускаем.

Теорема 6.3 (Карлин [1963]). *Пусть* $\{u_i\}_0^{2m}$ — *$ET$-система периодических на* $[a, b)$ *функций, и пусть* $f(t) \in C^{2m}$ — *неотрицательная периодическая функция, имеющая* $2r$ $(r < m)$ *нулей с учетом кратности. Тогда для каждой точки* $t^* \in [a, b)$ *существует единственный многочлен* $u^*(t; t^*)$, *обладающий свойствами*:

(i) $f(t) \geqslant u^*(t; t^*) \geqslant 0$;

(ii) $u^*(t; t^*)$ *имеет* $2m - 2r$ *нулей, отличных от нулей* $f(t)$, *являющихся нулями кратности* 2 *в* $m - r$ *различных точках, скажем* $t_1, \ldots, t_{m-r}$, *одна из которых* $t^*$;

(iii) $f(t) - u^*(t; t^*)$ *имеет по крайней мере в два раза больше нулей, чем* $f(t)$ *в каждом из* $m - r$ *открытых интервалах между соседними* $t_i$.

Следствие 6.3. *Если, дополнительно к предположениям теоремы* 6.3, $f$ — *многочлен, то для каждого* $t^*$ *существует един-*

*ственное представление*

$$f(t) = u^*(t;\ t^*) + u^{**}(t;\ t^*),$$

*где* $u(t^*;\ t^*) = 0$, *а* $u^*$ *и* $u^{**}$*—неотрицательные многочлены, каждый из которых имеет* $2m$ *нулей с учетом кратности. Эти нули строго перемежаются, если пренебречь нулями* $f(t)$.

**Крайние точки и граница** $\mathcal{P}_{2m+1}$. Если $\{u_i\}_0^{2m}$ — $T$-система периодических функций на $[a, b)$, то границу и крайние точки $\mathcal{P}_{2m+1}$ можно охарактеризовать точно так же, как если бы $u_i(t)$, $i=0, 1, \ldots, 2m$, были обычными периодическими функциями. Граница состоит из тех неотрицательных многочленов, которые обращаются в нуль при некотором $t \in [a, b)$, и теоремы 11.1 и 11.2 гл. II также имеют место. Теорема 11.2 может быть переформулирована в следующем виде.

Теорема 6.4. *Если* $\{u_i\}_0^{2m}$ — *ET-система периодических функций на* $[a, b)$, *то многочлен* $u$ *лежит на крайнем луче* $\mathcal{P}_{2m+1}$ *тогда и только тогда, когда* $u$ *имеет ровно* $2m$ *нулей с учетом кратности.*

## § 7. Чебышевские системы на $(-\infty, \infty)$

Оставшиеся параграфы этой главы посвящены изучению моментных пространств, образованных $T$-системами, при условии, что рассматриваемая область определения является бесконечной прямой $(-\infty, \infty)$. При слабых предположениях мы сведем этот случай к случаю периодической $T$-системы. Пусть

$$u_0(t),\ u_1(t),\ \ldots,\ u_{2m}(t), \qquad -\infty < t < \infty, \tag{7.1}$$

обозначает $T$-систему на $(-\infty, \infty)$ такую, что $\{u_i\}_0^{2m-1}$ также является $T$-системой. Мы предполагаем, что $u_{2m}(t) > 0$ для $|t| \geqq A$ и

$$\lim \frac{u_i(t)}{w(t)} = 0, \qquad i = 0, 1, \ldots, 2m-1,$$

где $w(t) > 0$ для $t \in (-\infty, \infty)$ и $w(t) = u_n(t)$ для $|t| \geqq A$.

Если мы перейдем от системы (7.1) к системе

$$v_i(t) = \begin{cases} u_i(t)/w(t), & t \in (-\infty, \infty), \\ \delta_{i,n}, & t = \pm\infty, \end{cases} \tag{7.2}$$

то $\{v_i\}_0^{2m}$ образуют $T$-систему на интервале $[-\infty, \infty]$, где $+\infty$ и $-\infty$ отождествляются. Ясно, что теория предыдущих параграфов применима к системе $v_i(t)$, определенной на $[-\infty, \infty]$. Иногда бывает удобно ввести функции $g_i(x) = v_i(\operatorname{tg} \pi x/2)$, $i = 0, 1, \ldots, 2m$, так, что $\{g_i(x)\}_0^{2m}$ может рассматриваться как периодическая система на $[-1, 1]$.

Мы приведем несколько результатов, которые являются простыми следствиями результатов §§ 1—6.

I. Моментная точка $\mathbf{c} = (c_0, \ldots, c_{2m}) \in \mathcal{M}_{2m+1} = \mathcal{M}_{2m+1}(u_0(t), \ldots \ldots, u_{2m}(t))$ принадлежит $\mathcal{D}\mathcal{M}_{2m+1}$ тогда и только тогда, когда индекс $I(\mathbf{c}) \leq m$. В этом случае существует единственное конечное представление $(c_0, c_1, \ldots, c_{2m-1}, c_{2m})$.

II. Пусть $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{2m+1}$. Существует единственное каноническое представление $\sigma^*$ с массой на бесконечности. Пусть $t_1^* < t_2^* < \ldots < t_m^*$ определяет конечные корни этого представления. Для каждого $t' \neq \neq t_i^*$ $(i = 1, \ldots, m)$ существует единственное каноническое представление $\sigma_{t'}$, включающее $t'$, остальные корни которого являются непрерывными функциями $t'$. Более того, корни любых двух канонических представлений строго чередуются. Если мы не будем обращать внимание на массу в бесконечности, то первые $2m$ моментов меры $d\sigma^*$ равны $c_0, \ldots, c_{2m-1}$, а ее момент порядка $2m$ не равен $c_{2m}$ (см. гл. V).

III. Масса $\rho(t_0)$ в точке $t_0$ канонической меры $\sigma_{t_0}$ может быть представлена в виде

$$\rho(t_0) = \min \int_{-\infty}^{\infty} u(t)\, d\sigma \qquad (\sigma \in V(\mathbf{c})),$$

где минимум берется по множеству всех неотрицательных многочленов, удовлетворяющих условию нормировки $u(t_0) = 1$.

Для $T$-системы $u_k(t) = t^k$, $h = 0, 1, \ldots, 2m$, определенной на $(-\infty, \infty)$, величину $\rho(t_0)$ можно просто вычислить при помощи определенных ортогональных многочленов. Если $t_0 = t_i^*$ для некоторого $i$, то минимизирующий многочлен имеет вид $b \prod_{j=1}^{m} (t - t_j)^2$, где $\{t_1, \ldots, t_{m-1}\} = \{t_1^*, \ldots, t_m^*\} - t_0$. Если $t \neq t_i^*$ ни для какого $i$, то минимизирующий многочлен имеет вид $b \prod_{j=1}^{m} (t - t_j)^2$, где $t_1, \ldots, t_m$ — все корни $\sigma_{t_0}$, кроме $t_0$. Величина $\rho^{-1}(t_0)$ будет тогда равна

$$\rho^{-1}(t_0) = \begin{cases} \sum\limits_{k=0}^{m-1} R_k^2(t_0), & \text{если } t_0 \in \{t_1^*, \ldots, t_m^*\}, \\ \sum\limits_{k=0}^{m} R_k^2(t_0), & \text{если } t_0 \notin \{t_1^*, \ldots, t_m^*\}, \end{cases} \tag{7.3}$$

где $R_k(t)$ ортогональны на $(-\infty, \infty)$ относительно любой меры $\sigma$, соответствующей $(c_0, \ldots, c_{2m}) \in \operatorname{Int} \mathcal{M}_{n+1}$.

IV. Для $T$-системы $\{t^i\}_0^{2m}$ на $(-\infty, \infty)$ функция

$$f(\xi) = \int_{-\infty}^{\xi+} [d\sigma_1(t) - d\sigma_2(t)]$$

меняет знак по крайней мере $2m$ раз, если $\sigma_1$ и $\sigma_2$—различные меры, имеющие одинаковые моменты $c_0, \ldots, c_{2m}$ (см. лемму 3.1 гл. III).

## § 8. Представления для многочленов в $\mathcal{P}_{2m+1}\ (-\infty, \infty)$

Система $\{v_k(t)\}_0^{2m}$, определенная в (7.2), является $T$-системой, которая удовлетворяет условию $v_k(+\infty) = v_k(-\infty)$, $k = 0, \ldots, 2m$, так что система $\{g_k(x)\}_0^{2m}$, определяемая равенством

$$g_k(x) = v_k\left(\operatorname{tg} \frac{\pi}{2} x\right), \qquad k = 0, 1, \ldots, 2m,$$

является $T$-системой периодических функций на $[-1, 1)$. В силу заключения § 6 (смотри также гл. V, § 8), мы можем сформулировать следующие утверждения.

Теорема 8.1. *Если $u(t) = \sum_{i=0}^{2m} a_i u_i(t)$ положителен на $(-\infty, \infty)$ и $a_{2m} > 0$, то существует единственное представление*

$$u(t) = u^*(t) + u^{**}(t),$$

*где коэффициентом при $u_{2m}$ в $u^{**}$ является $a_{2m}$, а $u^{**}$ и $u^*$ — неотрицательные многочлены, имеющие нули $\{t_i\}_1^m$ и $\{\bar{t}_i\}_1^{m-1}$ соответственно, которые удовлетворяют условию: $-\infty < t_1 < \bar{t}_1 < t_2 < \ldots < \bar{t}_{m-1} < t_m < \infty$.*

Замечание 8.1. Устремляя $t$ к бесконечности, легко показать, что коэффициент при $u_{2m}(t)$ в многочлене $u^*(t)$ равен нулю. В этом смысле дополнительный нуль $\bar{t}_m$ многочлена $u^*$ равен бесконечности. Это соответствует частному случаю теоремы 6.1 при $t^* = \infty$.

Следствие 8.1. *Если $u_i(t) = t^i$, $i = 0, 1, \ldots, 2m$, то каждый положительный многочлен $P_{2m}(t)$ на $(-\infty, \infty)$ с $a_{2m} > 0$ допускает единственное представление в виде*

$$P_{2m}(t) = a_{2m} \prod_{j=1}^{m} (t - t_j)^2 + \beta \prod_{j=1}^{m-1} (t - \bar{t}_j)^2,$$

*где $\beta > 0$, причем корни перемежаются следующим образом: $-\infty < t_1 < \bar{t}_1 < t_2 < \ldots < \bar{t}_{m-1} < t_m < \infty$.*

Утверждение теоремы 10.1 гл. V, относящееся к границе и крайним точкам $\mathcal{P}_{n+1}[0, \infty)$ для системы $\{t^i\}_0^n$, распространяется на интервал $(-\infty, \infty)$ как для четных, так и для нечетных $n$. Следует обратить внимание на то, что когда $n$ — нечетное, каждый многочлен $\sum_{i=0}^{n} a_i t^i$ в $\mathcal{P}_{n+1}(-\infty, \infty)$ имеет по крайней мере один бесконечный нуль, так как старший коэффициент должен равняться нулю. Из этого следует, что у $\mathcal{P}_{2m+2}(-\infty, \infty)$ нет внутренних точек. Фактически, пространство нечетных многочленов $\mathcal{P}_{2m+2}(-\infty, \infty)$ можно отождествлять с $\mathcal{P}_{2m+1}(-\infty, \infty)$.

Теорема о представлении для положительных многочленов на $(-\infty, \infty)$ в общих чертах была получена Лукачом (см. Сеге [1959], стр. 4). Более точные теоремы § 8 взяты из работы Карлина [1963]. В частном случае степенных функций для получения этих результатов можно также применять принцип аргумента из комплексного анализа.

## § 9. Двумерные сечения для $\{t^i\}_0^{2m}$ на $(-\infty, \infty)$

Материал §§ 9, 10, насколько нам известно, не отражен в литературе, хотя его составные части можно легко найти в классических трудах на эту тему в основном в рассуждениях о квазиортогональных многочленах (см. Шохат и Тамаркин [1943]). Мы дополним систему $\{t^i\}_0^{2m}$ на $(-\infty, \infty)$ функциями $u_{2m+1}(t) = t^{2m+1}$, $u_{2m+2} = t^{2m+2}$ и построим моментное пространство $\mathcal{M}_{2m+3}$. Граница $\mathcal{M}_{2m+3}$ сопоставляется каноническим мерам таким образом, как это описано ниже.

В силу следствия 8.1 мы знаем, что точка $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{2m+1}$ тогда и только тогда, когда определители Генкеля, соответствующие $(c_0, c_1, \ldots, c_{2m})$, удовлетворяют неравенствам

$$\Delta_k = \begin{vmatrix} c_0 & c_1 & \cdots & c_k \\ c_1 & c_2 & \cdots & c_{k+1} \\ \cdot & \cdot & \cdot & \cdot \\ c_k & c_{k+1} & \cdots & c_{2k} \end{vmatrix} > 0, \qquad k = 0, 1, \ldots, m. \tag{9.1}$$

Пусть $\mathbf{c}^0 = (c_0^0, \ldots, c_{2m}^0) \in \operatorname{Int} \mathcal{M}_{2m+1}$, и пусть $\xi$ — произвольное вещественное число. Положим $\xi = c_{2m+1}$ и определим $c_{2m+2}(\xi)$ из соотношения $\Delta_{m+1} = 0$. Единственная мера (индекс которой необходимо равен $m+1$), которая представляет граничную точку $(c_0^0, \ldots, c_{2m}^0, \xi, c_{2m+2}(\xi))$, принадлежащую $\mathcal{M}_{2m+3}$, представляет собой также каноническую меру точки $\mathbf{c}^0$. Более того, канонические меры, исключая $\sigma^*$ (см. § 7 гл. II) находятся во взаимно однозначном соответствии с действительной прямой. Точнее, каждой $(\xi, c_{2m+2}(\xi))$, $-\infty < \xi < \infty$, сопоставляется единственная каноническая мера $\sigma_\xi$ точки $\mathbf{c}^0$. С другой стороны, корни меры $\sigma^*$ совпадают с корнями многочлена

$$Q_m(t;\infty)=\begin{vmatrix} c_0 & \cdots & c_{m-1} & 1\\ c_1 & \cdots & c_m & t\\ \cdot & \cdot & \cdot & \cdot\\ c_m & \cdots & c_{2m-1} & t^m\end{vmatrix}. \tag{9.2}$$

Заметим, что $Q_m(t;\infty)$ не включает $c_{2m}$ и что $\sigma^*$ не будет представлять $c_{2m}$, если отбросить массу на бесконечности.

Многочлен $Q_m^2(t;\infty)$ порождает единственную опорную плоскость в $(c_0, c_1, \ldots, c_{2m-1}, \underline{c}_{2m})$ и $\underline{c}_{2m}$ находится из уравнения $\Delta_m=0$.

Каноническая мера $\sigma_\xi$ сконцентрирована в корнях многочлена

$$Q_{m+1}(t;\xi)=\begin{vmatrix} c_0 & \cdots & c_{m-1} & c_m & 1\\ c_1 & \cdots & c_m & c_{m+1} & t\\ \cdot & \cdot & \cdot & \cdot & \cdot\\ c_m & & c_{2m-1} & c_{2m} & t^m\\ c_{m+1} & & c_{2m} & \xi & t^{m+1}\end{vmatrix}. \tag{9.3}$$

Разлагая этот определитель по последнему столбцу, выводим формулу

$$Q_{m+1}(t;\xi)=Q_{m+1}(t;0)-\xi Q_m(t;\infty). \tag{9.4}$$

Заметим, что многочлены $Q_k(t;\infty)$, $k=0, 1, \ldots, m$, ортогональны на $(-\infty,\infty)$ относительно каждой меры $\sigma_\xi$, в то время как $Q_{m+1}(t;\xi)$ квазиортогонален, так как он ортогонален только степеням $1, t, \ldots, t^{m-1}$ (см. стр. 119). Поучительно сравнить полученные результаты с результатами лемм 2.2 и 2.3 гл. IV. В этих леммах мы доказали для случая конечного интервала, что главные представления обладают корнями в нулях определенных ортогональных многочленов, в то время как корни канонических представлений совпадают с корнями определенных квазиортогональных многочленов. Свойство перемежаемости корней также имеет место в случае бесконечного интервала.

Закончим этот параграф кратким обсуждением двумерных сечений:

$$A{:}\,x+iy=\int\limits_{-\infty}^{\infty}\frac{d\sigma(t)}{t-z} \tag{9.5}$$

$$(\sigma\in V(\mathbf{c}^0),\qquad \mathbf{c}^0\in \operatorname{Int}\mathcal{M}_{2m+1},\ \operatorname{Im} z\neq 0).$$

Заметим, что функции

$$1, t, \ldots, t^{2m},\ \operatorname{Re}(1/t-z),\qquad \operatorname{Im}(1/t-z),\qquad -\infty<t<\infty \tag{9.6}$$

образуют $T$-систему, так как любой многочлен из этой системы допускает не более $2m+2$ нулей. Деля на $w(t)>0$, удовлетворяющее условию $\lim\limits_{|t|\to\infty}(t^{2m}/w(t))=1$, мы можем свести рассмотрение (9.5) и

(9.6) к периодическому случаю на $[-\infty, \infty]$. Из теоремы 3.1 нам известно, что граница $A$ порождается множеством канонических мер, принадлежащих $\mathbf{c}^0$. Подчеркиваем, что канонические меры находятся во взаимно однозначном соответствии с действительной прямой, т. е. для каждого $\xi$ $(-\infty < \xi < \infty)$ существует соответствующая каноническая мера $\sigma_\xi$, корни которой совпадают с нулями многочлена (9.3). Для $\sigma \in V(\mathbf{c}^0)$ мы образуем

$$R_m(z;\xi) = \int_{-\infty}^{\infty} \frac{Q_{m+1}(t;\xi) - Q_{m+1}(z;\xi)}{t-z}\, d\sigma(t). \tag{9.7}$$

Так как подынтегральное выражение является многочленом степени, не большей чем $m$, то определение (9.7) не зависит от выбора $\sigma \in V(\mathbf{c}^0)$. Используя (9.4), получим, что

$$R_m(z;\xi) = R_m(z;0) - \xi R_{m-1}(z;\infty). \tag{9.8}$$

Но $Q_{m+1}(t;\xi)$ обращается в нуль в корнях $\sigma_\xi$ и, следовательно,

$$\int_{-\infty}^{\infty} \frac{Q_{m+1}(t;\xi)}{t-z}\, d\sigma_\xi(t) = 0.$$

Поэтому

$$Q_{m+1}(z;\xi)\int_{-\infty}^{\infty} \frac{d\sigma_\xi(t)}{t-z} = \int_{-\infty}^{\infty} \frac{Q_{m+1}(z;\xi) - Q_{m+1}(t;\xi)}{t-z}\, d\sigma_\xi(t) = -R_m(z;\xi).$$

Теперь, подставляя выражения из (9.4) и (9.8), получим

$$\int_{-\infty}^{\infty} \frac{d\sigma_\xi(t)}{t-z} = -\frac{R_m(z;0) - \xi R_{m-1}(z;\infty)}{Q_{m+1}(z;0) - \xi Q_m(z;\infty)}, \qquad \zeta \in (-\infty,\infty).$$

Эта формула представляет границу $A$ как образ действительной прямой при дробно-линейном отображении. Отсюда следует, что $A$ — круг.

## § 10. Двумерные сечения для $\{t^i\}_0^{2m+1}(-\infty, \infty)$

Мы считаем, что уместно привести некоторые дополнительные замечания, относящиеся к моментному пространству, порожденному $\{t^i\}_0^n$ на $(-\infty,\infty)$ для $n = 2m+1$. (Обращаем особое внимание на то, что здесь $n$ — нечетное, в отличие от случая четного $n$, который был разобран ранее.)

Мы сначала заметим, что для того, чтобы $(\alpha_0, \ldots, \alpha_n)$ определяло опорную плоскость к $\mathcal{M}_{n+1}$, необходимо, чтобы величина $\alpha_n$ равнялась нулю (так как $n$ — нечетное). Из соотношения (9.1) следует, что $(c_0, \ldots, c_{2m+1}) \in \operatorname{Int} \mathcal{M}_{2m+2}$ тогда и только тогда, когда

$$\Delta_k = \begin{vmatrix} c_0 & \cdots & c_k \\ \cdot & \cdot & \cdot \\ c_k & & c_{2k} \end{vmatrix} > 0, \qquad k = 0, 1, \ldots, m. \tag{10.1}$$

Мы также заметим, что если $\mathbf{c} = (c_0, \ldots, c_{2m+1}) \in \operatorname{Int} \mathcal{M}_{2m+2}$, то $\mathbf{c}$ содержится в моментном пространстве, порожденном $t^i$, $i = 0, 1, \ldots, 2m+1$, на интервале $[-T, T]$ для достаточно большого $T$. В этом случае $\mathbf{c}$ имеет нижнее главное представление относительно $[-T, T]$ индекса $(n+1)/2 = m+1$, состоящего из $m+1$ внутренних точек. Поэтому каждое $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$ допускает единственное представление индекса $m+1$; единственность следует из единственности нижних главных представлений в случае конечного интервала.

Вернемся к описанию двумерных сечений. Сопоставим $(c_0, \ldots, c_{2m-1}) \in \operatorname{Int}\mathcal{M}_{2m}$ многочлены

$$R_k(t) = Q_k(t; \infty) \cdot [\Delta_k \cdot \Delta_{k-1}]^{-1/2}, \qquad k = 0, 1, \ldots, m,$$

где $Q_k(t; \infty)$ и $\Delta_k$ определены в (9.2) и (10.1) соответственно. Эти многочлены ортогональны относительно любой меры $\sigma$, имеющей моменты $c_i$, $i = 0, 1, \ldots, 2m-1$. Теперь для фиксированного $t'$, не равного никакому из корней $t_1^* < t_2^* < \ldots < t_m^*$ многочлена $Q_m(t; \infty)$, точка $(c_0, \ldots, c_{2m-1}, c_{2m+\nu})$ $(\nu > 0)$, содержащаяся во внутренней части $\mathcal{M}_{2m+1}$, имеет единственное представление, включающее $t'$ и $m$ других конечных точек. Масса $\rho(t' \colon m+1)$ в $t'$, в соответствии с (7.3), равна

$$\rho(t', m+1) = \left[\sum_0^m R_k^2(t')\right]^{-1} = \left[\frac{1}{\rho(t', m)} + \frac{Q_m^2(t', \infty)}{\Delta_m \Delta_{m-1}}\right]^{-1} =$$

$$= \left[\frac{1}{\rho(t', m)} + \frac{Q_m^2(t', \infty)}{\nu \Delta_{m-1}^2}\right]^{-1}. \tag{10.2}$$

Последнее упрощение следует из того, что $\Delta_m = 0$ для $(c_0, \ldots, c_{2m-1}, c_{2m})$. Решая это уравнение относительно $\nu$, получим

$$\nu = \frac{\alpha_m \rho(t', m+1)}{1 - \dfrac{\rho(t', m+1)}{\rho(t', m)}},$$

где $\alpha_m = \dfrac{Q_m^2(t'; \infty)}{\Delta_{m-1}^2}$. При возрастании $\nu$ от 0 до $\infty$ $\rho(t', m+1)$, очевидно, возрастает непрерывно от 0 до $\rho(t', m)$, Тогда момент порядка $(2m+1)$, $c_{2m+1}$ определяется соотношением

$$Q_{m+1}(t', c_{2m+1}) = Q_{m+1}(t', 0) - c_{2m+1} Q_m(t', \infty) = 0. \tag{10.3}$$

Заметим, что корень $t \in (t_1^*, t_2^*)$ и величина $c_{2m+1} \in (-\infty, \infty)$, определяемая из (10.3), находятся во взаимно однозначном соответствии. Рассмотрим теперь отображение $\varphi$ области $(t', \lambda)$ $(t_1^* < t' < t_2^*,\ 0 < \lambda < \rho(t', m))$ в множество $(c_{2m} + \nu, c_{2m+1})$ $(0 < \nu < \infty, -\infty < c_{2m+1} < \infty)$. При фиксированных $\nu$ и $c_{2m+1}$ мы можем решить (10.3) относительно $t'$ и затем использовать $t'$ и $\nu$ для того, чтобы вычислить $\rho(t', m+1) = \lambda$ из (10.2). Это убеждает нас в том, что отображение $\varphi$ является однозначным Также легко видеть, что отображение $\varphi$ взаимно однозначно и непрерывно. Таким образом, нами доказана

Теорема 10.1. *Если $(c_0, \ldots, c_{2m-1}) \in \operatorname{Int} \mathcal{M}_{2m}$, то для каждых $t' \notin \{t_1^*, \ldots, t_m^*\}$ и $\lambda \in (0, \rho(t', m))$ существует представление $(c_0, \ldots, c_{2m-1})$, включающее*

*$t'$ с весом $\lambda$ и включающее $m$ других точек. При $t'$, пробегающем $(t_1^*, t_2^*)$, соответствующее представление определяет взаимно однозначное непрерывное отображение*

$$(t', \lambda) \to (\underline{c}_{2m} + \nu, c_{2m+1}), \quad \nu > 0$$

## § 11. Соотношение между $\mathcal{P}_{n+1}$ и $\mathcal{M}_{n+1}$

Мы определим сечения $M^n$ и $P^n$ пространств $\mathcal{M}_{n+1}(-\infty, \infty)$ и $\mathcal{P}_{n+1}(-\infty\ \infty)$, вводя нормирующие условия $c_0 = 1$ и $\sum_{k=0}^{n} k!\, a_k = 1$ соответственно. Эти условия мы накладывали в случае полубесконечного интервала (см. § 11 гл. V). Для каждого $(c_0, \ldots, c_{2m-1}) \in M^{2m-1}(-\infty, \infty)$ определим многочлены $p_k(t)$ так же, как в § 11 гл. V, нормируя многочлены

$$\begin{vmatrix} c_0 & c_1 \ldots & c_{k-1} & 1 \\ c_1 & c_2 & c_k & t \\ \cdot & \cdot\ \cdot\ \cdot & \cdot\ \cdot\ \cdot & \cdot \\ c_k & c_{k+1} & c_{2k-1} & t^k \end{vmatrix}$$

так, чтобы $p_k^2(t)$ принадлежал $P^{2k}$.

Теперь сформулируем две теоремы, аналогичные теоремам 11.1 и 11.2 гл. V.

Теорема 11.1. *Существует взаимно однозначное, непрерывное соответствие между внутренними точками пространства $M^{2m-1}$, упорядоченными парами многочленов $(p_m, p_{m-1})$ и симплексом строго перемежающихся корней*

$$t_1 < s_1 < t_2 < \ldots < s_{m-1} < t_m.$$

Теорема 11.2. *Существует гомеоморфизм между внутренними точками $P^{2m}$ и внутренними точками $M^{2m}$ такой, что многочлены вида*

$$\lambda p_m^2 + (1 - \lambda)\, p_{m-1}^2 \qquad (0 < \lambda < 1),$$

*где $p_m$ и $p_{m-1}$ — фиксированы, сопоставляются моментным точкам, имеющим такие же первые $2m - 1$ момента.*

Мы оставляем читателю задачу сформулировать аналог теоремы 12.1 гл. V.

Глава VII

# МОМЕНТНЫЕ ПРОСТРАНСТВА ДЛЯ ЧЕБЫШЕВСКИХ СИСТЕМ, ОПРЕДЕЛЕННЫХ НА ДИСКРЕТНОМ МНОЖЕСТВЕ

## § 1. Введение

В этой главе мы изучим строение моментного пространства для чебышевской системы функций, определенной на дискретном множестве $\overline{K}$ с $k = +\infty \in \overline{K}$, являющейся его единственной предельной точкой. Целесообразность изучения таких функций будет указана ниже. Сейчас мы перейдем к формальным определениям.

Определение 1.1. Система $\{u_i\}_0^n$ функций, определенных на $\overline{K}$, является *T-системой*, если

$$U\begin{pmatrix} 0, & 1, & \ldots, & n \\ k_0, & k_1, & \ldots, & k_n \end{pmatrix} > 0, \quad k_0 < k_1 < \ldots < k_n,$$

$$k_i \in \overline{K}, \qquad i = 0, 1, \ldots, n, \tag{1.1}$$

и функции $u_0, u_1, \ldots, u_n$ непрерывны в точке $k = \infty$, т. е.

$$\lim_{k\to\infty} u_i(k) = u_i(\infty), \qquad i = 0, 1, \ldots, n. \tag{1.2}$$

Для удобства обозначений мы отождествим $\overline{K}$ с неотрицательными целыми числами, включая точку на бесконечности. В соответствии с этим соглашением мы обозначим множество неотрицательных целых чисел без точки на бесконечности через $K$.

Обычная характеризация $T$-системы $u_0(t), u_1(t), \ldots, u_n(t)$ состоит в том, что каждый множитель обращается в нуль не более чем $n$ раз (смотри теорему 4.1 гл. I). Ясно, что для систем, определенных на интервале $[a, b]$, это свойство эквивалентно условию

$$U\begin{pmatrix} 0, & 1, & \ldots, & n \\ t_0, & t_1, & \ldots, & t_n \end{pmatrix} \neq 0, \qquad a \leq t_0 < t_1 < \ldots < t_n \leq b. \tag{1.3}$$

Так как оговорено, что каждая из функций $u_i(t)$ непрерывна на $[a, b]$, то соотношение (1.3) влечет, что определители в (1.3) имеют постоянный знак, не зависящий от выбора $t_0 < t_1 < \ldots < t_n$. Теперь мы имеем дело с той ситуацией, когда функции $u_i(\cdot)$ определены на дискретном множестве. В этом случае из того факта, что определитель (1.3) не обращается в нуль, не следует, вообще говоря, что определитель выражения (1.3) обладает фиксированным знаком. Однако для большей части нашего анализа существенно, что имеет место (1.1), т. е. что все определители (1.1) сохраняют постоянный знак. Сейчас определение $T$-системы при помощи неравенств определителей (см. определение 1.1 гл. I) более приемлемо, чем при помощи нулей многочленов и фактически формулировки гл. I не совпадают в новой ситуации.

Трудности, присущие анализу моментных пространств, когда рассматриваемый интервал является дискретным множеством, проистекают из того, что мы не можем использовать соображения непрерывности. Эта необходимость более сильного понятия $T$-системы, определенной на дискретном множестве, подобна необходимости усиления понятия строгой полной положительности. Ослабление предположений о полной положительности ядер играет решающую роль в этом анализе.

В этом случае мы будем изучать моментное пространство

$$\mathcal{M}_{n+1}(\overline{K}) = \{\mathbf{c} = (c_0, \ldots, c_n) \mid c_i = \\ = \int\limits_{\overline{K}} u_i(k)\, d\sigma(k) = \sum_{k \in \overline{K}} u_i(k)\, \sigma_k, \quad i = 0, 1, \ldots, n\}, \qquad (1.4)$$

где $\sigma$ пробегает множество ограниченных дискретных мер на $\overline{K}$.

Важный пример $T$-системы на дискретном множестве встречается при изучении абсолютно монотонных функций $f(x)$, определенных на $(0, \infty)$ или на интервале $(0, b]$, где $b$ конечно. Известно, что такие функции разлагаются в степенные ряды

$$f(x) = \sum_{k=0}^{\infty} a_k x^k, \qquad a_k \geqslant 0,$$

сходящиеся для $|x| \leqslant b$. Каждое множество $n+1$ величин

$$f(x_i) = \sum_{k=0}^{\infty} a_k x_i^k, \qquad i = 0, 1, \ldots, n,$$

можно рассматривать как точку в моментном пространстве, порожденном функциями

$$u_i(k) = x_k^k \qquad i = 0, 1, \ldots, n, \quad k = 0, 1, 2, \ldots \qquad (1.5)$$

В более знакомых обозначениях имеем

$$f(x_i) = \int_{K} u_i(k)\, d\sigma(k), \qquad i = 0, 1, \ldots, n,$$

где $\sigma(k)$ имеет массу $a_k$ в точке $k$. Принимая во внимание пример 1 в § 3 гл. I, мы видим, что система (1.5) для $0 < x_0 < x_1 < \ldots < x_n < \infty$ образует $T$-систему на неотрицательных целых числах.

Для фиксированного набора действительных чисел $c_0, c_1, \ldots, c_n$ естественно потребовать существования абсолютно монотонной функции на $(0, b]$, которая принимает значения $c_0, c_1, \ldots, c_n$ в точках $x_0 < x_1 < \ldots < x_n$. Эта задача является задачей интерполяции набора $n+1$ абсцисс и $n+1$ ординат абсолютно монотонной функцией. Иногда нам требуется, чтобы интерполирующая функция была многочленом с неотрицательными коэффициентами. Геометрическая моментная теория дает естественный язык для описания этой интерполяционной теории.

Интерполяционная задача разрешима, если точка $(c_0, c_1, \ldots, c_n)$ лежит в моментном пространстве, порожденном функциями (1.5). Дальнейшую информацию также можно получить, рассматривая канонические представления моментной точки $(c_0, c_1, \ldots, c_n)$ и соответствующие ему величины. Эти задачи будут сформулированы в § 10.

Общий класс $T$-систем $u_i(k)$, $i = 0, 1, \ldots, n$, определенных на целых числах, строится при помощи ядер $\varphi(x, k)$, определенных для $k = 0, 1, \ldots$ и $0 < x \leq b$, которые являются $STP$ по переменным $x$ и $k$ (см. определение 2.2 гл. I).

Для $0 < x_0 < x_1 < \ldots < x_n \leq b$ система функций $u_i(k) = \varphi(x_i, k)$, $i = 0, 1, \ldots, n$ есть $T$-система на $K$, так как по предположению для каждого $n \geqslant 0$

$$\Phi\begin{pmatrix} x_0, x_1, \ldots, x_n \\ k_0, k_1, \ldots, k_n \end{pmatrix} = \det[\,\|\varphi(x_i, k_j)\|_{i,j=0}^{n}\,] \tag{1.6}$$

положителен при любом выборе

$$0 \leq k_0 < k_1 < \ldots < k_n < \infty \text{ и } 0 < x_0 < x_1 < \ldots < x_n \leq b.$$

В этом случае, когда функции $u_i(k)$ определены только для $k \in K$, как в примере абсолютно монотонных функций, мы будем предполагать, что

(a) $u_n(k) > 0$ для всех $k > \widehat{k}$, где $\widehat{k}$ — фиксированное положительное целое;

(в) $\lim\limits_{n \to \infty} \dfrac{u_i(k)}{u_n(k)} = 0$, $i = 0, 1, \ldots, n-1$, и

(c) $\{u_i\}_0^{n-1}$ и $\{u_i\}_0^{n}$ — $T$-системы на $K$.

В этой ситуации мы, так же как и в гл. V, переходим к $T$-системе

$\{v_i(k)\}_0^n$, $k \in \overline{K}$, определяемой так:

$$v_i(k) = \begin{cases} \dfrac{u_i(k)}{w(k)}, & k \in K, \\ \delta_{in}, & k = \infty, \end{cases} \qquad i = 0, 1, \ldots, n,$$

где $w(k)$ положительна на $K$ и $w(k) = u_n(k)$ для $k > \hat{k}$. Система $\{v_i\}_0^n$ образует $T$-систему на замкнутом компакте $\overline{K}$, так что изучение системы $\{u_i\}_0^n$ по существу сводится к изучению системы $\{v_i\}_0^n$, которая удовлетворяет (1.2).

Основная часть этой главы посвящена анализу чебышевских систем, определенных на $\overline{K}$. Дальнейшее обсуждение абсолютно монотонных функций и систем, удовлетворяющих условиям (a), (b), (c) проводится в §§ 8 и 9.

Можно было бы изучать моментные пространства, порожденные чебышевской системой $u_1(t)$, $u_2(t)$, ..., $u_n(t)$, определенной на линейно упорядоченном множестве общего вида $\Gamma$ (частными случаями являлись бы дискретные множества и интервалы), причем эти системы несомненно имели бы теоретический интерес и можно было бы найти их практические приложения. Мы ограничимся частным случаем, когда $\Gamma$ состоит из неотрицательных целых, включая предельную точку $\infty$. Геометрия, лежащая в основе конструкции класса абсолютно монотонных функций, явилась причиной некоторых исследований этой главы. Основное внимание мы уделяем задаче интерполяции по заданному множеству ординат $c_1$, $c_2$, ..., $c_n$ в некоторых специально выбранных точках $0 < x_1 < x_2 < \ldots < x_n$ абсолютно монотонными функциями. Решение этой задачи естественным образом вкладывается в теорию моментных пространств, как это сделано в настоящей главе. Розенблюм [1951] и независимо от него Крейн и Рехтман [1955] исследовали различные близкие задачи. Происхождение многих из этих идей связано с работами Бернштейна [1926], [1950]. В §§ 2—8 продолжаются исследования Крейна и Рехтмана [1955], § 9 основывается на одной работе Розенблюма.

## § 2. Индекс множества $\Sigma \ddot{<} \overline{K}$

Когда рассматриваемое множество $\overline{K}$ дискретно, то обычное понятие нуля функции становится неадекватным и в этом случае более естественно иметь дело с понятием изменения знака. Для функции $f(t)$, определенной на линейном множестве $\Gamma$, обозначим (см. определение 3.1 гл. III) через $S(f)$ число действительных изменений знака этой функции. В этом разделе нам потребуется более сильное понятие, немного напоминающее понятие кратности нуля.

Определение 2.1. Для произвольной действительнозначной функции $f$, заданной на линейном множестве $\Gamma$, определим $S^+(f)$ соотношением

$$S^+(f) = \sup S^+(f(t_1), \ldots, f(t_p)),$$

где $S^+(f(t_1), \ldots, f(t_p))$ обозначает максимальное число изменений знака последовательности $f(t_1), \ldots, f(t_p)$, причем нуль дает произвольный знак. Верхняя грань берется по всевозможным наборам $t_1 < \ldots < t_p$, где $t_i \in \Gamma$, $p$ — произвольно.

При таком определении легко видеть, что если неотрицательный многочлен $u(t) = \sum_{i=0}^{n} a_i u_i(t)$ определен на $[a, b]$ и обращается в нуль только в различных точках $t_1, \ldots, t_p$, то $2I(t_1, \ldots, t_p) = S^+(u)$ (см. определение 2.1 гл. II).

Теперь мы можем проверить, что для $T$-системы $\{u_i\}_0^n$ на произвольном линейном множестве $\Gamma$ любой ненулевой многочлен $u = \sum_{i=0}^{n} a_i u_i$ удовлетворяет соотношению

$$S^+(u) \leq n. \tag{2.1}$$

Для того чтобы установить это, мы предположим противное. Тогда для некоторого множества $\{t_i\}_0^{n+1} \subset \Gamma$, удовлетворяющего $t_0 < t_1 < \ldots < t_{n+1}$, имеем $S^+(u(t_0), \ldots, u(t_{n+1})) = n + 1$.

В этом случае рассмотрим определитель:

$$\begin{vmatrix} u(t_0) & u_0(t_0) & \ldots & u_n(t_0) \\ u(t_1) & u_0(t_1) & \ldots & u_n(t_1) \\ \vdots & \vdots & & \\ u(t_{n+1}) & u_0(t_{n+1}) & \ldots & u_n(t_{n+1}) \end{vmatrix} = \sum_{i=0}^{n+1} (-1)^i u(t_i)\, U\begin{pmatrix} 0, & 1, & \ldots & & n \\ t_0, & t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_{n+1} \end{pmatrix}. \tag{2.2}$$

Этот определитель равен нулю, так как первый столбец является линейной комбинацией остальных столбцов. Теперь величины $u(t_0), \ldots, u(t_{n+1})$ имеют чередующиеся знаки в том смысле, что $u(t_i)\,u(t_{i+1}) \leq 0$. Так как определители, стоящие в правой части (2.2), строго положительны в силу чебышевского свойства, то отсюда следует, что $u(t_i) = 0$, $i = 0, 1, \ldots, n+1$. Однако это невозможно, так как определитель системы уравнений $u(t_j) = \sum_{i=0}^{n} a_i u_i(t_j) = 0$, $j = 0, 1, \ldots, n$, не равен нулю. Полученное противоречие доказывает справедливость (2.1).

Определение 2.2. Пусть $\Sigma = \{k_j\}_1^p$ — множество различных точек в $\overline{\Gamma}$. Построим функцию $f(k)$, $k \in \overline{K}$, обращающуюся в нуль только на $\Sigma$ и положительную в остальных точках; например, пусть $f$ будет характеристической функцией $\Sigma^c$, где $\Sigma^c$ — дополнение $\Sigma$ в $K$. Тогда *индексом* $\Sigma$ называется

$$I(\Sigma) = S^+(f). \tag{2.3}$$

Другой путь определения индекса данного конечного множества состоит в сопоставлении элементам $\overline{K}$ последовательности знаков в соответствии со следующим правилом. Каждое $k \notin \Sigma$ имеет положительный знак. С другой стороны, точки $k_i \in \Sigma$ имеют произвольные знаки, расставленные таким образом, что при $k$, пробегающем $\overline{K}$, достигается максимальное число изменений знака; например, обозначим его через $N$. Индекс множества $\Sigma$ определяется тогда равенством $I(\Sigma) = N$.

Мы предупреждаем читателя, что понятие индекса, определенное здесь, не совпадает с величиной индекса, определенной в гл. II. Они отличаются на множитель 2.

Теперь мы перейдем к выявлению вклада отдельных элементов $\Sigma = \{k_i\}_1^p$ в величину индекса. Поскольку каждое $k_i$, $i = 1, \ldots, p$, вносит по крайней мере единицу в индекс, то заключаем, что $I(\{k_i\}_{i=1}^p) \geqslant p$ для любого целого $p$.

Нетрудно убедиться, что каждая изолированная точка $k^* \in \Sigma$ (т. е. такая точка, что $k^* - 1 \in K$ и $k^* + 1 \in K$ не принадлежат $\Sigma = \{k_i\}_{i=1}^p$) вносит две единицы в величину индекса. В общем случае набор величин

$$k_{i_0},\ k_{i_0+1} = k_{i_0+1},\quad k_{i_0+2} = k_{i_0} + 2, \ldots, k_{i_0+\nu} = k_{i_0} + \nu, \tag{2.4}$$

принадлежащих $\Sigma$ и таких, что $k_{i_0} - 1$, $k_{i_0+\nu} + 1$ не принадлежат $\Sigma$, в то время как $k_{i_0} - 1 \in K$ вносит $\nu + 1$ $(\nu + 2)$ единицу в индекс $I(\Sigma)$, если $\nu$ — нечетное (четное). Однако если $k_{i_0} = 0$, то вклад в индекс равен $\nu + 1$ независимо от того, четное $\nu$ или нечетное.

В частности, изолированная точка $\Sigma = \{k_{i_0}\}$, $k_{i_0} > 0$, обладает индексом 2, в то время как изолированная пара последовательных точек также имеет индекс 2.

Проведенные рассуждения показывают, что множества $\Sigma = \{0\}$ и $\Sigma = \{\infty\}$ имеют каждое индекс 1. Более того, множества $\Sigma = \{k_0, k_0 + 1\}$, $\Sigma = \{k_0\}$ $(0 < k_0 < \infty)$ и $\Sigma = \{0, \infty\}$ имеют индекс 2. Множества индекса 3 имеют следующий вид:

$$\begin{aligned}
&\Sigma = \{0, k_0\, k_0 + 1\}, && k_0 > 0, \quad k_0 \in K,\\
&\Sigma = \{k_0, k_0 + 1, \infty\}, && k_0 \geqslant 0, \quad k_0 \in K,\\
&\Sigma = \{0, k_0\}, && 1 < k_0 < \infty,\\
&\Sigma = \{k_0, \infty\}, && k_0 > 0, \quad k_0 \in K.
\end{aligned}$$

## § 3. Многочлены с заданными нулями

Теорема 3.1. *Пусть $\{u_i\}_0^n$ — $T$-система на $\overline{K}$. Для любого множества $\Sigma = \{k_i\}_{i=1}^q \subset \overline{K}$, для которого $I(\Sigma) \leq n$, существует неотрицательный многочлен $u(k) = \sum_{i=0}^{n} a_i u_i(k)$, нулями которого являются только элементы множества $\Sigma$.*

Доказательство. Предположим, что $I(\Sigma) = p \leq n$. Любое конечное множество $\Sigma$ можно разбить на блоки последовательных величин. Если мы добавим к каждому нечетному блоку точек из $\Sigma$ (не содержащему $k=0$) точку, непосредственно предшествующую блоку, то мы сможем увеличить $\Sigma$ до множества $\Sigma_1$, состоящего из $p$ точек. Заметим, что четные блоки остаются неизменными. Теперь, если к каждому нечетному блоку (не содержащему $k=0$) добавить точку, непосредственно следующую за блоком, то мы определим таким образом другое множество $\Sigma_2$, также содержащее $p$ точек и такое, что $\Sigma_1 \cap \Sigma_2 = \Sigma$.

Теперь, если $n$ и $p$ имеют одинаковую четность, т. е. они либо оба четны, либо нечетны, мы можем последовательно расширять $\Sigma_1$ и $\Sigma_2$, добавляя достаточное число пар точек, имеющих большие значения, чтобы получить множества $\Sigma_1'$ и $\Sigma_2'$, имеющие каждое индекс $n$, содержащие по $n$ точек и удовлетворяющие соотношению $\Sigma_1' \cap \Sigma_2' = \Sigma$. Образуем многочлены

$$u^{(i)}(k) = c_i \begin{vmatrix} u_0(s_1^i) & u_0(s_2^i) & \dots & u_0(s_n^i) & u_0(k) \\ u_1(s_1^i) & u_1(s_2^i) & \dots & u_1(s_n^i) & u_1(k) \\ \dots & \dots & \dots & \dots & \dots \\ u_n(s_1^i) & u_n(s_2^i) & \dots & u_n(s_n^i) & u_n(k) \end{vmatrix} \quad i = 1,2; \tag{3.1}$$

где $\{s_j^1\}$, $\{s_j^2\}$ — элементы $\Sigma_1'$ и $\Sigma_2'$ соответственно, расположенные в возрастающем порядке.

Многочлен $u^{(i)}(k)$ обращается в нуль только на $\Sigma_i'$ и сохраняет постоянный знак на всем $\overline{K}$, так как блоки $\Sigma_i'$, которые не содержат ни $k=0$, ни $k=\infty$, имеют четное число элементов. Постоянная $c_i = \pm 1$ выбирается таким образом, чтобы $u^{(i)}(k)$ был везде неотрицателен, так что $c_i = -1$ тогда и только тогда, когда $s_n^i = \infty$. Требуемым многочленом является тогда многочлен $u = u^{(1)} + u^{(2)}$, который обращается в нуль только на $\Sigma$, так как $\Sigma_1' \cap \Sigma_2' = \Sigma$.

Теперь мы рассмотрим случай, когда $n$ и $p$ имеют различную четность; будем различать две ситуации:

(1) $n = 2m+1$ и $p$ — четное.

(2) $n = 2m$ и $p$ — нечетное.

В обоих случаях мы, как и раньше, увеличиваем $\Sigma$ так, чтобы получить $\Sigma_1'$ и $\Sigma_2'$, где $\Sigma_1' \cap \Sigma_2' = \Sigma$, с тем лишь изменением, что теперь $\Sigma_1'$ и $\Sigma_2'$ имеют индекс $n-1$ и ровно $n-1$ точек.

В первом случае имеется две возможности: (a) $k = \infty \in \Sigma_i'$, $i=1,2$, и первый блок содержит или $k=0$ или (в противном случае) нечетное число элементов, (b) $k = \infty \notin \Sigma_i'$, $i=1, 2$, и первый блок содержит четное число элементов. Для случая (a) мы образуем $u^{(i)}$, выбирая определенным образом значения в $n+1$ точке. Пусть $s_0$ — наименьшая точка, не принадлежащая $\Sigma_i'$. Построим $u^{(i)}$, положив его равным нулю на $\Sigma_i'$, единице в $s_0$ и в некоторой другой точке

$s_1 \notin \Sigma_i'$. Многочлен $u^{(i)}$ не может обращаться в нуль или принимать отрицательные значения вне $\Sigma_i'$, так как в противном случае было бы $S^+(u^{(i)}) \geqslant n+1$. Требуемый многочлен определяется как $u^{(1)}+u^{(2)}$.

Для случая (b) мы построим $u^{(1)}$ и $u^{(2)}$, как в (3.1), выбрав определенным образом их нули так, чтобы они совпадали с $\Sigma_1' \bigcup \{s_0\}$ и $\Sigma_2' \bigcup \{\infty\}$ соответственно.

Остается рассмотреть случай 2. Опять мы имеем дело с двумя подслучаями: (a) $k = \infty \in \Sigma_i'$, $i = 1, 2$, и 1-й блок содержит четное число элементов; (b) $k = \infty \notin \Sigma_i'$, $i = 1, 2$, и 1-й блок нечетен и содержит $k = 0$. В обоих подслучаях рассуждения проводятся точно так же, как и в первом случае. Этим заканчивается доказательство.

Проведенное выше построение можно обобщить таким образом, чтобы получить многочлены, которые удовлетворяют определенным правилам изменения знака. Следующая теорема относится к теоремам такого типа. Мы опускаем доказательство, так как необходимые рассуждения аналогичны проведенным выше.

Теорема 3.2. *Пусть $\{u_i\}_0^n$ — $T$-система на $\overline{K}$, и пусть $f$ — произвольная функция, определенная на $\overline{K}$ и удовлетворяющая условию $S^+(f) \leqslant n$. Тогда существует нетривиальный многочлен $u$ такой, что $u(k) f(k) \geqslant 0$ для $k \in \overline{K}$.*

## § 4. Граница и канонические представления точек в $\mathcal{M}_{n+1}(\overline{K})$

Напомним определение моментного пространства:

$$\mathcal{M}_{n+1}(\overline{K}) = \left\{ \mathrm{c} = (c_0, \ldots, c_n) \mid c_i = \int_{\overline{K}} u_i(k)\, d\sigma(k), \quad i = 0, \ldots, n \right\}, \tag{4.1}$$

где $\sigma$ пробегает множество ограниченных мер в $\overline{K}$. Пространство $\mathcal{M}_{n+1}(\overline{K})$, как легко видеть, содержит внутренние точки. Кроме того, из теоремы 3.1 следует, что существует строго положительный многочлен. Используя это, мы получим, что выпуклый конус $\mathcal{M}_{n+1}(\overline{K})$ замкнут. Теперь рассмотрим дискретную кривую

$$C_{n+1} = C_{n+1}(\overline{K}) = \{\mathrm{c} \mid c_i = u_i(k), \quad i = 0, 1, \ldots, n, \ k \in \overline{K}\}. \tag{4.2}$$

Следующая теорема о конической оболочке является частным случаем теоремы 2.1.

Теорема 4.1. *Выпуклым конусом, порожденным $C_{n+1}$, является $\mathcal{M}_{n+1}(\overline{K})$.*

С помощью теоремы 3.1 мы можем также установить при помощи ставшей уже стандартной техники следующую теорему.

Теорема 4.2. *Точка* $\mathbf{c}$ *принадлежит границе* $\mathcal{M}_{n+1}(\overline{K})$ *тогда и только тогда, когда* $\mathbf{c}$ *допускает единственное представление*

$$c_i = \sum_{j=1}^{p} \lambda_j u_i(k_j), \qquad i = 0, 1, \ldots, n, \tag{4.3}$$

*где* $\lambda_j > 0$, $k_j \in \overline{K}$ *и* $\Sigma = \{k_j\}_1^p$ *удовлетворяет условию* $I(\Sigma) \leq n$.

Определение 4.1. Точки $\{k_i\}_{i=1}^p$, фигурирующие в представлении типа (4.3), называются *корнями представления*, а $\{\lambda_j\}_1^p$ — соответствующими *весами*. *Индексом представления* (4.3) называется индекс его множества корней. *Индексом* произвольной *моментной точки* $\mathbf{c}$ называется минимум индексов представлений $\mathbf{c}$.

Теперь мы перейдем к изучению $\operatorname{Int} \mathcal{M}_{n+1}(\overline{K})$. Пусть $\mathbf{c}^0 = (c_0^0, c_1^0, \ldots, c_n^0)$ принадлежит $\operatorname{Int} \mathcal{M}_{n+1}(\overline{K})$, и пусть $S$ — сечение $\mathcal{M}_{n+1}(\overline{K})$, содержащее точку $\mathbf{c}^0$. Сечение $S$ можно образовать посредством наложения нормирующего условия $\sum_{i=0}^{n} a_i c_i = \alpha > 0$, где $\sum_{i=0}^{n} a_i u_i(k)$ — фиксированный положительный многочлен. Заметим, что сечение $S$ является выпуклым, замкнутым и ограниченным. Для каждого фиксированного $k_0 \in \overline{K}$ мы рассмотрим точку $\mathbf{c}(k_0) = \alpha(u_0(k_0), u_1(k_0), \ldots, u_n(k_0))$, где $k_0 \in \overline{K}$ выбрано так, чтобы $c(k_0) \in S$. Теперь проведем отрезок через $\mathbf{c}(k_0)$ и $\mathbf{c}^0$ до пересечения с границей $\mathcal{M}_{n+1}(\overline{K})$, например, в точке $\mathbf{c}^*(k_0)$. В силу теоремы 4.2 $\mathbf{c}^*(k_0) \in \in \operatorname{Bd} \mathcal{M}_{n+1}(\overline{K})$ имеет представление вида (4.3), имеющее индекс, не больший чем $n$. Точку $\mathbf{c}^0$, очевидно, можно представить как выпуклую комбинацию $c(k_0)$ и $c^*(k_0)$, которая дает для $\mathbf{c}^0$ представление индекса, не большего чем $n+2$. С другой стороны, так как $\mathbf{c}^0$ принадлежит $\operatorname{Int} \mathcal{M}_{n+1}(\overline{K})$, то в силу теоремы 4.2 $I(\mathbf{c}^0) \geqslant n+1$.

Определение 4.2. Любое представление $c^0$ индекса $n+1$ называется *главным*, а все представления индекса, не большего чем $n+2$, — *каноническими*. Предыдущие рассуждения доказывают первую часть следующей теоремы.

Теорема 4.3. *Пусть* $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}(\overline{K})$. *Для любого* $k_0 \in \overline{K}$ *существует каноническое представление* $\mathbf{c}^0$, *включающее* $k_0$ *в качестве корня. Среди всех представлений* $\mathbf{c}^0$, *включающих* $k_0$, *одно, полученное при помощи геометрического построения, описанного выше, имеет максимальную массу в точке* $k_0$ *и является единственным представлением, имеющим эту максимальную массу.*

Доказательство второго утверждения теоремы проходит следующим образом.

Построенное выше каноническое представление $\mathbf{c}^0$, включающее $k_0$, определяется из выпуклого представления:

$$\mathbf{c}^0 = \lambda\alpha\mathbf{c}(k_0) + \mu\mathbf{c}^*(k_0), \qquad \lambda > 0, \quad \mu > 0, \quad \lambda + \mu = 1, \tag{4.4}$$

где $\mathbf{c}(k_0) = (u_0(k_0), u_1(k_0), \ldots, u_n(k_0))$, а $\mathbf{c}^*(k_0)$ — граничная точка $\mathcal{M}_{n+1}(\bar{K})$. Пусть $u(k)$ — неотрицательный многочлен (см. теорему 3.1), обращающийся в нуль только в корнях единственного представления $\mathbf{c}^*(k_0)$ и нормированный таким образом, что $u(k_0) = 1$. Если $\sigma_{k_0}$ — каноническое представление, соответствующее соотношению (4.4), а $\sigma$ — любое другое представление $\mathbf{c}^0$, то

$$\sigma_{k_0}(\{k_0\}) = \int u(k)\, d\sigma_{k_0}(k) = \int u(k)\, d\sigma(k) \geqslant \sigma(\{k_0\}). \tag{4.5}$$

Равенство в (4.5) имеет место только тогда, когда спектр $\sigma$ совпадает со спектром $\sigma_{k_0}$. Но представление

$$c_i = \sum_{j=1}^{n} \lambda_j u_i(k_j), \qquad i = 0, 1, \ldots, n \quad (p \leq n+1),$$

с $\{k_j\}$ — корнями $\sigma_{k_0}$, однозначно разрешимо относительно $\lambda_1, \lambda_2, \ldots$ $\ldots, \lambda_p$, так что $\sigma = \sigma_{k_0}$.

Для $T$-системы, определенной на интервале $[a, b]$, мы каждому $t_0 \in (a, b)$ и моментной точке $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}$ сопоставили единственное каноническое представление $\mathbf{c}$. В настоящей ситуации мы не исключаем возможности, при которой внутренняя точка $\mathbf{c}$ обладает двумя различными каноническими представлениями с корнями, принадлежащими одному и тому же множеству $\Sigma$ индекса, не большего чем $n+2$. Такая ситуация действительно возможна. С помощью элементарных вычислений можно убедиться в том, что для $T$-системы $u_0(k) = (1/2)^k$, $u_1(k) \equiv 1$, моментная точка $c_0 = 7/12$, $c_1 = 1$ имеет бесконечное число представлений, включающих множество $\Sigma = \{0, 1, 2\}$. В частности, $\sigma\{0\} = \sigma\{1\} = \sigma\{2\} = 1/3$ и $\sigma\{0\} = 7/18$, $\sigma\{1\} = 1/6$, $\sigma\{2\} = 4/9$ являются двумя решениями. Заметим, что в этом случае множество $\Sigma$ содержит $n+2$ точки. Данное каноническое представление не обязано иметь максимальную массу в каждой из своих точек в $(0, \infty)$. К этому вопросу относится следующая теорема.

Теорема 4.4. *Пусть* $\sigma^0$ — *каноническое представление* $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}(\bar{K})$. *Пусть* $\Sigma$ *обозначает корни* $\sigma^0$, *и предположим, что* $k_0$, $k_0+1$ — *две последовательные точки такие, что* $\Sigma' = \Sigma - \{k_0, k_0+1\}$ *имеет индекс, не больший чем* $n$ (*это имеет место, в частности, если* $k_0$ *и* $k_0+1$ *оба являются корнями* $\sigma^0$). *Тогда существуют две постоянные* $\alpha > 0$ *и* $\beta > 0$ *такие, что*

$$\alpha\sigma^0(\{k_0\}) + \beta\sigma^0(\{k_0+1\}) \geqslant \alpha\sigma(\{k_0\}) + \beta\sigma(\{k_0+1\}), \tag{4.6}$$

*где* $\sigma$ — *любое другое представление* $\mathbf{c}^0$. *Равенство в* (4.6) *имеет место тогда и только тогда, когда спектр* $\sigma$ *принадлежит множеству* $\Sigma$.

Доказательство. Так как $\Sigma - \{k_0, k_0+1\}$ имеет индекс $n$, то мы можем построить, в соответствии с теоремой 3.1, неотрицательный многочлен $u$, который обращается в нуль только на этом множестве. Тогда, если $\alpha = u(k_0)$ и $\beta = u(k_0+1)$, то

$$\alpha\sigma^0(\{k_0\}) + \beta\sigma^0(\{k_0+1\}) = \int u\,d\sigma^0 =$$

$$= \int u\,d\sigma \geqslant \alpha\sigma(\{k_0\}) + \beta\sigma(\{k_0+1\}).$$

Равенство имеет место тогда и только тогда, когда спектр $\sigma$ принадлежит множеству $\Sigma$.

## § 5. Главные представления

Если $k_0 = 0$ является корнем, то каноническое представление, определяемое теоремой 4.3, является главным. Действительно, напомним, что $I(\mathbf{c}^*(0)) \leq n$, так как $\mathbf{c}^*(0)$ — граничная точка $\mathcal{M}_{n+1}(\overline{K})$. Присоединяя корень $k_0 = 0$ к корням $\mathbf{c}^*(0)$, можно увеличить индекс самое большее на единицу. Это показывает, что построение теоремы 4.3 дает нам главное представление, когда $k_0=0$. Аналогично, если $k_0 = \infty$ является корнем, то описанная выше процедура устанавливает существование главного представления для $\mathbf{c}^0$, включающее корень $\infty$.

Если $n = 2m$, то два главных представления, полученные таким образом, различны. Фактически множество корней $\Sigma(0) = \{k_i(0)\}_{i=1}^{p}$ точки $\mathbf{c}^*(0)$ (однозначно определенное в силу того, что $\mathbf{c}^*(0) \in \in \mathrm{Bd}\,\mathcal{M}_{n+1}(\overline{K})$) обладает тем свойством, что $0 \notin \Sigma(0)$, так как в противном случае $I(\mathbf{c}^0) = n$. Из этого следует, что $\infty \notin \Sigma(0)$, так как $I(\Sigma(0)) = n = 2m$. (Здесь и в дальнейшем важно помнить, что индекс любого множества $\Sigma = \{k_i\}_{i=1}^{p}$, где $k_i \geqslant 1$ и $k_p < \infty$, является всегда четным, в то время как индекс множества, включающего $\infty$, но не включающего нуль, всегда нечетный. Подобным образом, если 0 принадлежит $\Sigma$, а $\infty$ не принадлежит и если первый блок корней в $\Sigma$ содержит нечетное число элементов, то индекс является нечетным.)

Продолжая рассмотрение четного случая ($n = 2m$), мы получим различные результаты, относящиеся к главным представлениям.

Пусть $\mathbf{c}^0 \in \mathrm{Int}\,\mathcal{M}_{n+1}(\overline{K})$.

(i) *Главное представление, включающее точку $k_0 = 0$, единственно. Это главное представление характеризуется тем свойством, что среди всех представлений $\mathbf{c}^0$ вес, приписываемый нулевой точке, максимален.*

(ii) *Главное представление, включающее точку $k_0 = \infty$, единственно и его соответствующий вес максимален.*

(iii) *Существует ровно два главных представления.*

Доказательство.

(i) Рассмотрим главное представление $\sigma_0$, включающее корень нуль. Пусть $\Sigma = \{k_i\}_{i=1}^{p}$ — его корни, кроме нуля. Тогда $I(\Sigma) = n$. Используя теорему 3.1, мы получим неотрицательный многочлен

$u^*(k)$, обращающийся в нуль только на $\Sigma$ и нормированный таким образом, что $u^*(0) = 1$.

Так как $\int u\,d\sigma = \sum_{i=0}^{n} a_i c_i^0$ для любого многочлена $u(k) = \sum_{i=0}^{n} a_i u_i(k)$, где $\sigma$ — любое представление $\mathbf{c}^0$, то $\sigma_0(\{0\}) = \int u^* d\sigma_0 = \int u^* d\sigma \geqslant \sigma(\{0\})$, причем равенство имеет место, только когда $\sigma = \sigma_0$. Отсюда легко следует заключение (i).

(ii) Доказательство аналогично доказательству (i).

(iii) В силу (i) и (ii) множество корней любого другого главного представления не включает ни 0 ни $\infty$, т. е. корни содержатся в множестве $\tilde{K} = 1, 2, \ldots$ Индекс любого такого множества должен быть четным и, следовательно, не равным $n + 1 = 2m + 1$. Это означает, что не может существовать главного представления, не включающего ни 0 ни $\infty$, что и завершает доказательство.

Рассмотрим теперь случай, когда $n = 2m + 1$. Мы построим главное представление, включающее $k_0 = 0$, как и ранее, но в этом случае точка $k_0 = \infty$ также должна быть включена в это представление. Для подтверждения этого рассмотрим $k_0 = \infty$ и определим главное представление при помощи теоремы 4.3. Другие корни этого представления имеют индекс $n = 2m + 1$, что может быть, только если нуль является одним из корней. Рассуждая так же, как и при доказательстве утверждения (i), проведенным выше, мы можем установить, что построенное представление, использующее точку $k = \infty$, характеризуется тем свойством, что оно имеет максимальную массу либо в точке 0, либо в точке $\infty$. Из утверждения единственности в теореме 4.3 следует, что любое главное представление включает точки 0 и $\infty$ вместе или не включает их вообще и главное представление, включающее 0 и $\infty$, единственно.

Теперь мы построим второе главное представление, которое, очевидно, не имеет 0 и $\infty$ в качестве корней. Для этого мы введем убывающую последовательность замкнутых выпуклых конусов $\mathcal{M}_{n+1}(\overline{K}_r)$, $r = 0, 1, 2, \ldots$, образованных, как в (4.1), где участвующая в определении область $\overline{K}$ заменена на $\overline{K}_r = [r, r+1, r+2, \ldots, \infty]$. Пусть $r_0$ — единственная величина, удовлетворяющая соотношению

$$\mathbf{c}^0 \in \mathcal{M}_{n+1}(\overline{K}_{r_0}), \qquad \mathbf{c}^0 \notin \mathcal{M}_{n+1}(\overline{K}_{r_0+1}). \tag{5.1}$$

Возможны два случая:

а) $\mathbf{c}^0 \in \mathcal{D}\mathcal{M}_{n+1}(\overline{K}_{r_0}) \quad (r_0 \geqslant 1)$.

В силу теоремы 4.2 мы получаем представление $\mathbf{c}^0$ индекса $n$ относительно $\overline{K}_{r_0}$. Это представление должно включать $r_0$ в качестве корня, так как в противном случае мы получили бы, что $\mathbf{c}_0 \in \mathcal{D}\mathcal{M}_{n+1}(\overline{K})$, что противоречило бы гипотезе. Индекс этого представления относительно $\bar{K}$ не больше чем $n+1$ и поэтому равен $n+1$, так как $c^0 \in \operatorname{Int} \mathcal{M}_{n+1}(\bar{K})$. Более того, индекс может возрасти до $n+1$, только если блок, начинающийся с $r_0$, нечетен. Это представление не может включать $\infty$, так как нуль не является корнем и $n+1$ — четное.

b) $c^0 \in \operatorname{Int} \mathcal{M}_{n+1}(\bar{K}_{r_0})$.

Мы построим каноническое представление $\mathbf{c}^0$ относительно $\mathcal{M}_{n+1}(\bar{K}_{r_0})$, соответствующее построению теоремы 4.3, используя точку $r_0+1$. Все его корни, кроме $r_0+1$, обозначим

$$\Sigma = \{k_1', k_2, \ldots, k_p'\}, \quad k_i' \in \bar{K}_{r_0}. \tag{5.2}$$

Ясно, что $I_{r_0}(\Sigma) \leq n$, где индекс вычисляется обычным путем относительно $\bar{K}_{r_0}$.

Мы утверждаем, что множество $\Sigma$ включает точку $k_1' = r_0$. В противном случае мы могли бы найти представление $c^0$, включающее только точки $\bar{K}_{r_0+1}$, что противоречило бы второму соотношению в (5.1). Множество $\tilde{\Sigma} = (r_0+1, \Sigma)$, очевидно, имеет индекс $n+1$ относительно $\bar{K}$, так как $r_0 \in \Sigma$. Теперь $n+1$ — четное; это означает, что $\infty \notin \tilde{\Sigma}$. Представление, определенное таким образом, дает нам второе главное представление $c^0$.

Теперь мы докажем, что имеется ровно два главных представления для $n = 2m+1$. В силу симметрии достаточно показать, что не может существовать двух различных главных представлений $\sigma$ и $\sigma'$, соответствующие корни которых $\Sigma_1 = \{k_i\}_{i=1}^{p}$, $\Sigma_2 = \{k_i'\}_{i=1}^{q}$ удовлетворяли бы соотношению $k_1' = k_1 \geqslant 1$. (В том случае, когда $k_1' = k_1$, мы предполагаем, не умаляя общности, что $\lambda_1 \geqslant \lambda_1'$, где $\lambda_1'$ и $\lambda_1$ — веса, расположенные в точках $k_1'$ и $k_1$ соответственно).

Расширим множество $\Sigma_1$ $(I(\Sigma_1) = n+1)$ так, чтобы получить множество $\Sigma_1'$, добавляя по точке непосредственно справа к каждому нечетному блоку в $\Sigma_1$ Множество $\Sigma_1'$ тогда имеет $n+1$ элемент и индекс $n+1$. Обозначая $n$ точек $\Sigma_1' - \{k_1\}$ через $s_1, s_2, \ldots$ $\ldots, s_n$, мы образуем многочлен

$$\tilde{u}(k) = \varepsilon U\begin{pmatrix} 0, 1, \ldots, n-1, n \\ s_1, s_2, \ldots, s_n, k \end{pmatrix},$$

где $\varepsilon > 0$. Многочлен $\tilde{u}$ обращается в нуль только в $s_1, s_2, \ldots, s_n$ и является неотрицательным для $k > k_1$ и строго отрицательным при $k \leq k_1$, так как первый блок из $s$ значений нечетен, а все другие блоки четны. Значение $\varepsilon > 0$ выбирается таким образом, чтобы $\tilde{u}(k_1) = -1$.

Если $k_1' > k_1$, то $-\lambda = \int \tilde{u}\, d\sigma = \int \tilde{u}\, d\sigma' \geqslant 0$, что, очевидно, невозможно. Если $k_1' = k_1$, то

$$-\lambda_1 = \int \tilde{u}\, d\sigma = \int \tilde{u}\, d\sigma'' = -\lambda_1' + \varkappa, \tag{5.3}$$

где $\varkappa = \int\limits_{[k_2', \infty]} \tilde{u}\, d\sigma'(k) \geqslant 0$.

Очевидно, что (5.3) возможно только в том случае, когда $\lambda_1 = \lambda_1'$ и $\varkappa = 0$, что означает $\sigma = \sigma'$.

Приведенные рассуждения доказывают следующую теорему.

Теорема 5.1. *Для каждого* $\mathbf{c} \in \operatorname{Int} \mathcal{M}_{n+1}(\overline{K})$ *существует ровно два главных представления. Если* $n = 2m$, *то эти представления характеризуются максимальной массой в* $k = 0$ *и* $k = \infty$ *соответственно. Для* $n = 2m + 1$ *главное представление, включающее как* 0, *так и* $\infty$, *можно охарактеризовать наличием максимальной массы в одной из этих точек.*

Определение 5.1. Главное представление, включающее корень $k = \infty$, называется *верхним*, оставшееся главное представление называется *нижним*.

## § 6. Свойства перемежаемости корней канонических представлений

Теперь мы изучим структуру корней представлений точки $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}(\overline{K})$, обращая особое внимание на свойство перемежаемости корней различных представлений.

Рассмотрим два произвольных различных конечных представления $\mathbf{c}^0$, обозначенные

$$R_1 = \{\lambda_j; k_j\}_{j=1}^{p} \quad \text{и} \quad R_2 = \{\lambda_j'; k_j'\}_{j=1}^{q},$$

т. е.

$$c_i^0 = \sum_{j=1}^{p} \lambda_j u_i(k_j) = \sum_{j=1}^{q} \lambda_j u_i(k_j') \quad (i = 0, 1, \ldots, n).$$

В этом случае

$$0 = \sum_{j=1}^{p} \lambda_j u_i(k_j) - \sum_{j=1}^{q} \lambda_j' u_i(k_j') = \sum_{k=1}^{r} \gamma_k u_i(s_k), \quad i = 0, 1, \ldots, n, \tag{6.1}$$

где $\{s_k\}_{k=1}^{r} = \{k_j\} \cup \{k_j'\}$ расположены в возрастающем порядке, причем при $k_j = k_j'$ и $\lambda_j = \lambda_j'$ соответствующие члены вычеркнуты. Величины $\gamma_k^j \neq 0$ обозначают соответствующие коэффициенты или ненулевые разности коэффициентов. Очевидно, что $r \geqslant n + 2$, в противном случае соотношение (6.1) было бы неверным.

Так как $\gamma_l \neq 0$, $l = 1, 2, \ldots, r$, то из (6.1) следует, что

$$S^-(\{\gamma_k\}_{k=1}^r) \geqslant n + 1. \tag{6.2}$$

Здесь $S^-(\{\gamma_k\}_1^r)$ обозначает число действительных изменений знака в последовательности $\gamma_1, \gamma_2, \ldots, \gamma_r$.

Если предположить, что (6.2) несправедливо, то $S^-(\{\gamma_k\}_1^r) \leqslant n$, и мы можем разбить последовательность $\gamma_1, \ldots, \gamma_r$ $(r \geqslant n+2)$ на $n+1$ группу (см. теорему 4.4 гл. I) $\{\gamma_1, \ldots, \gamma_{\nu_1}\}$, $\{\gamma_{\nu+1}, \ldots, \gamma_{\nu_2}\}, \ldots$ $\ldots, \{\gamma_{\nu_n+1}, \ldots, \gamma_r\}$, где каждый блок имеет строго фиксированный знак, причем в каждой группе по крайней мере один элемент отличен от нуля. Из (6.1) получаем

$$\sum_{k=1}^{n+1} \varepsilon_k \sum_{i=\nu_{k-1}+1}^{\nu_k} |\gamma_i| u_j(s_i) = 0, \quad j = 0, 1, \ldots, n, \tag{6.3}$$

где $\varepsilon_k = \pm 1$, $\nu_0 = 0$ и $\nu_{n+1} = r$. Однако равенство (6.3) не может выполняться, так как определитель системы уравнений (6.3), где $\varepsilon_k$ рассматриваются как неизвестные, имеет значение

$$\sum_{l_0=1}^{\nu_1} \cdots \sum_{l_n=\nu_n+1}^{\nu_{n+1}} |\gamma_{l_0} \ldots \gamma_{l_n}| U\begin{pmatrix} 0, & 1, & \ldots, & n \\ s_{l_0}, & s_{l_1}, & \ldots, & s_{l_n} \end{pmatrix} > 0.$$

Отсюда следует, что соотношение (6.2) должно выполняться.

Вспоминая интерпретацию понятия индекса (см. описание во второй части определения 2.2), мы заключаем, что

$$S^-(\{\gamma_k\}_{k=1}^r) \leqslant \min\{I(\Sigma_1),\ I(\Sigma_2)\}, \tag{6.4}$$

где

$$\Sigma_1 = \{k_j\}_{j=1}^p \quad \text{и} \quad \Sigma_2 = \{k_j'\}_{j=1}^q. \tag{6.5}$$

Комбинация соотношений (6.2) и (6.4) дает нам следующую теорему.

Теорема 6.1. *Пусть* $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}(\overline{K})$, *и пусть* $R_1 = \{\lambda_j;\ k_j\}_{j=1}^p$ *и* $R_2 = \{\lambda_j';\ k_j'\}_{j=1}^q$ — *два различных представления* $\mathbf{c}^0$. *Пусть* $(\gamma_1, \gamma_2, \ldots, \gamma_r)$ — *последовательность, определяемая соотношением* (6.1) *и* $\Sigma_1, \Sigma_2$ *заданы соотношением* (6.5); *тогда*

$$n + 1 \leqslant S^-\{(\gamma_1, \gamma_2, \ldots, \gamma_r)\} \leqslant \min\{I(\Sigma_1);\ I(\Sigma_2)\}. \tag{6.6}$$

Следствие 6.1. *Если* $R_1$ *или* $R_2$ — *канонические меры, то*

$$n + 1 \leqslant S^{-1}\{(\gamma_1, \gamma_2, \ldots, \gamma_r)\} \leqslant n + 2. \tag{6.7}$$

Теперь предположим, что $R_1$ — главное представление, а $R_2$ — произвольное представление $\mathbf{c}^0$. Тогда из (6.6) следует, что

$$S^-(\{\gamma_k\}_1^r) = n + 1. \tag{6.8}$$

В частности, пусть

$$S = \{k_{j_0}, k_{j_0}+1, \ldots, k_{j_0}+q\} \tag{6.9}$$

— последовательное множество корней, принадлежащих представлению $R_1$, где $k_{j_0}-1$ и $k_{j_0}+q+1$ принадлежат $\overline{K}$, но не является корнями $R_1$. Очевидно,

$$I(S) = \begin{cases} q+2, & \text{если } q \text{ — четное,} \\ q+1, & \text{если } q \text{ — нечетное.} \end{cases}$$

Для четного $q$ мы, в соответствии с (6.8) и (6.9), должны получить

$$S^-(\gamma_{l_0-1}, \gamma_{l_0}, \ldots, \gamma_{l_0+q+1}) = q+2, \tag{6.10}$$

где $l_0$ — индекс, для которого $s_{l_0} = k_{j_0}$, в то время как для нечетного $q$ мы должны получить

$$S^-(\gamma_{l_0}, \ldots, \gamma_{l_0+q+1}) = q+1,$$

либо

$$S^-(\gamma_{l_0-1}, \ldots, \gamma_{l_0+q}) = q+1.$$

Такой же вывод остается в силе и когда $k_{j_0}=0$ или $k_{j_0}=\infty$. Соотношения (6.8), (6.10) и (6.11) выражают свойство перемежаемости корней главного и произвольного представлений. Конечно, если $R_1$ и $R_2$ различные главные представления, то каждое перемежается с другим в вышеуказанном смысле.

## § 7. Теорема Маркова — Крейна

Теперь мы рассмотрим возможность расширить $T$-систему $\{u_i\}_0^n$, добавляя к ней дополнительно функцию $u_{n+1}(k) = \Omega(k)$, определенную на $\overline{K}$ так, чтобы $\{u_i\}_0^{n+1}$ образовывали $T$-систему.

Пусть $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}(\overline{K})$. Множество всех мер $\sigma$, которые представляют $\mathbf{c}^0$, обозначается $\mathbf{V}(\mathbf{c}^0)$. Тогда множество

$$\Gamma = \{\gamma \mid \gamma = \int_{\overline{K}} \Omega(k)\, d\sigma(k),\ \sigma \in V(\mathbf{c}^0)\} \tag{7.1}$$

является ограниченным невырожденным интервалом $\Gamma = \{\underline{\gamma} \leqslant \gamma \leqslant \leqslant \overline{\gamma}\}$. Легко доказать, что верхнее значение $\overline{\gamma}$ достигается только для $\sigma = \overline{\sigma}$ — верхнего главного представления $\mathbf{c}^0$, а $\underline{\gamma}$ достигается только для $\sigma = \underline{\sigma}$ — нижнего главного представления.

Теперь мы усилим наши предположения об исходных функциях так, чтобы они удовлетворяли соотношениям:

Условие 7.1.

(а) $u_0, u_1, \ldots, u_p, u_{n+1}$ и $u_0, u_1, \ldots, u_p$ являются $T$-системами на $\overline{K}$ для каждого $p = 0, 1, \ldots, n$;

(в) $u_{n+1}(k) > 0$ для $k \in \overline{K}$.

При этих предположениях мы можем доказать соответствующую теорему Маркова — Крейна для дискретных множеств (см. § 2 гл. III).

Нам потребуется следующая лемма (доказательство ее, по существу такое же, как и доказательство леммы 6.1 гл. I). Фактически эту лемму можно рассматривать как следствие из леммы 6.1 гл. I).

Лемма 7.1. *Если выполнено условие* (7.1), *то*

$$(-1)^r \begin{vmatrix} u_0(k_0) & \ldots & u_0(k_{n-r+1}) & u_0(k_{n-r+2}) & \ldots & u_0(k_{n+1}) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ u_n(k_0) & \ldots & u_n(k_{n-r+1}) & u_n(k_{n-r+2}) & \ldots & u_n(k_{n+1}) \\ \Omega(k_0) & \ldots & \Omega(k_{n-r+1}) & 0 & \ldots & 0 \end{vmatrix} =$$

$$= (-1)^r N_r > 0, \qquad (7.2)$$

*где* $r$ *определяет число нулей в последней строке и*, $0 \leqslant k_0 < k_1 < \ldots \ldots < k_{n+1} \leqslant \infty$.

Теорема 7.1. *Пусть выполнено условие* (7.1), *и пусть* $\mathbf{c}^0 = (c_0^0, \ldots, c_n^0) \in \operatorname{Int} \mathcal{M}_{n+1}(\overline{K})$. *Если для каждого целого* $k_0 \in (0, \infty)$ $\sigma_{k_0}$ *обозначает каноническое представление* $\mathbf{c}^0$, *построенное в теореме* 4.3, *а* $\sigma \in V(\mathbf{c}^0)$ — *другое представление, то*

$$\int_0^{k_0^-} \Omega(k)\, d\sigma_{k_0}(k) < \int_0^{k_0^-} \Omega(k)\, d\sigma(k) \leqslant \int_0^{k_0^+} \Omega(k)\, d\sigma(k) < \int_0^{k_0^+} \Omega(k)\, d\sigma_{k_0}(k), \tag{7.3}$$

*за исключением того случая, когда* $k_0$ *является первым корнем* $\sigma_{k_0}$, *а* $\sigma_{k_0}$ *имеет индекс* $n+2$. *В этом случае оба члена в левой части равны нулю и, следовательно, в левой части неравенства имеет место равенство.*

Доказательство. Пусть $\Sigma$ обозначат множество корней $\sigma_{k_0}$, за исключением точки $k_0$. Множество $\Sigma$ имеет индекс $n$ или $n-1$ в зависимости от того, имеет $\sigma_{k_0}$ индекс $n+2$ или $n+1$.

Мы дополним множество $\Sigma$ до множества $\Sigma_0$, добавляя по точке непосредственно после каждого нечетного блока, расположенного справа от $k_0$, и по точке непосредственно перед каждым нечетным блоком, расположенным слева от $k_0$. Если $\sigma_{k_0}$ имеет индекс $n+2$, то $\Sigma_0$ содержит $n$ точек, если же $\sigma_{k_0}$ имеет индекс $n+1$, то $\Sigma_0$ содержит $n-1$ точку. В последнем случае для четного $n$ мы всегда можем добавить либо $k=0$, либо $k=\infty$, а для нечетного $n$ можем добавить минимальную точку, не входящую в $\Sigma_0$. Получившееся множество, которое мы обозначим $\Sigma_1$, имеем $n$ точек, не содержит $k_0$ и все его точки, за исключением точек $k=0$ и $k=\infty$, при движении от точки $k_0$ можно разбить на последовательные пары.

Теперь мы рассмотрим правую часть неравенства в (7.3). Пусть $k_1 < k_2 < \ldots < k_h < k_0 < l_1 < \ldots l_q - n + 1$ точек $\Sigma_1 \bigcup \{k_0\}$ и рассмотрим многочлен $u^{**}$ (см. теорему 6.2 гл. I), определяемый системой уравнений

$$\sum_{i=0}^{n} a_i u_i(k) = \begin{cases} \Omega(k), & k = k_1, \ldots, k_h, k_0, \\ 0, & k = l_1, \ldots, l_q. \end{cases} \tag{7.4}$$

У этой системы нетривиальное решение всегда существует, так как $\{u_i\}_0^n$ — $T$-система. Многочлен $u^{**}$ обладает следующим свойством:

$$u^{**}(k) \geqslant \Omega(k), \quad k \in [0, k_0], \tag{7.5}$$

причем равенство справедливо только для $k_1, \ldots, k_h, k_0$,

$$u^{**}(k) \geqslant 0, \quad k \in [k_0, \infty],$$

причем равенство справедливо только для $l_1, \ldots, l_q$. Чтобы убедиться в этом, мы сначала предположим, что $u^{**}(l) \leqslant \Omega(l)$ для некоторого $l \in [0, k_0]$, не равного $k_1, \ldots, k_h, k_0$. Если $k_h < l < k_0$, то мы, применяя лемму 7.1, получим

$$(-1)^q \begin{vmatrix} u_0(k_1) \ldots u_0(k_h) & u_0(l) & u_0(k_0) & u_0(l_1) \ldots u_0(l_q) \\ \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot & \cdot \; \cdot & \cdot \; \cdot \; \cdot & \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \; \cdot \\ u_n(k_1) \ldots u_n(k_h) & u_n(l) & u_n(k_0) & u_n(l_1) \ldots u_n(l_q) \\ \Omega(k_1) \ldots \Omega(k_h) & \Omega(l) & \Omega(k_0) & 0 \quad \ldots \quad 0 \end{vmatrix} > 0. \tag{7.6}$$

Если вычесть из последней строки линейную комбинацию первых $n+1$ строк $\sum_0^n a_i u_i$, то единственным ненулевым членом в последней строке останется $\Omega(l) - u^{**}(l)$, который стоит в $(h+1)$-м столбце. Производя разложение по последней строке, получим

$$(-1)(\Omega(l) - u^{**}(l)) U\begin{pmatrix} 0, & 1, & \ldots, & n \\ k_1, & k_2, & \ldots, & l_q \end{pmatrix} > 0.$$

Это неравенство, однако, противоречит предположению о том, что $\Omega(l) - u^{**}(l) \geqslant 0$.

Другие случаи при $l \in [0, k_0]$ рассматриваются аналогично, если заметить, что столбец с аргументом $l$ в (7.6) всегда сдвигается на четное число. Доказательство того, что $u^{**}(k) > 0$ на $(k_0, \infty)$, за исключением $l_1, \ldots, l_q$, проводится при помощи таких же рассуждений.

Используя многочлен $u^{**}$, мы получаем

$$\int\limits_{[0, k_0]} \Omega(k)\, d\sigma_{k_0}(k) = \int\limits_{[0, k_0]} u^{**}(k)\, d\sigma_{k_0}(k) = \int\limits_{[0, \infty]} u^{**}(k)\, d\sigma_{k_0}(k) =$$

$$= \int\limits_{[0, \infty]} u^{**}(k)\, d\sigma(k) = \int\limits_{[0, k_0]} u^{**}(k)\, d\sigma(k) + \int\limits_{(k_0, \infty]} u^{**}(t)\, d\sigma(t). \tag{7.7}$$

Из неравенств, которым удовлетворяет $u^{**}$, следует

$$\int\limits_{[0,k_0]} u^{**}d\sigma \geqslant \int\limits_{[0,k_0]} \Omega\, d\sigma \quad \text{и} \quad \int\limits_{(k_0,\infty)} u^{**}d\sigma \geqslant 0. \tag{7.8}$$

Далее, так как мера $\sigma$ не может быть сосредоточена только на множестве $\Sigma_1 \cup k_0$, которое имеет $n+1$ точек, то строгое неравенство должно иметь место хотя бы для одного из неравенств (7.8). Поэтому, используя (7.7), получим неравенство в правой части (7.3).

Чтобы доказать неравенство в левой части (7.3), предположим сначала, что $k_0$ не является первым корнем $\sigma_{k_0}$. Построим многочлен $u^*$ такой же, как и в (7.4), за исключением того, что $u^*(k_0)=0$. Тогда доказательство можно провести точно так же, как и раньше, если использовать тот факт, что $u^*$ удовлетворяет неравенствам

$$u^*(k) \leqslant \Omega(k), \quad k \in [0, k_0)$$

(причем равенство справедливо только в точках $k_1, \ldots, k_h$),

$$u^*(k) \leqslant 0, \quad k \in [k_0, \infty]$$

(причем равенство справедливо только в точках $k_0, l_1, \ldots, l_q$).

Случай, когда $k_0$ — первый корень $\sigma_{k_0}$, можно рассмотреть отдельно. Если индекс $\sigma_{k_0}$ равен $n+2$, то $\mathbf{c}^0$ может быть внутренней точкой для $\mathcal{M}_{n+1}[k_0, \infty]$ так что, возможно, существуют меры, отличные от $\sigma_{k_0}$, которые представляют $\mathbf{c}^0$ и не имеют массы на $[0, k_0)$. Если $\sigma_k$ имеет индекс $n+1$, то $\mathbf{c}^0$ — граничная точка $\mathcal{M}_{n+1}[k_0, \infty]$ и описанная выше ситуация невозможна.

## § 8. Дополнения

Как отмечалось во введении, результаты предыдущих параграфов можно применять к системе функций, определенных только на $K$, которые удовлетворяют следующему условию.

Условия 8.1.

(a) $u_n(k) > 0$ для $k > \hat{k}$,

(b) $\lim\limits_{k\to\infty} u_i(k)/u_n(k) = \delta_{in}$,

(c) $\{u_i\}_0^{n-1}$ и $\{u_i\}_0^n$ — $T$-системы на $K$. Определим

$$v_i(k) = \begin{cases} \dfrac{u_i(k)}{w(k)}, & k \in K, \\ \delta_{in}, & k = \infty, \end{cases}$$

где $w(k)$ положительна на $K$ и $w(k) = u_n(k)$ для $k \geqslant \hat{k}$. Тогда $\{v_i\}_0^n$ — $T$-система на $\overline{K}$. Применяя методы гл. V, можно вывести следующие утверждения:

(1) $\mathcal{M}_{n+1}^{(u)}(K) = \mathcal{C}(C_{n+1}(K))$, $\mathcal{M}_{n+1}^{(u)}(K) = \mathcal{M}_{n+1}^{(v)}(K)$ и $\mathcal{M}_{n+1}^{(u)}(K) = \mathcal{M}_{n+1}^{(u)}(K) + R_\infty$, $R_\infty = \{(0, \ldots, 0, \lambda) \mid \lambda \geqslant 0\}$.

2) Для каждой точки $\mathbf{c}^0 \in \operatorname{Int} \mathcal{M}_{n+1}^{(u)}(K)$ существует два главных представления индекса $n+1$. «Нижнее главное» представление является настоящим представлением, в то время как верхнее главное представление порождает только моменты $\mathbf{c}_0^0, \ldots, \mathbf{c}_{n-1}^0$, если не обращать внимания на массу в $k=\infty$. Для каждого $k_0 \in (0, \infty)$ существует «каноническое представление» $\mathbf{c}^0$, которое имеет индекс $n+1$ или $n+2$ и которое включает точку $k_0$.

(3) Если $\mathbf{c}^0 \in \operatorname{Int} M_{n+1}^{(u)}(K)$ и функция $\Omega(k) = u_{n+1}(k)$ определена на $K$ таким образом, что $\{u_i\}_0^{n+1}$ является $T$-системой, то множество

$$\Gamma = \left\{\gamma \mid \gamma = \int\limits_K u_{n+1} d\sigma, \ \sigma \in \mathbf{V}(\mathbf{c}^0)\right\}$$

ограничено снизу и минимальное значение $\underline{\gamma}$ достигается только на нижнем главном представлении $\underline{\sigma}$ точки $\mathbf{c}^0$.

## § 9. Другой класс экстремальных задач, решениями которых являются канонические меры

В этом параграфе мы рассмотрим ряд экстремальных задач, в которых экстремальные меры являются каноническими мерами. В рассуждениях теоремы Маркова — Крейна изучаемое множество состояло из мер, моменты которых относительно $u_0, u_1, \ldots, u_n$ были фиксированы. В этом параграфе соответствующие множества будут состоять из мер, моменты которых принадлежат некоторому подмножеству, которое не обязано быть одной точкой моментного пространства.

Излагаемый здесь материал аналогичен материалу в § 5 гл. III и в § 6 гл. V. Будет видно, что, по существу, утверждения теорем здесь такие же, как и в гл. V, с одним лишь исключением, что интегралы распространяются теперь на множество $K$ вместо полубесконечного интервала $[0, \infty)$. За исключением леммы 9.1, доказательства теорем повторяют доказательства, приведенные в гл. V и будут опущены. Лемма 9.1 требует более тонкого анализа из-за дискретности множества $K$. Мы сосредоточим внимание в этом параграфе на системах типа тех, которые обсуждались в § 8. Результаты, соответствующие случаю компактного интервала, которые обсуждались в § 5, гл. III, также могут быть получены для $T$-систем на $\overline{K}$.

Пусть $u_0, \ldots, u_n, u_{n+1}$ — $n+2$ функции, определенные на неотрицательных целых числах, удовлетворяющих следующему условию.

Условие 9.1.

(a) $\{u_i\}_0^n$ и $\{u_i\}_0^{n+1}$ являются обе $T$-системами,

(b) для $i = 0, 1, \ldots, n+1$ определители $(n+1)$-го порядка, образованные из $u_0, \ldots, u_{i-1}, u_{i+1}, \ldots, u_{n+1}$ все имеют один знак, обозначаемый $\varepsilon_i$ ($\varepsilon_i = \pm 1$) (см. условие 5.1 (b) гл. III),

(c) $\lim\limits_{k\to\infty} \dfrac{u_i(k)}{u_{n+1}(k)} = 0, \quad i = 0, 1, \ldots, n,$

(d) $u_{n+1}(k) > 0$ для $k = 0, 1, 2, \ldots$

Мы ограничимся рассмотрением мер $\sigma$, сконцентрированных на неотрицательных целых числах, для которых $\int u_{n+1} d\sigma < \infty$.

Теорема 9.1. *Пусть $A$, $B$ и $C$ определяют разбиение множества целых чисел $0, 1, \ldots, n$. Если $\{u_i\}_0^{n+1}$ удовлетворяет условию 9.1 и класс мер, удовлетворяющих условию*

$$\int u_i d\sigma \begin{cases} = d_i, & i \in A, \\ \leqslant d_i, & i \in B, \\ \geqslant d_i, & i \in C, \end{cases}$$

*не пуст, то*

$$\underline{\alpha} = \min_{\sigma \in V} \int u_{n+1} d\sigma$$

*достигается в $V$ только на мере $\underline{\sigma}$ индекса, не большего чем $n+1$.*

Доказательство единственности $\sigma$ в теореме 9.1 основывается на следующей лемме. Мы сопоставим каждой моментной точке $\mathbf{c} = (c_0, \ldots, c_n)$, порожденной $u_0, u_1, \ldots, u_n$, величину

$$\underline{c}_{n+1}(\mathbf{c}) = \inf_{\sigma \in V(\mathbf{c})} \int u_{n+1} d\sigma.$$

Лемма 9.1. *Если $u_0, \ldots, u_{n+1}$ удовлетворяет условиям* 9.1 (a), (c) *и* (d) *и условию* 9.1 (b) *при фиксированном $i$, то $(-1)^{n-i}\varepsilon_i \underline{c}_{n+1}(\mathbf{c})$ строго возрастает по $c_i$.*

Доказательство. Пусть $\sigma$ и $\sigma'$ — представления индекса, не большего чем $n+1$, такие, что $\int u_{n+1} d\sigma = \underline{c}_{n+1}(\mathbf{c})$ и $\int u_{n+1} d\sigma' = \underline{c}_{n+1}(\mathbf{c}')$, где $\mathbf{c} = (c_0, \ldots, c_{i-1}, c_i, c_{i+1}, \ldots, c_n)$, $\mathbf{c}' = (c_0, \ldots, c_{i-1}, c'_i, c_{i+1}, \ldots, c_n)$ и $c'_i > c_i$.

Пусть $\Sigma$ и $\Sigma'$ обозначают множества корней $\sigma$ и $\sigma'$ соответственно, и пусть $s_1, \ldots, s_r$—совокупность всех корней, расположенных в порядке возрастания. Тогда

$$\int u_i d\sigma - \int u_i d\sigma' = \sum_{j=1}^{r} \gamma_j u_i(s_j), \quad i = 0, 1, \ldots, n,$$

где $\gamma_j$, $j = 0, 1, \ldots, r$, веса $\sigma$ либо $\sigma'$ с соответствующим знаком.

Ясно, что $S^-(\gamma) \leqslant \min\{I(\Sigma), I(\Sigma')\}$, (см. (6.4)). Здесь $S^-(\gamma)$ обозначает число изменений знака в последовательности $\{\gamma_i\}_1^r$. Мы утверждаем, что $S^-(\gamma) \leqslant n$, так как, если бы $S^-(\gamma) = n+1$, то одно из чисел $I(\Sigma)$ или $I(\Sigma')$ превышало бы $n+1$. Чтобы убедиться в этом, предположим, что $S^-(\gamma) = n+1$. Тогда последовательность $\gamma_1, \ldots, \gamma_r$ можно разделить на $n+2$ блоков, для которых $\gamma_i$ в одном блоке имеют один и тот же знак и по крайней мере один элемент в каждом блоке не равен нулю, причем последовательные блоки имеют противоположные знаки. Не умаляя общности, мы можем предположить, что первый блок положителен. Теперь, если $n+1=2m$, то существует по крайней мере $m+1$ положительных $\gamma_i$, которые разделены отрицательными $\gamma_i$. В этом случае каждое положительное $\gamma_i$ вкладывает две единицы в индекс $\Sigma$, исключая, возможно, первое из них, так что $I(\Sigma)\geqslant 2m+1=n+1$. Если $n+1=2m+1$, то имеется $m+1$ отрицательных $\gamma_i$, вкладывающих по две единицы в индекс $\Sigma$, так что $I(\Sigma')\geqslant n+2$. Поэтому $S^-(\gamma)\leqslant n$.

Теперь пусть $q$ таково, что $S^-(\gamma)=q-1\leqslant n$. Разобъем $\{\gamma_j\}_1^r$ на $q$ блоков $\{\gamma_1,\ldots,\gamma_{\nu_1}\}$, $\{\gamma_{\nu_1+1}\ldots,\gamma_{\nu_2}\},\ldots,\{\gamma_{\nu_{q-1}+1},\ldots,\gamma_{\nu_q}\}$, удовлетворяющих условиям, сформулированным выше. Тогда

$$\sum_{j=1}^{r}\gamma_j u_i(s_j)=\sum_{p=0}^{q-1}(-1)^p v_{ip},\qquad i=0,1,\ldots,n+1,$$

где $v_{ip}=\sum\limits_{j=\nu_p+1}^{\nu_{p+1}}|\gamma_j|\,u_i(s_j),\qquad p=0,\ldots,q-1\qquad(\nu_0=0,\ \nu_q=r).$

Рассмотрим определитель

$$\begin{vmatrix} 0 & v_{00} & \cdots & v_{0n}\\ \vdots & & & \\ & v_{10} & \cdots & v_{1n}\\ & \vdots & & \vdots\\ c_i-c_i' & \vdots & & \vdots\\ 0 & & & \\ \vdots & \vdots & & \vdots\\ 0 & & & \\ \underline{c}_{n+1}-\underline{c}_{n+1}' & v_{n+1,0} & & v_{n+1,n}\end{vmatrix},$$

где $v_{ip}=u_i(s_p')$ для $p=q,q+1,\ldots,n$ и $s_n'>s_{n-1}'>\ldots>s_q'>s_r$. Так как первый столбец линейно зависит от остальных столбцов, то этот определитель равен нулю. Заметим также (используя разложение по столбцу), что минор $c_i-c_i'$, умноженный на множитель $\varepsilon_i$, и минор $\underline{c}_{n+1}-\underline{c}_{n+1}'$ оба положительны. Отсюда можно заключить, что из $c_i-c_i'<0$ следует, что $(-1)^{n-i}\varepsilon_i(\underline{c}_{n+1}-\underline{c}_{n+1}')<0$, что заканчивает доказательство.

Следствие 9.1 (а) *Сохраняя обозначения теоремы* 9.1 *и вводя обозначение* $\underline{D}=\{i\mid\int u_i d\underline{\sigma}=d_i,\ 0\leqslant i\leqslant n\}$, *мы получим, что*

*$\underline{\sigma}$ является единственной мерой в $V_1$, на которой достигается минимум*

$$\underline{\beta} = \min_{\sigma \in V_1} \int u_{n+1} d\sigma,$$

*где $V_1$ — класс мер $\sigma$, удовлетворяющих условию*

$$\int u_i d\sigma \begin{cases} = d_i, & i \in A \cap \underline{D}, \\ \leqslant d_i, & i \in \underline{B} \cap \underline{D}, \\ \geqslant d_i, & i \in C \cap \underline{D}. \end{cases}$$

Предположим, что $v_0, \ldots, v_n$ — другое множество функций, определенных на неотрицательных целых числах таких, что $v_0, \ldots$ $\ldots, v_n, u_{n+1}$ удовлетворяет условию 9.1. Тогда получим

Следствие 9.1 (b) *Если $\underline{\sigma}$, $A$, $B$, $C$ и $\underline{D}$ определены как в следствии 9.1 (a) и $e_i = \int v_i d\underline{\sigma}$, $0 \leqslant i \leqslant n$, то $\underline{\sigma}$ — единственная мера в $W_1$, на которой достигается*

$$\underline{\gamma} = \min_{\sigma \in W_1} \int u_{n+1} d\sigma,$$

*где $W_1$ — класс мер $\sigma$, удовлетворяющих условию*

$$\int v_i d\sigma \begin{cases} = e_i, & i \in A \cap \underline{D}, \\ \leqslant e_i, & i \in B \cap \underline{D}, \\ \geqslant e_i, & i \in C \cap \underline{D}. \end{cases}$$

Предположим теперь, что мы можем дополнить систему $\{u_i\}_0^{n+1}$ функцией $v_i$ так, чтобы образовавшееся множество функций удовлетворяло следующим условиям:

Условие 9.2.

(a) $\{u_0, \ldots, u_{i-1}, v_i, u_i, \ldots, u_n\}$ и $\{u_0, \ldots, u_{i-1}, v_i, u_i, \ldots, u_{n+1}\}$ являются обе $T$-системами,

(b) определители $(n+1)$-го порядка, образованные из любого фиксированного упорядоченного набора $n+1$ функций, выбранных из $\{u_0, \ldots, u_{i-1}, v_i, u_i, \ldots, u_n\}$, имеют строго один и тот же знак (см. условие 5.1 (b) гл. III),

(c) $\lim\limits_{k \to \infty} [v_i(k)]/[u_{n+1}(k)] = 0$.

Теорема 9.2. *Если функции $u_0, \ldots, u_{i-1}, v_i, u_i, \ldots, u_{n+1}$ удовлетворяют условию 9.2 и условиям 9.1 (a), (c) и (d), а $\underline{\sigma}$ и $V$ определены как в теореме 9.1, то*

$$\min_{\sigma \in V} (-1)^{n-i-1} \int v_i d\sigma = \underline{\delta}_i$$

*достигается только на мере $\underline{\sigma}$.*

## § 10. Применения к теории интерполяции абсолютно монотонных функций

Предположим, что функции $u_i$, $i = 0, 1, \ldots, n$, порождены функцией $\varphi(x, k)$, $0 < x \leqslant b < \infty$, $k = 0, 1, 2, \ldots$, при помощи соотношения $u_i(k) = \varphi(x_i, k)$, где $0 < x_0 < x_1 < \ldots < x_n \leqslant b$. Предполагаем, что определители

$$\Phi\begin{pmatrix} x_1, \ldots, x_p \\ k_1, \ldots, k_p \end{pmatrix} = \det\left[\| \varphi(x_i, k_j) \|_{i,j=1}^{p}\right] \tag{10.1}$$

положительны при любом выборе $0 \leqslant k_1 < k_2 < \ldots < k_p$, $0 < x_1 < \ldots$ $\ldots < x_p \leqslant b$ и при произвольном $p$. Дополнительно предположим, что

$$\lim_{k \to \infty} \frac{\varphi(x, k)}{\varphi(b, k)} = 0 \tag{10.2}$$

для $0 < x < b$.

В этом параграфе мы изучаем различные аспекты интерполяции функции

$$f(x) = \sum_{k=0}^{\infty} \varphi(x, k)\, a_k = \int \varphi(x, k)\, d\sigma(k), \quad a_k \geqslant 0, \tag{10.3}$$

в определенным образом выбранных точках $x_1 < x_2 < \ldots < x_p$.

Основные приложения относятся к важному случаю абсолютно монотонных функций, когда

$$\varphi(x, k) = x^k, \qquad k = 0, 1, 2, \ldots, \qquad 0 < x \leqslant b. \tag{10.4}$$

При этом специальном выборе $\varphi(x, k)$ (10.3) принимает вид

$$f(x) = \sum_{k=0}^{\infty} a_k x^k, \quad a_k \geqslant 0, \quad \sum_{k=0}^{\infty} a_k b^k < \infty. \tag{10.5}$$

Прежде чем начинать рассуждения, связанные с задачей интерполяции для абсолютно монотонных функций, мы сформулируем несколько положений для общего случая (10.3).

Выбрав две точки $0 < x_1 < x_2 \leqslant b$, мы положим $V = \left\{\sigma \mid \int\limits_K \varphi(x_i, k)\, d\sigma(k) = c_i,\ i = 1, 2\right\}$. Величины $c_1$ и $c_2$ считаются фиксированными, и предполагается, что $V$ не пусто. Единственная мера $\underline{\sigma}$ индекса, не большего чем 2, соответствующая $(c_1, c_2)$, очевидно, имеет веса $a_l > 0$ и $a_{l+1} \geqslant 0$ в двух последовательных точках $l$ и $l + 1$. Веса $a_l$ и $a_{l+1}$ удовлетворяют условиям

$$c_1 = a_l \varphi(x_1, l) + a_{l+1} \varphi(x_1, l + 1);$$

$$c_2 = a_l \varphi(x_2, l) + a_{l+1} \varphi(x_2, l + 1).$$

Так как $a_l > 0$ и $a_{l+1} \geqslant 0$, то $l$ должно удовлетворять соотношению

$$\frac{\varphi(x_2, l)}{\varphi(x_1, l)} \leqslant \frac{c_2}{c_1} < \frac{\varphi(x_2, l+1)}{\varphi(x_1, l+1)}. \tag{10.6}$$

Лемма 10.1. *Если* $f(x) = \int \varphi(x, k)\, d\sigma(k)$, $0 < x_1 < x_2 \leqslant b$ *и*

$$P(x) = -\frac{\begin{vmatrix} \varphi(x_1, l) & \varphi(x_1, l+1) & f(x_1) \\ \varphi(x_2, l) & \varphi(x_2, l+1) & f(x_2) \\ \varphi(x, l) & \varphi(x, l+1) & 0 \end{vmatrix}}{\Phi\begin{pmatrix} x_1, & x_2 \\ l, & l+1 \end{pmatrix}},$$

*где* $l$ *определяется соотношением* (10.6) *с* $c_2 = f(x_2)$ *и* $c_1 = f(x_1)$, *то* $f(x) \leqslant P(x)$ *для* $x \in (x_1, x_2)$ *и* $f(x) \geqslant P(x)$ *для* $x \in (0, x_1) \cup (x_2, b]$. *Более того, если имеет место равенство для некоторого* $x \neq x_1$ *или* $x_2$, *то* $P(x) \equiv f(x)$.

Доказательство. Этот результат получается простым применением теоремы 9.2, если положить $\varphi(x_{i+1}, k) = u_i(k)$, $i = 0, 1$; $u_2(k) = \varphi(b, k)$ и

$$\varphi(x, k) = \begin{cases} v_0(k), & x \in (0, x_1); \\ v_1(k), & x \in (x_1, x_2); \\ v_2(x), & x \in (x_2, b]. \end{cases}$$

Если рассмотреть аналогичную ситуацию, когда заданы $(c_1, c_2, c_3)$, то при условии, что $(c_1, c_2, c_3)$ находятся во внутренней части моментного пространства, порожденного функциями $\varphi(x_1, k)$, $\varphi(x_2, k)$ и $\varphi(x_3, k)$ $(0 < x_1 < x_2 < x_3 \leqslant b)$, мера $\sigma$ индекса 3 имеет веса $a_0$, $a_l$, $a_{l+1}$ в точках $0, l, l+1$ $(l \geqslant 1)$, причем $a_{l+1} > 0$, $a_l \geqslant 0$, $a_0 > 0$. Целое $l$ определяется из условий

$$\begin{vmatrix} c_1 & \varphi(x_1, 0) & \varphi(x_1, l) \\ c_2 & \varphi(x_2, 0) & \varphi(x_2, l) \\ c_3 & \varphi(x_3, 0) & \varphi(x_3, l) \end{vmatrix} > 0 \geqslant \begin{vmatrix} c_1 & \varphi(x_1, 0) & \varphi(x_1, l+1) \\ c_2 & \varphi(x_2, 0) & \varphi(x_2, l+1) \\ c_3 & \varphi(x_3, 0) & \varphi(x_3, l+1) \end{vmatrix}.$$

Тогда имеет место результат, аналогичный лемме 10.1. Мы проинтерпретируем этот результат в случае абсолютно монотонных функций (10.5).

Степенью функции (10.5) или (10.3) мы называем индекс множества $\{k \mid k \geqslant 0,\ a_k > 0\}$.

Лемма 10.2. *Если* $f(x)$ — *абсолютно монотонная функция на* $(0, b]$ *степени, большей чем* 3, $x_1$, $x_2$, $x_3$ *удовлетворяют нера-*

*венствам* $0 < x_1 < x_2 < x_3 \leqslant b$ *и если*

$$P(x) = \frac{\begin{vmatrix} 1 & x^l & x^{l+1} & 0 \\ 1 & x_1^l & x_1^{l+1} & f(x_1) \\ 1 & x_2^l & x_2^{l+1} & f(x_2) \\ 1 & x_3^l & x_3^{l+1} & f(x_3) \end{vmatrix}}{\Phi\begin{pmatrix} x_1, & x_2, & x_3 \\ 0, & l, & l+1 \end{pmatrix}},$$

*где* $l$ *определяется соотношениями*

$$\frac{f(x_2)(x_3^l - x_1^l) - f(x_1)(x_3^l - x_2^l)}{x_2^l - x_1^l} < f(x_3) \leqslant$$

$$\leqslant \frac{f(x_2)(x_3^{l+1} - x_1^{l+1}) - f(x_1)(x_3^{l+1} - x_2^{l+1})}{x_2^{l+1} - x_1^{l+1}}, \qquad (10.7)$$

*то* $f(x) > P(x)$ *для* $x \in (x_1, x_2) \cup (x_3, b]$ *и* $f(x) < P(x)$ *для* $x \in (0, x_1) \cup \cup (x_2, x_3)$.

Неравенства (10.7) значительно упрощаются, если $x_1$, $x_2$ и $x_3$ образуют геометрическую прогрессию, т. е. если $x_2 = \rho x_1$, $x_3 = \rho^2 x_1$ для некоторого $\rho > 1$. В этом случае (10.7) примет вид

$$\rho^l < \frac{f(x_3) - f(x_2)}{f(x_2) - f(x_1)} \leqslant \rho^{l+1}.$$

Лемму 10.2 можно сформулировать в более общем виде (см. § 7 гл. V).

Теорема 10.1. *Если* $f(x)$ — *абсолютно монотонная функция на* $(0, b]$ *степени, большей чем* $k$, *и* $0 < x_1 < x_2 < \ldots < x_k \leqslant b$, *то существует абсолютно монотонная функция* $P(x)$ *степени* $k$ *такая, что*

$$f(x_i) = P(x_i), \ i = 1, \ldots, k,$$

$$P(x) < f(x) \text{ для } x \in (x_k, b] \cup (x_{k-2}, x_{k-1}) \cup \ldots,$$

$$P(x) > f(x) \text{ для } x \in (x_{k-1}, x_k) \cup (x_{k-3}, x_{k-2}) \cup \ldots$$

Замечание 10.1. Приведенную выше лемму можно применить к функции

$$[M_2(r)]^2 = \frac{1}{2\pi}\int_0^{2\pi} |F(re^{i\theta})|^2 d\theta,$$

которая абсолютно монотонна относительно $r^2$ для $0 \leqslant r^2 \leqslant b$, при условии, что $F(z)$ аналитична в области $|z| \leqslant b$.

Мы приведем еще две интересные теоремы из теории абсолютно монотонных функций.

Теорема 10.2. *Предположим, что* $0 < x_2 < x_1 < \ldots < x_n = 1$ *и* $c_0, c_1, \ldots, c_n$ *таковы, что существует по крайней мере одна* *абсолютно монотонная функция степени не меньшей, чем* $n+1$ *на* $(0, 1]$ *удовлетворяющая условиям:*

$$f(x_i) = c_i, \qquad i = 0, 1, \ldots, n. \tag{10.8}$$

*Тогда для каждого* $N = 0, 1, 2, \ldots$ *существует единственный многочлен*

$$P_N(x) = \sum_{j=1}^{n(N)} \mu_j x^j, \quad \mu_j \geqslant 0, \qquad j = 1, 2, \ldots, n(N), \tag{10.9}$$

*индекса, не большего чем* $n+2$, *удовлетворяющий условиям* $P_N(x_i) = c_i$, $i = 0, 1, \ldots, n-1$, *и такой, что* $\mu_N(P_N) > \mu_N(f)$ *для любой* $f$, *удовлетворяющей* (10.8), *отличной от* (10.9). *Индекс* $P_N(x)$ *будет либо* $n$, *либо* $n+1$, *если* $P(1) < c_n$ *и* $n+1$ *или* $n+2$, *если* $P(1) = c_n$.

Доказательство. В силу предположений теоремы точка $(c_0, \ldots, c_n)$ — внутренняя точка моментного пространства, порожденного функциями $x_0^k, \ldots x_n^k$, заданными для $k \in K$. В этом случае многочлен (10.9) определен в силу результатов § 8 и теоремы 4.3 при $k_0 = N$. Масса $\mu_N(P_N)$, соответствующая точке $k = N$, максимальна, так что $\mu_N(P_N) > \mu_N(f)$ для любой $f$, удовлетворяющей (10.8), отличной от $P_N$. Утверждение относительно индекса $P_N$ выводится из результатов § 8.

Замечание 10.2. Индекс $P_N(x)$ будет равняться $n+1$, если $P(1) < c_n$, и $n+2$, если $P(1) = c_n$, так как противное возможно, только если рассматривается главное представление $c_0, c_1, \ldots, c_n$.

Теорема 10.3. *Пусть* $0 < x_0 < x_1 < \ldots < x_n$. *Предположим, что* $c_0, c_1, \ldots, c_n$ *— некоторые действительные числа.*

(i) *Абсолютно монотонная функция на* $(0, \infty)$ *такая, что* $f(x_i) = c_i$, $i = 0, 1, \ldots, n-1$, $f(x_n) \leqslant c_n$ *существует тогда и только тогда, когда*

$$\sum_{i=0}^{n} d_i x_i^k \geqslant 0, \quad k = 0, 1, \ldots \;\rightarrow \sum_{i=0}^{n} d_i c_i \geqslant 0. \tag{10.10}$$

(ii) *Если*

$$\sum_{i=0}^{n} d_i x_i^k \geqslant 0, \quad k = 0, 1, \ldots, \quad \sum_{i=0}^{n} d_i^2 > 0 \rightarrow \sum_{i=0}^{n} d_i c_i \tag{10.11}$$

и две последние суммы положительны, *то существует абсолютно монотонный многочлен* $P$ *степени* $n+1$ *такой, что*

$$P(x_i) = c_i, \quad i = 0, 1, 2, \ldots n.$$

Доказательство. Утверждение теоремы следует из замечания, что соотношения (10.10) и (10.11) являются условиями, выражающими дуальность между пространством моментов и пространством многочленов (см. лемму 9.2 гл. II).

Глава VIII

# МОМЕНТНЫЕ ПРОСТРАНСТВА, ПОРОЖДЕННЫЕ ОГРАНИЧЕННЫМИ МЕРАМИ

## § 1. Введение

В этой главе мы исследуем геометрию моментного пространства

$$\Phi_{n+1} = \left\{ \mathrm{c} = (c_0, c_1, \ldots, c_n) \,|\, c_i = \int_a^b u_i(t)\, \varphi(t)\, d\mu(t), \qquad i = 0, 1, \ldots, n \right\}, \tag{1.1}$$

где $\mu$ фиксированная, конечная, регулярная мера, удовлетворяющая условию $\mu(G) > 0$ для каждого открытого множества $G \subset (a, b)$, интеграл $\int_a^t d\mu$ непрерывен и $\varphi(t)$ пробегает множество $\Phi$ всех борелевских измеримых функций, подчиняющихся ограничению

$$0 \leqslant \varphi(t) \leqslant 1, \quad t \in (a, b).$$

Другими словами, мы будем изучать моментные пространства, порожденные мерами с равномерно ограниченными плотностями. Точнее, пространство $\Phi_{n+1}$ порождено такими мерами $\sigma$, которые абсолютно непрерывны относительно $\mu$ и для которых производная Радона—Никодима $\varphi = d\sigma/d\mu$ ограничена единицей. Элементы множества $\Phi$ рассматриваются как эквивалентные и будут отождествляться, если они различаются на множестве $\mu$ меры нуль. Далее будет считаться само собой разумеющимся, что все утверждения, в которых участвуют функции из $\Phi$, справедливы почти наверное относительно меры $\mu$.

Хотя мы предположили, что $\int_a^t d\mu$ непрерывен, теория обобщается и на тот случай, когда мера $\mu$ имеет веса в некоторых точках. Однако в этом случае изучение становится более трудным и рассмотрение общего случая здесь не оправдало бы затраченных усилий.

В этой главе мы предполагаем, что система $\{u_i\}_0^n$ образует $T$-систему на открытом интервале $(a, b)$, где $a$ и $b$, возможно, бесконечны. Это означает, что функции $u_0, u_1, \ldots, u_n$ непрерывны на $(a, b)$ и что определители (1.1) гл. I положительны, если значения $t$ принадлежат открытому интервалу $(a, b)$. Мы будем предполагать, что функции $u_0, u_1, \ldots, u_n$ $\mu$-интегрируемы на $(a, b)$, так что интегралы типа $\int_a^b u_i(t)\,\varphi(t)\,d\mu(t)$ существуют для всех $\varphi \in \Phi$.

Моментные пространства $\Phi_{n+1}$ естественно изучать в связи с некоторыми статистическими задачами. В частности, лемма Неймана — Пирсона в испытании статистических гипотез является частным случаем экстремальной задачи § 8. Ряд дальнейших применений будет указан в гл. XII. Сейчас мы сосредоточим внимание на формальной теории. Имеется тесная связь между рассуждениями этой главы и теорией распределений типа Полиа, относящейся к различным задачам параметрических статистических решений, рассмотренных у Карлина [1956], [1957] и [1958]. В контексте теории статистических решений существенно рассматривать моментные пространства вида (1.1), где, однако, допустимо, чтобы основная мера имела атомы. За дальнейшими результатами по этому вопросу мы отсылаем к Карлину [1956].

Крейн [1951] всесторонне исследовал моментные пространства (1.1) в связи с экстремальной задачей § 11. Ссылки на другие классические работы об ограниченных моментных пространствах можно найти в книге Шохата и Тамаркина [1943] (гл. III).

Некоторые методы и рассуждения, использованные в этой главе, близки к теореме Ляпунова о множестве значений векторной меры. Формулировка и доказательство этого результата, а также ряда примыкающих к нему результатов будут приведены в §§ 12—14.

## § 2. Граница и общие свойства моментного пространства $\Phi_{n+1}$

Так как множество $\Phi$ слабо компактно, то из этого следует

Теорема 2.1. *Множество $\Phi_{n+1}$ компактно и выпукло в $R^{n+1}$.*

Для моментных пространств $\mathcal{M}_{n+1}$, исследованных в гл. I—VII, важную роль играют меры $\sigma$, имеющие конечное число точек сосредоточения масс. Например, каждая граничная точка соответствует единственной мере $\sigma$ индекса, не превосходящего $n/2$, а каждой внутренней точке соответствуют две главные меры индекса $(n + 1)/2$. В нашем случае роль канонических мер играют функции $\varphi \in \Phi$, которые равны единице на конечном числе интервалов и нулю в остальных. Ниже мы покажем, что каждая точка $c \in \Phi_{n+1}$ может быть порождена функцией такого типа.

О п р е д е л е н и е 2.1. Пусть $\varphi \in \Phi$ — функция, принимающая значение единица на конечном числе интервалов и нуль в остальных. Точки, в которых $\varphi$ изменяет свое значение (т. е. концы интервалов постоянства), называются *узлами* $\varphi$. *Индексом* $\varphi$ обозначается число отдельных невырожденных интервалов, на которых $\varphi(t) = 1$, при условии, что интервал, примыкающий к концу (т. е. интервал, замыкание которого содержит конец основного интервала), считается за 1/2. *Индекс* $I(\mathbf{c})$ *произвольной точки* $\mathbf{c} \in \Phi_{n+1}$ определяется как минимальное значение индекса, соответствующее тем $\varphi$, которые принимают значение единица в конечном числе интервалов и нуль в других точках и которые представляют точку $\mathbf{c}$.

Следующая теорема служит для характеризации границы $\Phi_{n+1}$.

Т е о р е м а 2.2. *Необходимым и достаточным условием для того, чтобы* $\tilde{\mathbf{c}} = (\tilde{c}_0, \tilde{c}_1, \ldots, \tilde{c}_n)$ *принадлежала границе* $\Phi_{n+1}$, *является условие* $I(\tilde{\mathbf{c}}) \leqslant n/2$. *Более того, каждая граничная точка соответствует единственной функции* $\bar{\varphi} \in \Phi$ (*с точностью до эквивалентности*).

Д о к а з а т е л ь с т в о. Предположим, что точка $\tilde{\mathbf{c}} = (\tilde{c}_0, \tilde{c}_1, \ldots, \tilde{c}_n)$ лежит на границе $\Phi_{n+1}$. Следовательно, существует опорная плоскость к $\Phi_{n+1}$ в точке $\tilde{\mathbf{c}}$, откуда следует, что существуют действительные постоянные $a_0, a_1, \ldots, a_n, d \left(\sum_{i=0}^{n} a_i^2 > 0\right)$, удовлетворяющие условию

$$\sum_{i=0}^{n} a_i c_i + d \leqslant 0 \text{ для всех } \mathbf{c} \in \Phi_{n+1}, \qquad \sum_{i=0}^{n} a_i \tilde{c}_i + d = 0. \tag{2.1}$$

Вычитая, получим

$$\int_a^b \sum_{i=0}^{n} a_i u \quad t) - \tilde{\varphi}(t)]\, d\mu(t) \leqslant 0 \tag{2.2}$$

для всех $\varphi \in \Phi$ и $\tilde{\varphi} \in V(\tilde{\mathbf{c}})$, где

$$V(\tilde{\mathbf{c}}) = \left\{ \varphi \,|\, \tilde{c}_i = \int_a^b u_i \varphi\, d\mu, \quad i = 0, 1, \ldots, n, \ \varphi \in \Phi \right\}. \tag{2.3}$$

Положив $u(t) = \sum_{i=0}^{n} a_i u_i(t)$, мы утверждаем, что (2.2) имеет место тогда и только тогда, когда

$$\tilde{\varphi}(t) = \begin{cases} 1, & \text{если } u(t) > 0, \\ 0, & \text{если } u(t) \leqslant 0. \end{cases} \tag{2.4}$$

Чтобы доказать это утверждение, заметим, что (2.2) является простым следствием (2.4).

Обратное утверждение устанавливается следующим образом. Если $\tilde{\varphi}(t) < 1$ на подмножестве $T \subset \{t \mid u(t) > 0\}$ таком, что $\mu(T) > 0$, то положим

$$\varphi_1(t) = \begin{cases} 1, & t \in T, \\ \tilde{\varphi}(t), & t \notin T. \end{cases}$$

Ясно, что $\varphi = \varphi_1$ не удовлетворяет (2.2). Отсюда следует, что $\tilde{\varphi}(t) = 1$, если $u(t) > 0$. Для доказательства того, что $\tilde{\varphi}(t) = 0$, если $u(t) < 0$, используются аналогичные рассуждения. Таким образом, из (2.2) следует (2.4).

Так как многочлен $u(t)$ нетривиальный, то он обращается в нуль на открытом интервале $(a, b)$ менее чем в $n+1$ точке. Это множество не имеет положительной меры, так как $\mu$ — непрерывная мера. Поэтому равенство (2.4) однозначно определяет функцию $\tilde{\varphi}$. Более того, существует не более, чем $n+1$ интервалов, таких, что $\tilde{\varphi}$ изменяет свои значения на соседних интервалах. Из этого следует в силу определения 2.1, что $I(\tilde{\mathbf{c}}) \leqslant n/2$.

Теперь предположим, что $I(\tilde{\mathbf{c}}) \leqslant n/2$, а $\tilde{\varphi} \in V(\tilde{\mathbf{c}})$ имеет индекс $I(\tilde{\mathbf{c}})$. Мы можем построить многочлен $u(t) = \sum_{i=0}^{n} a_i u_i(t)$, узловые нули которого совпадают с узлами $\tilde{\varphi}$, причем $u(t) > 0$, если $\tilde{\varphi}(t) = 1$, т. е. выполнено (2.4). Тогда легко проверить, что справедливо (2.2). Положив $d = -\int_a^b u(t)\,\tilde{\varphi}(t)\,d\mu(t)$, видим, что (2.1) является непосредственным следствием (2.2). Кроме того, (2.1) можно рассматривать как отношение, характеризующее опорную гиперплоскость к $\Phi_{n+1}$ в точке $\tilde{\mathbf{c}}$, так что $\tilde{\mathbf{c}} \in \mathrm{Bd}\,\Phi_{n+1}$. Доказательство теоремы закончено.

Из теоремы 2.2 следует, что любая функция $\varphi \in \Phi$, удовлетворяющая условию $0 < \varphi(t) < 1$ на $(a, b)$, необходимо порождает внутреннюю точку $\Phi_{n+1}$.

Следствие 2.2 (а) *Пространство $\Phi_{n+1}$ имеет внутренние точки.*

Это следствие вытекает также из линейной независимости функций $u_0, u_1, \ldots, u_n$.

Следующий результат является следствием доказательства теоремы 2.2 даже в большей степени, чем утверждение самой теоремы.

Следствие 2.2 (b) *Существует соответствие между многочленами* $u(t)=\sum_{i=0}^{n} a_i u_i(t)$ *и границей* $\Phi_{n+1}$ *в том смысле, что гиперплоскость* $\sum_{i=0}^{n} a_i x_i + d$, *где* $d=-\int_a^b u_+(t)\,d\mu(t)$, *а* $u_+(t)=\max\{0, u(t)\}$, *является опорной к пространству* $\Phi_{n+1}$ *в точке* $\mathbf{c}$, *представляемой функцией*

$$\varphi(t)=\begin{cases}1, & u(t)>0,\\ 0, & u(t)\leqslant 0.\end{cases} \tag{2.5}$$

Используя метод теоремы 2.2, можно вывести простой способ определения принадлежности точки $\mathbf{c}=(c_0, c_1, \ldots, c_n)$ к $\Phi_{n+1}$. Этот критерий является аналогом леммы 9.2 гл. II, которая выражает дуальность между моментным пространством $\mathcal{M}_{n+1}$ и пространством $\mathcal{P}_{n+1}$ неотрицательных многочленов.

Теорема 2.3. (Крейн [1951]). *Для того чтобы вектор* $\mathbf{c}^0$ *принадлежал* $\Phi_{n+1}$, *необходимо и достаточно, чтобы для каждого многочлена*

$$u(t)=\sum_{i=0}^{n} a_i u_i(t) \tag{2.6}$$

*выполнялось неравенство*

$$\sum_{i=0}^{n} a_i c_i^0 \leqslant \int_a^b u_+(t)\,d\mu(t), \tag{2.7}$$

*где* $u_+(t)=\max\{0, u(t)\}$.

Доказательство. Необходимость очевидна. Чтобы доказать достаточность, нужно проверить, что вектор $\mathbf{c}^0$ и нулевой вектор $\mathbf{0}$, который, очевидно, принадлежит $\Phi_{n+1}$, лежат по одну сторону от каждой опорной гиперплоскости к $\Phi_{n+1}$.

Пусть $\sum_{i=0}^{n} a_i x_i + d$ — любая опорная гиперплоскость к $\Phi_{n+1}$, т. е. предположим, что $\sum_{i=0}^{n} a_i c_i + d \leqslant 0$ для $\mathbf{c}\in\Phi_{n+1}$, причем равенство имеет место для некоторого $\mathbf{c}\in\mathcal{D}\Phi_{n+1}$. Если положить $u(t)= \sum_{i=0}^{n} a_i u_i(t)$, то граничной точке $\mathbf{c}$ соответствует функция $\varphi$, определенная в (2.5), и $d=-\int_a^b u_+(t)\,d\mu(t)$. Из (2.7) мы выводим, что $\sum_{i=0}^{n} a_i c_i^0 + d \leqslant 0$. Поэтому $\mathbf{c}^0$ и нулевой вектор $\mathbf{0}$ лежат по одну сторону от любой опорной плоскости к $\Phi_{n+1}$.

## § 3. Главные представления внутренних точек

Приступим к детальному изучению внутренности $\Phi_{n+1}$.

Теорема 3.1. *Каждая внутренняя точка* $\mathbf{c}^0 = \{c_0^0, \ldots, c_n^0\} \in \Phi_{n+1}$ *имеет ровно два представления индекса* $(n+1)/2$; *одно из них равно единице на сегменте, примыкающем к концу интервала* $b$, *другое равно нулю на сегменте, примыкающем к концу интервала* $b$. (Мы обозначим первый из них $\bar{\varphi}$, а последний $\underline{\varphi}$ и будем ссылаться на них как на верхнее и нижнее главные представления соответственно.)

Доказательство. Сначала мы предположим, что $n = 2m$ и заметим, что $\mathbf{c}^0$ лежит на границе $\Phi_{n+1}(a, d_0)$ для некоторого $d_0 \in (a, b)$, где

$$\Phi_{n+1}(a, d) = \left\{ \mathbf{c} = (c_0, \ldots, c_n) \mid c_i = \int\limits_a^d u_i \varphi \, du, \quad i = 0, 1, \ldots, n, \varphi \in \Phi \right\}.$$

Если бы это было неверно, то мы смогли бы разложить связный интервал $(a, b)$ на два непересекающихся непустых открытых множества: $A = \{d \mid \mathbf{c}^0 \in \operatorname{Int} \Phi_{n+1}(a, d)\}$ и $B = \{d \mid \mathbf{c}^0 \notin \Phi_{n+1}(a, d)\}$. Обращаясь к теореме 2.2, мы получаем, что $\mathbf{c}$ имеет представление индекса, не большего чем $n/2$, относительно $(a, d_0)$. Если распространить это представление на интервал $(a, b)$, положив его равным нулю на интервале $(d_0, b)$, то получившееся в результате представление $\underline{\varphi}$ должно иметь индекс $(n+1)/2$, так как $\mathbf{c}^0 \in \operatorname{Int} \Phi_{n+1}$. К тому же мы видим, что так как $n = 2m$, то $\underline{\varphi}$ должно равняться единице на интервале, примыкающем к точке $a$.

Обращаясь к другому концу интервала $(a, b)$, мы можем построить представление $\bar{\varphi}$ индекса $(n+1)/2$, которое равно нулю на интервале, примыкающем к $a$, и равно единице на интервале, примыкающем к $b$. Это достигается при помощи рассуждений, использующих пространства

$$\Phi_{n+1}(d, b) = \left\{ \mathbf{c} = (c_0, \ldots, c_n) \mid c_i = \int\limits_d^b u_i \varphi \, d\mu, \quad i = 0, 1, \ldots, n, \varphi \in \Phi \right\}.$$

Ясно что представления $\underline{\varphi}$ и $\bar{\varphi}$ различны.

Если $n$ — нечетно, то приведенные выше конструкции приводят к функциям, которые обе равны нулю на интервале, примыкающем к $b$. Фактически эти два представления совпадают. Поэтому в данном случае мы располагаем лишь представлением $\underline{\varphi}$ индекса $(n+1)/2$, которое равно нулю на интервале, примыкающем к $b$. Для того чтобы получить второе главное представление, заметим, что $\mathbf{c}$ является граничной точкой пространства $\psi_{n+1}(a, d_0)$ для некоторого $d_0 \in (a, b)$, где $\psi_{n+1}(a, d_0)$ определяется как ограничение функций $\varphi$, равных единице для $t \in (d_0, b)$. Отсюда следует, что точка $\mathbf{c}$ координатами $c_i^0 - \int\limits_{d_0}^{b} u_i(t)\,d\mu(t)$, $i = 0, 1, \ldots, n$ является граничной точкой пространства $\Phi_{n+1}(a, \underline{d}_0)$. Рассуждая, как и выше, мы приходим к представлению $\bar{\varphi}$, которое равно единице на интервале, примыкающем к $b$ и имеет индекс $(n+1)/2$.

Для того чтобы доказать, что имеется ровно два главных представления, мы поступим следующим образом. Если дано произвольное представление $\varphi$ точки $\mathbf{c}^0$, которое имеет индекс $(n+1)/2$, то мы сравним его с представлением, построенным выше и имеющим тот же тип; например, если $\varphi$ равно единице на сегменте, примыкающем к концу $b$, то мы сравниваем его с $\bar{\varphi}$.

Мы утверждаем, что разность $\varphi - \bar{\varphi}$, где $\bar{\varphi}$ — верхнее главное представление (или $\varphi - \underline{\varphi}$, когда $\underline{\varphi}$ — нижнее главное представление), может иметь не более $n$ изменений знака, т. е. $S(\varphi - \bar{\varphi}) \leqslant n$ (см. определение 3.1 т. III).

Пусть для определенности $n = 2m$, а $\varphi$ и $\bar{\varphi}$ оба являются верхними главными представлениями индекса $(n+1)/2$. Пусть узлы $\bar{\varphi}$ обозначаются

$$a < \bar{\eta}_1 < \bar{\xi}_1 < \bar{\eta}_2 < \bar{\xi}_2 < \ldots < \bar{\eta}_m < \bar{\xi}_m < \bar{\eta}_{m+1} < \bar{\xi}_{n+1} = b,$$

где

$$\bar{\varphi}(t) = \begin{cases} 1, & t \in (\bar{\eta}_i, \bar{\xi}_i),\ i = 1, 2, \ldots, m+1, \\ 0 & \text{в противном случае,} \end{cases}$$

а узлы $\varphi$ обозначаются

$$a < \eta_1 < \xi_1 < \eta_2 < \xi_2 < \ldots < \eta_m < \xi_m < \eta_{m+1} < \xi_{m+1} = b,$$

где

$$\varphi(t) = \begin{cases} 1, & t \in (\eta_i, \xi_i),\ i = 1, 2, \ldots, m+1, \\ 0 & \text{в противном случае.} \end{cases}$$

Не умаляя общности, мы предположим, что $\eta_{m+1} \leqslant \bar{\eta}_{m+1}$. Тогда $\varphi(t) - \bar{\varphi}(t) \geqslant 0$ для $\bar{\xi}_m \leqslant t < b$ и, следовательно, изменение знака $\varphi - \bar{\varphi}$ может происходить только на интервале $a < t \leqslant \bar{\xi}_m$. Более того, функция $\varphi(t) - \bar{\varphi}(t)$ теперь может изменять знак только в точках $\bar{\eta}_1, \bar{\xi}_1, \bar{\eta}_2, \bar{\xi}_2, \ldots, \bar{\eta}_m, \bar{\xi}_m$. Поэтому $S(\varphi - \bar{\varphi}) \leqslant 2m = n$.

Если $S(\varphi - \bar{\varphi}) = p \leqslant n$, то имеется $p$ точек, $a < t_1 < \ldots < t_p < b$, таких, что $\varphi - \bar{\varphi}$ принимает противоположные знаки на последовательных интервалах $(a_1, t_1)$, $(t_1, t_2), \ldots, (t_p, b)$, и $\varphi - \bar{\varphi}$ не равно нулю на некотором подынтервале каждого из этих интервалов. Теперь мы определим нетривиальный многочлен $u^*$, который имеет узловые нули в точках $t_1, \ldots, t_p$ и имеет постоянный знак на каждом интервале. Однако из соотношения $\int\limits_a^b u_i \varphi \, d\mu = \int\limits_a^b u_i \bar{\varphi} \, d\mu$ $(i = 0, 1, \ldots, n)$ следует, что $0 = \int\limits_a^b u^* (\varphi - \bar{\varphi}) \, d\mu$; но это противоречит тому, что $(\varphi - \bar{\varphi}) u^*$ сохраняет постоянный знак и не равно нулю на множестве положительной меры $\mu$ (так как $\mu(G) > 0$ для каждого открытого множества $G$). Противоречие можно устранить, только если $\varphi$ и $\bar{\varphi}$ совпадают, т. е. эквивалентны почти наверное по мере $\mu$.

## § 4. Другой подход к построению главных представлений

Главные представления связаны с границей множества

$$R(\mathbf{c}^0) = \left\{ \omega \mid \omega = \int\limits_a^b \Omega(t) \varphi(t) \, d\mu(t), \varphi \in V(\mathbf{c}^0) \right\}, \tag{4.1}$$

где предполагается, что $\{u_0, u_1, \ldots, u_n, u_{n+1} = \Omega\}$ образуют $T$-систему на $(a, b)$. Очевидно, что множество $R(\mathbf{c}^0)$ является невырожденным интервалом $\underline{\omega} \leqslant \omega \leqslant \bar{\omega}$. Точка

$$\bar{\mathbf{c}}^0 = (c_0^0, c_1^0, \ldots, c_n^0, \bar{\omega}) \tag{4.2}$$

допускает единственное представление индекса $(n+1)/2$; то же относится к точке

$$\underline{\mathbf{c}}^0 = (c_0^0, c_1^0, \ldots, c_n^0, \underline{\omega}). \tag{4.3}$$

Теперь докажем, что $\bar{\mathbf{c}}^0$ ($\underline{\mathbf{c}}^0$) соответствует верхнему (нижнему) главному представлению $\mathbf{c}^0$.

Пусть $\varphi^*$ и $\varphi_*$ определяют представления, соответствующие $\bar{\mathbf{c}}^0$ и $\underline{\mathbf{c}}^0$ соответственно. Прямая проверка показывает, что разность $\varphi^* - \varphi_*$ изменяет знак не более чем $n + 1$ раз. Более того, рассуждения, использованные при доказательстве теоремы 3.1, показывают также, что $\varphi^* - \varphi_*$ имеет не менее $n + 1$ изменений знака. Обозначим точки, в которых происходит изменение знака разности $\psi(t) = \varphi^*(t) - \varphi_*(t)$ через $\xi_1 < \xi_2 < \ldots < \xi_{n+1}$, и положим $\xi_0 = a$, $\xi_{n+2} = b$, т. е. $\psi$ имеет противоположные знаки на последовательных интервалах $L_i = (\xi_i, \xi_{i+1})$, $i = 0, 1, \ldots, n+1$, и отлична от нуля на некотором подынтервале каждого из $L_i$. Тогда

$$\begin{gathered} 0 = \int\limits_a^b u_i(t) \psi(t) \, d\mu(t) = \varepsilon \sum_{j=0}^{n+1} (-1)^j K(i, j), \qquad i = 0, \ldots, n, \\ \bar{\omega} - \underline{\omega} = \int\limits_a^b \Omega(t) \psi(t) \, d\mu(t) = \varepsilon \sum_{j=0}^{n+1} (-1)^j K(n+1, j), \end{gathered} \tag{4.4}$$

где

$$K(i, j) = \int\limits_{\xi_j}^{\xi_{j+1}} u_i(t)\,|\psi(t)|\,d\mu(t), \qquad (i, j = 0, 1, \ldots, n+1),$$

а $\varepsilon$ принимает значение $+1$ или $-1$. Заметим для дальнейших ссылок, что $\varphi^* - \varphi_*$ имеет знак $(-1)^{n+1}\varepsilon$ на последнем интервале $[\xi_{n+1}, b]$.

Так как $\{u_i\}_0^n$ и $\{u_i\}_0^{n+1}$ — $T$-системы, то

$$\begin{aligned} \det \| K(i, j) \|_{i,j=0}^n &= K\begin{pmatrix} 0, 1, \ldots, n \\ 0, 1, \ldots, n \end{pmatrix} > 0, \\ \det \| K(i, j) \|_{i,j=0}^{n+1} &= K\begin{pmatrix} 0, 1, \ldots, n+1 \\ 0, 1, \ldots, n+1 \end{pmatrix} > 0. \end{aligned} \tag{4.5}$$

Далее

$$K\begin{pmatrix} 0, 1, \ldots, n \\ 0, 1, \ldots, n \end{pmatrix} =$$

$$= \int\limits_{\xi_n}^{\xi_{n+1}} \ldots \int\limits_{\xi_1}^{\xi_2} \int\limits_{a}^{\xi_1} U\begin{pmatrix} 0, 1, \ldots, n \\ t_0, t_1, \ldots, t_n \end{pmatrix} \prod_{i=0}^{n} |\psi(t_i)|\, dt_0 \ldots dt_n,$$

откуда следует первое из неравенств (4.5). Второе неравенство получается аналогично. Если рассматривать последовательность $\varepsilon, -\varepsilon, \varepsilon, \ldots, (-1)^{n+1}\varepsilon$ как решение системы уравнений (4.4), то мы можем разрешить эту систему относительно $(-1)^{n+1}\varepsilon$ и, принимая во внимание (4.5), получим, что $(-1)^{n+1}\varepsilon > 0$. Из этого следует, что $\varphi^*(t) \geqslant \varphi_*(t)$ для $t \in [\xi_{n+1}, b]$, и строгое неравенство имеет место в некоторой точке этого интервала. Учитывая, что $\varphi^*$ и $\varphi_*$ обе имеют индекс $(n+1)/2$, а $\varphi^* - \varphi_*$ имеет $n+1$ изменение знака, мы заключаем, что $\varphi^*(t) - \varphi_*(t) = 1$ на интервале, примыкающем к точке $b$. Поэтому мера $\varphi^*$, соответствующая $\bar{c}^0$, является верхним главным представлением, а $\varphi_*$, соответствующая $\underline{c}_0$, является нижним главным представлением.

## § 5. Пополнение чебышевской системы, определенной на открытом интервале

Цель этого параграфа показать, что предположение о существовании дополнительной функции $\Omega$, для которой $u_0, u_1, \ldots, u_n, \Omega$ является $T$-системой на $(a, b)$, несущественно в рассуждениях § 4. Точнее, мы докажем, что если $\{u_i\}_0^k$—$ET$-система на $[a, b]$ для каждого $k = 0, 1, \ldots, n$, то существует функция $\Omega$ такая, что $\{u_0, u_1, \ldots$ $\ldots, u_n, \Omega\}$ является $T$-системой на $(a, b)$. Из этого факта и результатов § 4 непосредственно следует существование главных представлений. Дальнейшее обсуждение построений этого параграфа переносится в гл. XI. Не умаляя общности, мы будем в дальнейшем предполагать, что интервал $(a, b)$ конечен (в противном случае делаем замену переменных; см. стр. 155 гл. V или стр. 203 гл. VI).

Переходя к непосредственной реализации намеченного плана, мы сначала преобразуем каждую из функций $u_0, u_1, \ldots, u_n$, вводя дополнительную систему функций

$$v_i(t) = v_i(t;\delta) = \int_a^b u_i(y) \frac{\exp\left\{-\frac{1}{2}\left(\frac{t-y}{\delta}\right)^2\right\}}{\delta\sqrt{2\pi}}\, dy, \quad i = 0, 1, \ldots, n, \ \delta > 0, \tag{5.1}$$

определенную для $t \in [a, b]$. В силу основной композиционной формулы (3.13) гл. I система $\{v_i\}_0^n$ обладает свойством

$$V^*\begin{pmatrix} 0, 1, \ldots, n \\ t_0, t_1, \ldots, t_n \end{pmatrix} > 0, \quad a \leqslant t_0 \leqslant t_1 \leqslant \ldots \leqslant t_n \leqslant b, \tag{5.2}$$

т. е. система $\{v_i\}_0^n$ является $ET$-системой на замкнутом интервале $[a, b]$. (О смысле обозначения (*) см. стр. 17).

Лемма 5.1. *Существует невырожденная $(n+1) \times (n+1)$-матрица $\mathbf{B} = \mathbf{B}(\delta) = \|b_{ij}\|_{i,j=0}^n$ такая, что система $\{\psi_i\}_0^n$, определенная равенством $\psi_i(t) = \sum_{j=0}^n b_{ij} v_j(t)$, $i = 0, 1, \ldots, n$, обладает тем свойством, что определители Вронского $W(\psi_0)$, $W(\psi_0, \psi_1), \ldots$ $\ldots, W(\psi_0, \ldots, \psi_n)$ строго положительны на $[a, b)$ (заметим, что мы включаем конец интервала $a$, но не $b$).*

Доказательство. Достаточно найти $\mathbf{B}$ такое, чтобы определители Вронского не обращались в нуль на $[a, b)$, так как мы можем затем умножить каждую $\psi_i$ соответственно на $+1$ или $-1$, чтобы сделать эти определители положительными. Пусть $\mathbf{B}$ является решением матричного уравнения

$$\mathbf{J} = \mathbf{B}\mathbf{V}(b),$$

где $J_{ik} = \delta_{k,n-i}$ для $i, k = 0, 1, \ldots, n$ ($\delta_{ij}$ означает функцию Кронекера), а $\mathbf{V}(b)$ — матрица, соответствующая определителю Вронского

$$W(v_0, \ldots, v_n)(b) = V^*\begin{pmatrix} 0, 1, \ldots, n \\ b, b, \ldots, b \end{pmatrix}.$$

При таком определении $\mathbf{B}$ любая линейная комбинация $\psi_0, \ldots, \psi_k$ имеет нуль в точке $b$ кратности, не меньшей чем $n-k$. Теперь, если $W(\psi_0, \ldots, \psi_k)(y) = 0$ для некоторого $y \in [a, b)$, то существует невырожденная линейная комбинация $\psi_0, \ldots, \psi_k$ с нулем кратности $k$ в точке $y$. В этом случае мы можем образовать невырожденную линейную комбинацию $v_0, \ldots, v_n$, которая имеет по крайней мере $n+1$ нуль с учетом кратности на $[a, b]$. Так как в силу (5.2) и теоремы 4.3 гл. I это невозможно, то отсюда следует, что определитель Вронского $W(\psi_0, \ldots, \psi_k)$ не может обращаться в нуль на $[a, b)$. Доказательство закончено.

Теперь мы введем дифференциальные операторы

$$D_j f = \frac{d}{dt} \frac{f}{w_j}$$

и определим функции $w_0, w_1, \ldots, w_n$ на $[a, b)$ рекурсивно соотношениями

$$\begin{aligned} w_0 &= \psi_0, \\ w_1 &= D_0 \psi_1, \\ w_2 &= D_1 D_0 \psi_2, \\ &\cdot \\ &\cdot \\ &\cdot \\ w_n &= D_{n-1} D_{n-2} \ldots D_0 \psi_n. \end{aligned}$$

Функции $w_i$, $i = 0, 1, \ldots, n$, определены корректно. Они строго положительны на $[a, b)$, как это видно из представлений

$$w_1 = \frac{W(\psi_0, \psi_1)}{\psi_0^2},$$

$$w_k = \frac{W(\psi_0, \ldots, \psi_k)\, W(\psi_0, \ldots, \psi_{k-2})}{W^2(\psi_0, \ldots, \psi_{k-1})}, \qquad k = 2, \ldots, n, \tag{5.3}$$

которые проверяются по индукции с использованием тождества Сильвестра (см. стр. 43).

Теперь возьмем произвольную непрерывную строго положительную функцию $w_{n+1}$ на $[a, b)$ и определим

$$\psi_{n+1}(t) = w_0(t) \int_a^t w_1(\xi_1) \int_a^{\xi_1} w_2(\xi_2) \ldots \int_a^{\xi_n} w_{n+1}(\xi_{n+1})\, d\xi_{n+1} \ldots d\xi_1.$$

Ясно, что

$$w_{n+1} = D_n \ldots D_1 D_0 \psi_{n+1}$$

и с помощью (5.3) находим, что

$$w_{n+1} = \frac{W(\psi_0, \ldots, \psi_{n+1})\, W(\psi_0, \ldots, \psi_{n-1})}{W^2(\psi_0, \ldots, \psi_n)}. \tag{5.4}$$

Далее, используя соотношения (5.3) и (5.4), получаем, что

$$W(\psi_0, \ldots, \psi_k) = w_0^{k+1} w_1^k \ldots w_k, \qquad k = 0, 1, \ldots, n+1,$$

так что определитель Вронского $w(\psi_0, \ldots, \psi_{n+1})$ строго положителен на $[a, b)$.

Приведенные выше рассуждения доказывают следующую лемму, которая имеет и самостоятельный интерес.

Лемма 5.2. *Если* $\psi_0, \ldots, \psi_n$ *содержатся в* $C^n[a, b)$ *и* $W(\psi_0) > 0, \ldots, W(\psi_0, \ldots, \psi_n) > 0$ *на* $[a, b)$, *то существует функция* $\psi_{n+1} \in C^{n+1}[a, b)$ *такая, что* $W(\psi_0, \ldots, \psi_{n+1}) > 0$ *на* $[a, b)$.

Для того чтобы приступить к обсуждению главных представлений внутренних точек $\Phi_{n+1}$, нужно сначала показать, что $\{\psi_i\}_0^n$ и $\{\psi_i\}_0^{n+1}$ являются $T$-системами на $[a, b)$. То, что $\{\psi_i\}_0^n$ — $T$-система, следует непосредственно из (5.2) и леммы 5.1. Из условий на определители Вронского $W(\psi_0, \ldots, \psi_k) > 0$, $k = 0, 1, \ldots, n+1$, следует, что $\{\psi_i\}_0^k$ для каждого $k = 0, 1, \ldots, n+1$ является $ET$-системой на $[a, b)$ (см. теорему 1.2 гл. XI). Однако для наших целей будет достаточно доказать следующую лемму.

Лемма 5.3. *Системы* $\{\psi_i\}_0^k$, $k = 0, 1, \ldots, n+1$, *являются* $T$*-системами на* $(a, b)$.

Доказательство проводим индукцией по $k$. Для $k = 0$ результат очевиден. Если положить $\Phi(i, t) = \psi_i(t)$, то, вынося $\psi_0$ из каждого столбца $\Phi\begin{pmatrix} 0, 1, \ldots, k \\ t_0, t_1, \ldots, t_k \end{pmatrix}$ и вычитая каждый столбец из предшествующего ему, мы получаем, используя теорему о среднем, что

$$\Phi\begin{pmatrix} 0, 1, \ldots, k \\ t_0, t_1, \ldots, t_k \end{pmatrix} = \prod_{j=0}^{k} \psi_0(t_j) \prod_{j=0}^{k-1} (t_{j+1} - t_j)\, \widehat{\Phi}\begin{pmatrix} 0, 1, \ldots, k-1 \\ \eta_0, \eta_1, \ldots, \eta_{k-1} \end{pmatrix},$$

где

$$\widehat{\psi}_i(t) = \widehat{\Phi}(i, t) = \frac{d}{dt}\left(\frac{\psi_{i+1}(t)}{\psi_0(t)}\right), \qquad i = 0, 1, \ldots, k-1, \tag{5.5}$$

и $a < t_0 < \eta_0 < t_1 < \ldots < t_{k-1} < \eta_{k-1} < t_k < b$.

Система (5.5) называется *приведенной системой* и оказывается полезной во многих отношениях (см. например, (3.1) гл. III). Теперь равенство

$$W(\psi_0, \ldots, \psi_l) = (\psi_0)^{l+1} W(1, \psi_1/\psi_0, \ldots, \psi_l/\psi_0) = (\psi_0)^{l+1} W(\widehat{\psi}_0, \ldots, \widehat{\psi}_{l-1}), \qquad l = 1, 2, \ldots, n+1 \tag{5.6}$$

(см. ниже § 6) показывает, что $W(\widehat{\psi}_0, \ldots, \widehat{\psi}_{k-1}) > 0$, $k = 1, 2, \ldots, n+1$. Это позволяет нам применить предположение индукции к $\hat{\psi}_0, \ldots, \hat{\psi}_{k-1}$, имеем $\hat{\Phi}\begin{pmatrix} 0, 1, \ldots, k-1 \\ \eta_0, \eta_1, \ldots, \eta_{k-1} \end{pmatrix} > 0$ и, следовательно, $\Phi\begin{pmatrix} 0, 1, \ldots, k \\ t_0, t_1, \ldots, t_k \end{pmatrix} > 0$. Теорема доказана.

Справедливость равенства (5.6), которое приводит порядок определителя Вронского к единице, является следствием правила Лейбница

$$\left(\frac{\psi_j}{\psi_0}\right)^{(m)} = \frac{\psi_j^{(m)}}{\psi_0} + \sum_{p=1}^{m} C_m^p \, \psi_j^{(m-p)} \left(\frac{1}{\psi_0}\right)^{(p)}.$$

Мы делим каждый столбец $W(\psi_0, \ldots, \psi_k)$ на $\psi_0$ и затем прибавляем к каждому столбцу подходящую линейную комбинацию всех предшествующих столбцов. Действуя таким образом последовательно справа налево, получаем (5.6).

Изложенная техника пополнения $T$-системы, определенной на открытом интервале $(a, b)$ таким образом, чтобы пополненная система содержала некоторую дополнительную функцию, относится к понятию «свойство $W$», связанному с определенными линейными дифференциальными уравнениями, исследованными Полиа [1922].

Возвращаясь к исходной задаче, мы рассмотрим точку $\mathbf{c} = (c_0, \ldots, c_n) \in \operatorname{Int} \Phi_{n+1}$, которая представляется относительно системы $\{u_i\}_0^n$ некоторой функцией $\varphi \in \Phi$. Пусть

$$c_i(\delta) = \int_a^b \psi_i(t; \delta)\, \varphi(t)\, d\mu(t), \qquad i = 0, 1, \ldots, n,$$

где $\psi_i(t; \delta)$ получаются из $v_i(t; \delta)$ при помощи невырожденного преобразования $\mathbf{B}(\delta)$ в соответствии с леммой 5.1.

Точка с координатами $\int v_i \varphi\, d\mu$, $i = 0, 1, \ldots, n$, является внутренней точкой моментного пространства, порожденного системой $\{v_i\}_0^n$, так что для последовательности $\delta_k \to 0$ точки $\mathbf{c}(\delta_k) = (c_0(\delta_k), \ldots, c_n(\delta_k))$ являются внутренними точками соответствующих моментных пространств, порожденных системами $\{\psi_i(t; \delta_k)\}_{i=0}^n$. С помощью дополнительной функции $\psi_{n+1}(t; \delta_k)$ мы, используя построение в § 4, определяем верхнее и нижнее главные представления $\overline{\varphi}(t; \delta_k)$ и $\underline{\varphi}(t; \delta_k)$ точки $\mathbf{c}(\delta_k)$ и выбираем подпоследовательности так, чтобы узлы сходились. Отсюда следует, что $\overline{\varphi}(t; \delta_k)$ и $\underline{\varphi}(t; \delta_k)$ сходятся к функциям $\overline{\varphi}$ и $\underline{\varphi}$ соответственно.

Представления $\underline{\varphi}$ и $\overline{\varphi}$ должны иметь индекс не больший, чем $(n+1)/2$ и, кроме того, $\overline{\varphi}$ принимает значение единица на интервале, примыкающем к точке $b$ или имеет индекс не больший, чем

$n/2$, а $\underline{\varphi}$ равна нулю на интервале, примыкающем к $b$, или имеет индекс не больший, чем $n/2$.

Для каждого $k$

$$\int_a^b \sum_{j=0}^n b_{ij}(\delta_k)\, v_j(t;\,\delta_k)\, \varphi(t)\, d\mu(t) =$$
$$= \int_a^b \sum_{j=0}^n b_{ij}(\delta_k)\, v_j(t;\,\delta_k)\, \overline{\varphi}(t;\,\delta_k)\, d\mu(t), \qquad i = 0, 1, \ldots, n,$$

и, так как матрица $\mathbf{B}(\delta_k)$ не вырождена, то

$$\int_a^b v_i(t;\,\delta_k)\, \varphi(t)\, d\mu(t) = \int_a^b v_i(t;\,\delta_k)\, \overline{\varphi}(t;\,\delta_k)\, d\mu(t), \qquad i = 0, 1, \ldots, n.$$

Переходя к пределу по $k$, получаем $c_i = \int_a^b u_i \overline{\varphi}\, d\mu$ для $i = 0, 1, \ldots, n$. Такой же результат справедлив для $\underline{\varphi}$, и таким образом, $\overline{\varphi}$ и $\underline{\varphi}$ являются требуемыми верхним и нижним главными представлениями точки $\mathbf{c}$.

## § 6. Свойства перемежаемости узлов канонических представлений

Некоторые из вопросов этого параграфа рассмотрены в гл. II работы Крейна [1951].

Здесь мы предполагаем, что $\{u_i\}_0^n$ и $\{u_i\}_0^{n+1}$ являются $T$-системами $\mu$-интегрируемых функций на $(a, b)$. Как указано в § 5, предположение о существовании дополнительной функции $u_{n+1}$ не требует никаких дополнительных ограничений.

Для каждого $\omega$ в интервале $(\underline{\omega}, \overline{\omega})$ (см. § 4) рассмотрим точку

$$\mathbf{c}(\omega) = (c_0^0, c_1^0, \ldots, c_n^0, \omega) \in \operatorname{Int} \Phi_{n+2} \tag{6.1}$$

($\Phi_{n+2}$ означает моментное пространство, построенное по $\{u_i\}_0^{n+1}$). Каждое $\mathbf{c}(\omega)$ имеет два главных представления (верхнее и нижнее) индекса $(n+2)/2$, которые мы будем называть каноническими представлениями $\mathbf{c}^0$. Следующая теорема описывает соотношение между различными каноническими представлениями точки $\mathbf{c}^0$.

Теорема 6.1. *Пусть $\varphi$ и $\varphi^*$ — два различных верхних канонических представления точки* $\mathbf{c}^0$ (соответствующие $\mathbf{c}(\omega)$ *и* $\mathbf{c}(\omega^*)$, где $\omega^* > \omega$). *Тогда $\varphi - \varphi^*$ меняет знак ровно $n + 1$ раз.*

Доказательство. Мы проводим доказательство только в случае, когда $n = 2m$; если $n = 2m + 1$, то доказательство проводится аналогично.

Пусть узлы верхних канонических представлений точек $\mathbf{c}(\omega)$ и $\mathbf{c}(\omega^*)$ расположены в возрастающем порядке. Таким образом,

$$\begin{aligned} &(a=\eta_0<\xi_0<\eta_1<\xi_1<\ldots<\eta_{m+1}<\xi_{m+1}=b)\sim \mathbf{c}(\omega),\\ &(a=\eta_0^*<\xi_0^*<\eta_1^*<\xi_1^*<\ldots<\eta_{m+1}^*<\xi_{m+1}^*=b)\sim \mathbf{c}(\omega^*), \end{aligned} \tag{6.2}$$

где $\varphi(t)=1$ на $(\eta_{m+1}, \xi_{m+1})$ и $\varphi^*(t)=1$ на $(\eta_{m+1}^*, \xi_{m+1}^*)$. Для того чтобы запомнить эти обозначения, мы отметим, что верхнее каноническое представление равно единице на интервалах вида $(\eta_i, \xi_i)$, т. е.

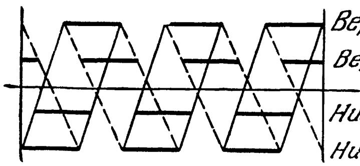



Рис. 12.

интервалах, которые начинаются с некоторой точки $\eta_i$ и заканчиваются в следующей точке $\xi_i$. Интервалы, на которых это представление равно нулю, имеют вид $(\xi_j, \eta_{j+1})$. Типичное верхнее каноническое представление проиллюстрировано на рис. 12.

Очевидно, что

$$0=\int_a^b u_i(t)\,[\varphi(t)-\varphi^*(t)]\,d\mu(t), \qquad i=0, 1, \ldots, n, \tag{6.3}$$

и, следовательно, в качестве конечного шага доказательства теоремы 3.1 мы получаем, что $S(\varphi(t)-\varphi^*(t))\geqslant n+1$.

Мы будем различать два случая. Сначала предположим, что $\eta_{m+1}^*\leqslant\eta_{m+1}$. Тогда $\varphi(t)-\varphi^*(t)\leqslant 0$ для $t>\xi_m$. Отсюда видно, что $\varphi(t)-\varphi^*(t)$ может менять знак только в промежутке $a\leqslant t\leqslant\xi_m$. Однако только точки $\xi_i$ и $\eta_i$ $(i\leqslant m)$ могут быть точками изменения знака, так как только они являются точками, в которых $\varphi$ изменяет свои значения и, таким образом, $S(\varphi(t)-\varphi^*(t))\leqslant n+1$.

Предположим теперь, что имеет место другая возможность: $\eta_{m+1}^*>\eta_{m+1}$. Тогда аналогичные рассуждения снова приводят нас к тому, что $S(\varphi(t)-\varphi^*(t))\leqslant n+1$. Таким образом, для всех случаев мы установили, что

$$S(\varphi(t)-\varphi^*(t))=2m+1=n+1. \tag{6.4}$$

Этим заканчивается доказательство теоремы.

Применив рассуждения, использующие уравнения (4.4), можно, в силу неравенства $\omega^*>\omega$ показать, что в действительности

$$\eta_{m+1}^*<\eta_{m+1}. \tag{6.5}$$

Из (6.4) и (6.5) выводим, что узлы $\mathbf{c}(\omega^*)$ удовлетворяют соотношениям

$$\begin{aligned} &\xi_i(\omega^*) < \xi_i(\omega) < \xi_{i+1}(\omega^*), && i = 0, 1, \ldots, m, \\ &\eta_i(\omega) < \eta_{i+1}(\omega^*) < \eta_{i+1}(\omega), && i = 0, 1, \ldots, m, \end{aligned} \tag{6.6}$$

для $\underline{\omega} < \omega < \omega^* < \bar{\omega}$. Соотношение (6.6) также справедливо для $\omega = \underline{\omega}$ или $\omega^* = \bar{\omega}$.

Непрерывность $\xi_i(\omega)$ и $\eta_i(\omega)$ является следствием существования и единственности верхнего главного представления $\bar{\varphi}_\omega$ для $\mathbf{c}(\omega)$, а непрерывность в точках $\underline{\omega}$ и $\bar{\omega}$ также легко устанавливается. Из соотношений (6.6), очевидно, следует, что $\eta_i(\omega)$ и $\xi_i(\omega)$ являются монотонно убывающими функциями $\omega$. Приведенные рассуждения доказывают следующую теорему.

Теорема 6.2 ($n = 2m$). *При убывании $\omega$ от $\bar{\omega}$ до $\underline{\omega}$ $\xi_i(\omega)$ возрастает непрерывно от $\xi_i(\bar{\omega})$ до $\xi_i(\underline{\omega})$, а $\eta_i(\omega)$ возрастает непрерывно от $\eta_i(\bar{\omega})$ до $\eta_i(\underline{\omega})$.*

Типичное верхнее каноническое представление изображено на рис. 12 для $n = 4$. Интервалы $(\eta_i, \xi_i)$ являются отрезками между пунктирными линиями. На рисунке также показано поведение нижнего главного из канонических представлений, которые будут обсуждаться ниже.

Мы исследуем поведение нижнего канонического представления $\underline{\varphi}_\omega(t)$, когда $\omega$ пробегает отрезок $[\underline{\omega}, \bar{\omega}]$. Обозначим узлы $\underline{\varphi}_\omega(t)$ через

$$a < \alpha_0(\omega) < \beta_1(\omega) < \alpha_1(\omega) < \beta_2(\omega) < \ldots < \beta_{m+1}(\omega) < b.$$

В соответствии с этими обозначениями нижнее каноническое представление равно единице на интервалах $(\alpha_i, \beta_{i+1})$ $(i = 0, 1, \ldots, m)$ и нулю на интервалах $(\beta_i, \alpha_i)$ $(i = 0, 1, \ldots, m+1,\ \beta_0 = \alpha_0,\ \alpha_{m+1} = b)$. Применяя рассуждения, аналогичные тем, которые привели к (6.6), мы получаем, что

$$\begin{aligned} &\alpha_i(\omega) < \alpha_i(\omega^*) < \alpha_{i+1}(\omega), && i = 0, 1, \ldots, m, \\ &\beta_i(\omega^*) < \beta_{i+1}(\omega) < \beta_{i+1}(\omega^*), && i = 0, 1, \ldots, m, \end{aligned} \tag{6.7}$$

для $\omega < \omega^*$. Окончательно мы получаем следующую теорему.

Теорема 6.3 ($n = 2m$). *Когда $\omega$ пробегает $[\underline{\omega}, \bar{\omega}]$ слева направо*, то

$\alpha_i(\omega)$ *изменяется от* $\eta_i(\underline{\omega})$ *до* $\eta_{i+1}(\bar{\omega})$ $\quad (i = 0, 1, \ldots, m)$,

$\beta_i(\omega)$ *изменяется от* $\xi_{i-1}(\underline{\omega})$ *до* $\xi_i(\bar{\omega})$ $\quad (i = 1, 2, \ldots, m+1)$, *причем изменение каждой функции является непрерывным и строго возрастающим. (Здесь $\xi_{m+1}(\bar{\omega}) = b$, как было оговорено.)*

Теорема 6.4 ($n = 2m$). *Для каждого $t_0 \in (a, b)$ существует только два канонических представления (индекса $(n+1)/2$ или $(n+2)/2$) $\varphi^*$ и $\varphi^{**}$ такие, что $\varphi^* = 1$ на интервале, начинающемся с $t_0$, а $\varphi^{**} = 1$ на интервале, заканчивающемся в $t_0$.*

Доказательство. Из соотношения (6.6) следует, что $a = \xi_0(\bar{\omega}) < \xi_0(\underline{\omega}) < \xi_1(\bar{\omega}) < \ldots < \xi_m(\underline{\omega}) < \xi_{m+1}(\bar{\omega}) = b$. Если $t_0 \in (\xi_i(\bar{\omega}), \xi_i(\underline{\omega}))$, то из теоремы 6.2 следует, что $\xi_i(\omega)$ непрерывна и монотонна, так что существует единственное $\omega$ такое, что $t_0 = \xi_i(\omega)$. Получающееся в результате верхнее главное представление точки $\mathbf{c}(\omega)$ является требуемым $\varphi^{**}$. Можно применить те же рассуждения и в случае, когда $t_0 \in (\xi_{i-1}(\underline{\omega}), \xi_i(\bar{\omega}))$ для некоторого $i$, если использовать свойства $\beta_i(\omega)$, сформулированные в теореме 6.3. Существование $\varphi^*$ можно установить, если определить интервал вида $(\eta_i(\bar{\omega}), \eta_i(\underline{\omega}))$, содержащий $t_0$, и использовать непрерывность канонических представлений.

Этим заканчивается описание узловых точек канонических представлений и поведения этих представлений в случае $n = 2m$. Мы формулируем соответствующие результаты для случая $n = 2m + 1$ без доказательства.

Для каждого $\omega$ ($\underline{\omega} < \omega < \bar{\omega}$) существует верхнее каноническое представление индекса $(n+2)/2$, узлы которого располагаются в порядке

$$a < \eta_0(\omega) < \xi_0(\omega) < \eta_1(\omega) < \xi_1(\omega) < \ldots < \eta_{m+1}(\omega) < b,$$

и нижнее каноническое представление с узловыми точками

$$a < \beta_0(\omega) < \alpha_1(\omega) < \beta_1(\omega) < \alpha_2(\omega) < \ldots < \beta_{m+1}(\omega) < b.$$

Предполагается, что значение единица они принимают на интервалах $(\eta_i, \xi_i)$ и $(\alpha_i, \beta_i)$, $i = 0, 1, \ldots, m+1$, соответственно ($\alpha_0 = a$, $\xi_{m+1} = b$). Узлы верхнего и нижнего главных представлений перемежаются следующим образом:

$$\left.\begin{aligned} \eta_i(\bar{\omega}) < \eta_i(\underline{\omega}) < \eta_{i+1}(\bar{\omega}), \\ \xi_i(\bar{\omega}) < \xi_i(\underline{\omega}) < \xi_{i+1}(\bar{\omega}), \end{aligned}\right. \qquad i = 0, 1, \ldots, m.$$

Наконец, когда $\omega$ пробегает $[\underline{\omega}, \bar{\omega}]$, то

$\eta_i(\omega)$ убывает от $\eta_i(\underline{\omega})$ до $\eta_i(\bar{\omega})$ ($i = 0, 1, \ldots, m+1$),

$\xi_i(\omega)$ убывает от $\xi_i(\underline{\omega})$ до $\xi_i(\bar{\omega})$ ($i = 0, 1, \ldots, m$),

$\alpha_i(\omega)$ возрастает от $\eta_{i-1}(\underline{\omega})$ до $\eta_i(\bar{\omega})$ ($i = 1, \ldots, m+1$),

$\beta_i(\omega)$ возрастает от $\xi_{i-1}(\omega)$ до $\xi_i(\bar{\omega})$, $(i=0, 1, \dots, m+1)$ (как было оговорено ранее $\xi_{m+1}(\bar{\omega})=b$, $\eta_{m+1}(\omega)=b$, $\eta_0(\bar{\omega})=\xi_{-1}(\underline{\omega})= =a$). Функции $\eta_i(\omega)$, $\xi_i(\omega)$, $\alpha_i(\omega)$ и $\beta_i(\omega)$ непрерывны и строго монотонны.

Утверждение теоремы 6.4 справедливо также и для нечетных $n$.

## § 7. Двумерные сечения

Пусть $\{u_i\}_0^k$ для $k=n, n+1, n+2$ является $T$-системой $\mu$-интегрируемых функций на открытом интервале $(a, b)$.

Если $n=2m$, то для каждой точки $\mathbf{c}=(c_0, \dots, c_n) \in \operatorname{Int} \Phi_{n+1}$ имеются верхнее главное представление $\bar{\varphi}$ и нижнее главное представление $\underline{\varphi}$ с узлами

$$a<\bar{\eta}_1<\bar{\xi}_1<\bar{\eta}_2<\dots<\bar{\eta}_{m+1}<\bar{\xi}_{m+1}=b,$$

$$a=\underline{\eta}_0<\underline{\xi}_0<\eta_1<\dots<\underline{\eta}_m<\underline{\xi}_m<b$$

соответственно. Для каждого $t' \in (a, \bar{\eta}_1)$ существует каноническое представление точки $\mathbf{c}$ с узлами

$$a<\alpha_0(t')<\beta_1(t')<\alpha_1(t')<\beta_2(t')<\dots<\alpha_m(t')<\beta_{m+1}(t')<b,$$

где $\alpha_0(t')=t'$ (см. теорему 6.4).

Далее для каждого $\beta \in (\alpha_0(t'), \beta_1(t'))$ точка с координатами

$$c_i-\int_{t'}^{\beta} u_i(t)\, d\mu(t), \qquad i=0, 1, \dots, n, \tag{7.1}$$

лежит во внутренней части моментного пространства $\Phi_{n+1}(\beta, b) \subset \Phi_{n+1}$, которое определяется как сужение на $(\beta, b)$ тех функций $\varphi \in \Phi$, которые равны нулю на $(\alpha, \beta)$.

Моментная точка (7.1) имеет нижнее главное представление относительно $\Phi_{n+1}(\beta, b)$ с узлами

$$\beta<\beta_1(t')<\alpha_1(t')<\beta_2(t')<\dots<\alpha_m(t')<\beta_{m+1}(t')<b.$$

Если теперь рассмотреть верхнее главное представление точки (7.1) относительно интервала $(\beta, b)$ и присоединить к этому интервалу интервал $(t', \beta)$ со значением единица, то получившееся в результате представление будет иметь индекс $(n+3)/2$ и равняться единице на интервале, примыкающем к b. Обозначим функцию для этого представления через $\varphi(t;\, t', \beta)$. Как и в § 7 гл. II, положим

$$B=\{(t', \beta) \mid t' \in (a, \bar{\eta}_1), \beta \in (t', \beta_1(t'))\},$$

$$A=\left\{(c_{n+1}, c_{n+2}) \mid c_i=\int_a^b u_i \varphi\, d\mu,\ i=n+1,\ n+2,\ \varphi \in V(\mathbf{c})\right\},$$

где $V(\mathbf{c})=\left\{\varphi \mid c_i=\int_a^b u_i \varphi\, d\mu_i\ i=0, 1, \dots, n,\ \varphi \in \Phi\right\}$.

Мы докажем, что отображение $\delta$ $B$ в $A$, определяемое соотношением

$$\delta(t', \beta)=(c_{n+1}, c_{n+2}),$$

где $c_i = \int_a^b u_i(t)\,\varphi(t; t', \beta)\,d\mu(t)$, $i = n+1$, $n+2$, взаимно однозначно, непрерывно и является отображением на внутреннюю часть $A$. Доказательство взаимной однозначности отображения $\delta$ проводится следующим образом. Если $\delta(t', \beta') = \delta(t'', \beta'')$, то $\varphi(t; t', \beta')$ и $\varphi(t; t'', \beta'')$ имеют общие первые $n+3$ момента и, очевидно, оба образуют верхние главные представления одной и той же точки из $\Phi_{n+3}$. В силу теоремы 3.1 это возможно только тогда, когда $\varphi(t; t', \beta') = \varphi(t; t'', \beta'')$.

Если $\varphi$ — любое верхнее представление точки $\bar{c}$ индекса $(n+3)/2$, то $\varphi - \bar{\varphi}$ меняет знак ровно $n+1$ раз, так что в $(a, \bar{\eta}_1)$ начинается только один интервал со значением $\varphi$, равным единице. Поэтому $(t', \beta') = (t'', \beta'')$ и отображение $\delta$ взаимно однозначно.

Отображение $\delta$ является «отображением на», так как каждая точка $(c_0, \ldots, c_{n+2})$ имеет верхнее главное представление $\bar{\varphi}_1$ индекса $(n+3)/2$, у которого первый интервал со значением единица должен начинаться в некоторой точке $t' \in (a, \bar{\eta}_1)$ и заканчиваться до $\bar{\beta}(t')$. Первый интервал со значением единица начинается в $(a, \bar{\eta}_1)$, потому что $\bar{\varphi}_1 - \bar{\varphi}$ меняет знак $n+1$ раз. Он должен заканчиваться до $\beta(t')$, так как разность между $\bar{\varphi}_1$ и каноническим представлением, начинающимся в $t'$, изменяет знак $n+1$ раз. Для доказательства непрерывности предположим, что $(t'_n, \beta_n) \to (t', \beta)$. Очевидно, можно выбрать подпоследовательность такую, что $\varphi(t; t'_m, \beta_m)$ сходится к некоторой функции $\varphi$, индекс которой не превосходит $(n+3)/2$. В действительности индекс должен равняться $(n+3)/2$, так как $\varphi - \bar{\varphi}$ изменяет знак $n+1$ раз. Также $\varphi$ не может быть нижним главным представлением точки из $\Phi_{n+3}$. Наконец, представление индекса $(n+3)/2$ с единичным значением на интервале $(t', \beta)$ и моментами $(c_0, \ldots, c_n)$ однозначно определено, так что $\varphi(t) = \varphi(t; t', \beta)$. Таким образом, все пределы $\varphi(t; t'_n, \beta_n)$ одинаковы, откуда следует, что отображение $\delta$ непрерывно.

Можно провести такой же анализ, использующий нижние представления, если рассмотреть $t' \in (\underline{\xi}_m, b)$ и $\eta \in (\eta_m(t'), t')$. Если ограничить внимание нижними каноническими представлениями, то приведенные выше рассуждения применимы также для $n = 2m+1$ Детали мы опускаем.

## § 8. Неравенства чебышевского типа

Пусть $c^0 \in \operatorname{Int} \Phi_{n+1}$, и пусть $V(c^0)$ обозначает множество всех функций $\varphi \in \Phi$, для которых $c_i^0 = \int_a^b u_i(t)\,\varphi(t)\,d\mu(t)$ $(i = 0, 1, \ldots, n)$.

Мы рассмотрим три типа связанных между собой экстремальных задач.

(1) Пусть $\Omega(t)$ выбрано так, что

$$(u_0, u_1, \ldots, u_n, \Omega) \tag{8.1}$$

является слабой $T$-системой.

Сначала мы исследуем следующую задачу: определить те $\varphi \in V(c^0)$, для которых достигается минимум и максимум интеграла

$$\int_a^b \Omega(t)\,\varphi(t)\,d\mu(t). \tag{8.2}$$

Решение является простым. Значение (8.2) изменяется на отрезке $R(\mathbf{c}^0) = [\underline{\omega}, \overline{\omega}]$. Очевидно, максимум $\overline{\omega}$ достигается для $\overline{\varphi}$ — верхнего главного представления $\mathbf{c}^0$, а минимум для $\underline{\varphi}$ — нижнего главного представления. Экстремальные решения $\overline{\varphi}$ и $\underline{\varphi}$ единственны при условии, что $\{u_0, u_1, \ldots, u_n, \Omega\}$ является $T$-системой.

(2) Если отбросить предположение (8.1) и допустить только, что $\Omega(t)$ непрерывна и линейно не зависит от $u_0, u_1, \ldots, u_n$, то максимум (8.2) достигается для $\tilde{\varphi} \in V(\mathbf{c}^0)$, задаваемой соотношением

$$\tilde{\varphi}(t) = \begin{cases} 1, & \text{если } \Omega(t) > \tilde{u}(t), \\ 0, & \text{если } \Omega(t) < \tilde{u}(t), \end{cases} \tag{8.3}$$

где $\tilde{u}(t)$ — определенным образом выбранный многочлен. Функция $\tilde{\varphi}(t)$ не определена, когда $\Omega(t) = \tilde{u}(t)$.

Доказательство (8.3) проводится следующим образом. Рассмотрим замкнутое ограниченное выпуклое множество размерности $n+2$

$$R_{n+2} = \left\{ \left( \int_a^b u_0 \varphi(t)\, d\mu(t), \ldots, \int_a^b u_n(t)\varphi(t)\, d\mu(t), \int_a^b \Omega(t)\varphi(t)\, d\mu(t) \right) \right\}, \tag{8.4}$$

которое получается при изменении $\varphi \in \Phi$. Мы построим опорную плоскость к $R_{n+2}$ в точке

$$\left( c_0^0, c_1^0, \ldots, c_n^0, \max_{\varphi \in V(\mathbf{c}^0)} \int_a^b \Omega(t)\varphi(t)\, d\mu(t) \right). \tag{8.5}$$

Аналогично (2.2) можно получить неравенства

$$\int_a^b [\alpha\Omega(t) - \tilde{u}(t)](\tilde{\varphi}(t) - \varphi(t))\, d\mu(t) \geqslant 0 \tag{8.6}$$

для всех $\varphi \in \Phi$, где $\tilde{\varphi}$ — тот элемент $V(\mathbf{c}^0)$, на котором достигается $\max\limits_{\varphi \in V(\mathbf{c}^0)} \int_a^b \Omega(t)\varphi(t)\, d\mu(t)$, а $\tilde{u}$ — соответствующим образом выбранный многочлен.

Очевидно, $\alpha \neq 0$, так как в противном случае мы получили бы из (8.6), что $\mathbf{c}^0 \in \mathscr{D}\Phi_{n+1}$, что противоречило бы предположению. Более того, мы утверждаем, что $\alpha > 0$. Предположим, что справедливо противное, т. е. что $\alpha < 0$. Тогда из (8.6) для $\varphi \in V(\mathbf{c}^0)$ можно вывести, что

$$\max_{\varphi \in V(\mathbf{c}^0)} \int_a^b \Omega(t)\varphi(t)\, d\mu(t) = \min_{\varphi \in V(\mathbf{c}^0)} \int_a^b \Omega(t)\varphi(t)\, d\mu(t).$$

Это означает, что точка (8.5) из $R_{n+2}$ совпадает с точкой (8.5), в которой max заменен на min. Отсюда следует, что в (8.6) имеет место равенство. Это может быть только тогда, когда $\Omega(t)$ линейно зависит от $u_0, u_1, \ldots, u_n$, что противоречит предположениям.

Поэтому $\alpha > 0$ и, не умаляя общности, мы можем предположить, что $\alpha = 1$. Характеризация (8.3) легко следует из (8.6) (ср. доказательство теоремы 2.2).

(3) Мы ищем минимум и максимум значения интеграла

$$\int_a^{\xi_0} \Omega(t)\,\varphi(t)\,d\mu(t), \qquad \varphi \in V(\mathbf{c}^0), \tag{8.7}$$

где $\xi_0$ — фиксировано ($a < \xi_0 < b$).

Предположим, что $u_0, u_1, \ldots, u_n$ и $\Omega(t)$ удовлетворяют условию 6.1 гл. I на интервале $(a, b)$. Пусть $\tilde{\varphi}_{\xi_0}$ обозначает каноническое представление точки $\mathbf{c}^0$ такое, что интервал с единичным значением $\tilde{\varphi}_{\xi_0}$ начинается в $\xi_0$ (его существование и единственность гарантируются теоремой 6.4). Начиная с этого места, мы ограничим наше внимание случаем $n = 2m$ и предположим, что узлы $\tilde{\varphi}_{\xi_0}$ расположены в следующем порядке:

$$a = \tilde{\eta}_0 < \tilde{\xi}_0 < \tilde{\eta}_1 < \tilde{\xi}_1 < \ldots < \tilde{\eta}_{i_0} = \xi_0 < \tilde{\xi}_{i_0} < \ldots < \tilde{\eta}_{m+1} < \tilde{\xi}_{m+1} = b.$$

В силу теоремы 6.1 гл. I существует многочлен $\tilde{u}(t)$, удовлетворяющий условиям

$$\tilde{u}(t) = \begin{cases} \Omega(t), & t = \tilde{\xi}_0, \quad \tilde{\eta}_1, \ldots, \tilde{\eta}_{i_0-1}, \quad \tilde{\xi}_{i_0-1}, \\ 0, & t = \tilde{\xi}_{i_0}, \quad \tilde{\eta}_{i_0+1}, \ldots, \tilde{\eta}_{m+1}, \end{cases} \tag{8.8}$$

$$\begin{aligned} &\tilde{u}(t) \geqslant \Omega(t), \quad (\tilde{\eta}_0, \tilde{\xi}_0), \quad (\tilde{\eta}_1, \tilde{\xi}_1), \ldots, (\tilde{\eta}_{i_0-1}, \tilde{\xi}_{i_0-1}), \\ &\tilde{u}(t) \leqslant \Omega(t), \quad (\tilde{\xi}_0, \tilde{\eta}_1), \quad (\tilde{\xi}_1, \tilde{\eta}_2), \ldots, (\tilde{\xi}_{i_0-1}, \tilde{\eta}_{i_0}), \\ &\tilde{u}(t) \geqslant 0, \quad t \in (\tilde{\eta}_{i_0}, \tilde{\xi}_{i_0}), \ldots, (\tilde{\eta}_{m+1}, b), \\ &\tilde{u}(t) \leqslant 0, \quad t \in (\tilde{\xi}_{i_0}, \tilde{\eta}_{i_0+1}), \ldots, (\tilde{\xi}_m, \tilde{\eta}_{m+1}). \end{aligned} \tag{8.9}$$

Пусть $\varphi \in V(\mathbf{c}^0)$. Тогда, очевидно,

$$\varphi(t) - \tilde{\varphi}_{\xi_0}(t) \begin{cases} \leqslant 0, & t \in (\tilde{\eta}_i, \tilde{\xi}_i) \quad (i = 0, 1, \ldots, m+1), \\ \geqslant 0, & t \in (\tilde{\xi}_i, \tilde{\eta}_{i+1}) \quad (i = 0, 1, \ldots, m). \end{cases}$$

Заметим также, что требования (8.8) не накладывают ограничений на значение $\tilde{u}$ в точке $\tilde{\eta}_{i_0} = \xi_0$. Многочлен $\tilde{u}$ равен $\Omega$ и 0 на $(a, \xi_0)$

и $(\xi_0, b)$, соответственно, только в узловых точках, описанных в (8.8). С помощью (8.9) мы получаем

$$\int_a^{\xi_0} \Omega(t)\,[\varphi(t) - \tilde{\varphi}_{\xi_0}(t)]\,d\mu(t) = \int_a^{\xi_0} [\Omega(t) - \tilde{u}(t)]\,[\varphi(t) - \tilde{\varphi}_{\xi_0}(t)]\,d\mu(t) -$$

$$- \int_{\xi_0}^{b} \tilde{u}(t)\,[\varphi(t) - \tilde{\varphi}_{\xi_0}(t)]\,d\mu(t) > 0. \quad (8.10)$$

Такие же рассуждения проводятся и для случая $n = 2m + 1$. В итоге мы можем получить соответствующую верхнюю оценку. Резюмируем проведенный анализ в формулировке следующей теоремы.

Теорема 8.1 (Крейн [1951]). *Пусть* $c^0 \in \operatorname{Int} \Phi_{n+1}$, *и предположим, что* $\xi_0 \in (a, b)$ *фиксировано. Пусть на* $(a, b)$ *выполнено условие* 6.1 *гл I. Обозначим через* $\tilde{\varphi}_{\xi_0}(t)$ *единственное каноническое представление, у которого интервал с единичным значением начинается в* $\xi_0$, *и пусть* $\overset{*}{\varphi}_{\xi_0}(t)$ — *единственное каноническое представление, у которого интервал с единичным значением заканчивается в* $\xi_0$. *Тогда для любого* $\varphi \in V(c^0)$, *отличного от* $\overset{*}{\varphi}_{\xi_0}$ *и* $\tilde{\varphi}_{\xi_0}$, *мы имеем*

$$\int_a^{\xi_0} \Omega(t)\,\tilde{\varphi}_{\xi_0}(t)\,d\mu(t) < \int_a^{\xi_0} \Omega(t)\,\varphi(t)\,d\mu(t) < \int_a^{\xi_0} \Omega(t)\,\overset{*}{\varphi}_{\xi_0}(t)\,d\mu(t), \qquad (8.11)$$

*кроме того случая, когда первые два члена равны нулю. Это происходит тогда, когда индекс* $\tilde{\varphi}_{\xi_0}$ *равен* $(n+2)/2$ *и* $\tilde{\varphi}_{\xi_0}(t) = 0$ *для* $t < \xi_0$.

*Результаты* §5 *гл.* III *для непрерывных функций на компактном интервале переносятся на рассматриваемый случай лишь с небольшими изменениями формулировок.*

## § 9. Множества мер, определяемых моментными неравенствами

В экстремальной задаче этого параграфа можно, конечно, узнать общую задачу Неймана — Пирсона (см., например, Данциг и Вальд [1951]). Решение задачи, сформулированной в (8. 7), принадлежит Крейну.

Пусть $u_i$, $0 \leqslant i \leqslant n+1$, $-n+2$, $\mu$-интегрируемые функции на конечном открытом интервале $(a, b)$, которые удовлетворяют следующему условию.

Условие 9.1.

(a) $\{u_i\}_0^n$ и $\{u_i\}_0^{n+1}$ — $T$-системы,

(b) для каждого $i = 0, 1, \ldots, n+1$ определители $(n+1)$-го порядка, образованные из $u_0, \ldots, u_{i-1}, u_{i+1}, \ldots, u_{n+1}$ все имеют строго один знак $\varepsilon_i$.

Теорема 9.1. *Пусть система* $\{u_i\}_0^{n+1}$ *удовлетворяет условию* 9.1 *и* $V$—*класс функций* $\varphi \in \Phi$, *удовлетворяющих соотношениям*

$$\int_a^b u_i \varphi \, d\mu \begin{cases} = d_i, & i \in A, \\ \leqslant d_i, & i \in B, \\ \geqslant d_i, & i \in C, \end{cases}$$

*где $A, B, C$ образуют разбиение множества $0, 1, \ldots, n$. Тогда, если $V$ непусто то*

$$\underline{\alpha} = \inf_{\varphi \in V} \int_a^b u_{n+1}\varphi\, d\mu \tag{i}$$

*достигается на $V$ только на функции $\underline{\varphi}$ индекса $\leqslant (n+1)/2$, которая соответствует нижнему главному представлению в том случае, когда индекс равен $(n+1)/2$. Кроме того,*

$$\bar{\alpha} = \sup_{\varphi \in V} \int_a^b u_{n+1}\varphi\, d\mu \tag{ii}$$

*достигается на $V$ только на функции $\bar{\varphi}$ индекса $\leqslant (n+1)/2$, которая соответствует верхнему главному представлению в том случае, когда этот индекс равен $(n+1)/2$.*

Доказательство. Мы будем доказывать только часть (i); доказательство (ii) совершенно аналогично. Так как $u_{n+1}$ $\mu$-интегрируема, то $\underline{\alpha}$-конечно, и поэтому существует последовательность $\{\varphi_k\}$ такая, что $\underline{\alpha} = \lim\limits_{k\to\infty} \int_a^b u_{n+1}\varphi_k\, d\mu$.

Но $\Phi$ слабо компактно так, что мы можем выбрать подпоследовательность, $\varphi_{k'}$ и функцию $\underline{\varphi}$ так, чтобы $\lim\limits_{k'\to\infty} \int_a^b u_i\varphi_{k'}\, d\mu = \int_a^b u_i\underline{\varphi}\, d\mu$, $i = 0, 1, \ldots, n+1$.

Если $\left(\int_a^b u_0\underline{\varphi}\, d\mu, \ldots, \int_a^b u_n\underline{\varphi}\, d\mu\right)$ — граничная точка $\Phi_{n+1}$, то $\underline{\varphi}$ однозначно определена и имеет индекс, не больший чем $n/2$, в то время как, если эта точка является внутренней точкой $\Phi_{n+1}$, то в силу (8.1) $\underline{\varphi}$ является нижним главным представлением. Единственность, так же как и большинство последующих результатов, основывается на следующей лемме. Пусть $\underline{c}_i$ и $\bar{c}_i$ определяются формулами

$$\underline{c}_i = \inf_{\varphi \in V_i(c)} \int_a^b u_i\varphi\, d\mu,$$

$$\bar{c}_i = \sup_{\varphi \in V_i(c)} \int_a^b u_i\varphi\, d\mu,$$

где $V_i(c) = \{\varphi \mid \varphi \in \Phi, \int_a^b u_j\varphi\, d\mu = c_j,\ 0 \leqslant j \leqslant n+1,\ j \neq i\}$.

Лемма 9.1. *Если функции $u_i$, $0 \leqslant i \leqslant n+1$, удовлетворяют условию 9.1 (a) и условию 9.1 (b) для двух фиксированных значений $i$ и $j\ (i \neq j)$, то $(-1)^{j-i}\varepsilon_i\varepsilon_j\underline{c}_i$ и $(-1)^{j-i}\varepsilon_i\varepsilon_j\bar{c}$ — строго убывающие функции координаты $c_j$.*

Доказательство. Мы докажем только первое утверждение. Предположим, что координаты, определяющие $\underline{c}_i$, все фиксированы, за исключением $c_j$, которую мы увеличим до $c_j^*$. Новое значение $\underline{c}_i$ обозначим $\underline{c}_i^*$, и пусть $\varphi$ и $\varphi^*$ будут обозначать функции, которые соответствуют значениям $\underline{c}_i$ и $\underline{c}_i^*$. Как $\varphi$, так и $\varphi^*$ имеют индекс, не больший чем $(n+1)/2$, Более того, если их индексы равны $(n+1)/2$, то они обе являются верхними главными или нижними главными, так что разность $\varphi - \varphi^*$ имеет $m$ изменений знака относительно $\mu$, где $m \leqslant n$ (ср. доказательство теоремы 3 1). Поэтому можно написать

$$\int_a^b u_k \varphi \, d\mu - \int_a^b u_k \varphi^* d\mu = \pm \sum_{p=0}^{m} (-1)^p \int_{x_p}^{x_{p+1}} u_k \, |\varphi - \varphi^*| \, d\mu$$

для $k = 0, 1, \ldots, n+1$ и разложить один или более из интервалов $(x_p, x_{p+1})$ так, чтобы получить

$$\int_a^b u_k \varphi \, d\mu - \int_a^b u_k \varphi^* d\mu = \sum_{p=0}^{n} \eta_p \int_{y_p}^{y_{p+1}} u_k \, |\varphi - \varphi^*| \, d\mu,$$

где $\eta_p = \pm 1$, $y_0 = a$, $y_{n+1} = b$ и $\varphi(t) - \varphi^*(t) \not\equiv 0$ на любом из интервалов $(y_p, y_{p+1})$.

Мы разложим определитель

$$\begin{vmatrix} 0, & \int_{y_0}^{y_1} u_0 \, |\varphi - \varphi^*| \, d\mu, & \ldots, & \int_{y_n}^{y_{n+1}} u_0 \, |\varphi - \varphi^*| \, d\mu \\ \cdots & \cdots & \cdots & \cdots \\ c_j - c_j^* & & & \\ \cdots & \cdots & \cdots & \cdots \\ \underline{c}_i - \underline{c}_i^* & & & \\ \cdots & \cdots & \cdots & \cdots \\ 0, & \int_{y_0}^{y_1} u_{n+1} \, |\varphi - \varphi^*| \, d\mu, & \ldots, & \int_{y_n}^{y_{n+1}} u_{n+1} \, |\varphi - \varphi^*| \, d\mu \end{vmatrix}$$

по первому столбцу, замечая, что его значение равно нулю, так как первый столбец можно выразить в виде линейной комбинации остальных столбцов. Утверждение леммы теперь следует из того, что миноры ненулевых членов в первом столбце имеют знаки $\varepsilon_i$ и $\varepsilon_j$ соответственно.

Утверждение единственности в теореме 9.1 доказывается с помощью приведенной выше леммы. Заметим, что если $\underline{\varphi}$ и $\varphi$ оба соответствуют значению $\underline{\alpha}$ и различны, то $\int u_j \underline{\varphi} \, d\mu \neq \int u_j \varphi \, d\mu$ для некоторого $j$ $(0 \leqslant j \leqslant n)$. Однако в этом случае можно применить лемму 9.1, чтобы уменьшить минимальное значение $\underline{\alpha}$, и таким образом, мы получим противоречие.

Сохраняя обозначения теоремы 9.1 и полагая

$$\underline{D} = \left\{ i \,\middle|\, \int u_i \underline{\varphi} \, d\mu = c_i, \quad 0 \leqslant i \leqslant n \right\},$$

$$\overline{D} = \left\{ i \,\middle|\, \int u_i \overline{\varphi} \, d\mu = c_i, \quad 0 \leqslant i \leqslant n \right\},$$

мы получаем следующие следствия. (Их доказательства являются дословным повторением доказательств следствий 5.1 (а) и (b) гл. III и будут опущены).

Следствие 9.1(а). *Пусть $V_1$ — класс тех $\varphi$, для которых*

$$\int_a^b u_i \, \varphi \, d\mu \begin{cases} = d_i, & i \in A \cap \underline{D}, \\ \leqslant d_i, & i \in B \cap \underline{D}, \\ \geqslant d_i, & i \in C \cap \underline{D}, \end{cases}$$

*и пусть $V_2$ определяется аналогично, только вместо $\underline{D}$ используется $\overline{D}$. Тогда*

(i) $\underline{\varphi}$ *— единственная функция в $V_1$, на которой достигается значение*

$$\underline{\beta} = \inf_{\varphi \in V_1} \int_a^b u_{n+1} \varphi \, d\mu,$$

(ii) $\overline{\varphi}$ *— единственная функция в $V_2$, на которой достигается значение*

$$\overline{\beta} = \sup_{\varphi \in V_2} \int_a^b u_{n+1} \varphi \, d\mu.$$

Следствие 9.1 (b). *Пусть $\overline{\varphi}$, $\underline{\varphi}$, $A$, $B$, $C$, $\underline{D}$ и $\overline{D}$ такие, как в следствии* 9.1 (a), *и пусть $v_0, v_1, \ldots, v_n$ — функции, для которых $v_0, \ldots, v_n, u_{n+1}$ удовлетворяют тем же условиям, что и $u_0, \ldots, u_n, u_{n+1}$. Если $e_i = \int v_i \underline{\varphi} \, d\mu$ и $f_i = \int v_i \overline{\varphi} \, d\mu$ $(i = 0, \ldots, n)$, то $\underline{\varphi}$ $(\overline{\varphi})$ — единственная функция из $W_1$ $(W_2)$, на которой достигается значение*

$$\underline{\gamma} = \inf_{\varphi \in W_1} \int_a^b u_{n+1} \varphi \, d\mu \left( \overline{\gamma} = \sup_{\varphi \in W_2} \int_a^b u_{n+1} \varphi \, d\mu \right),$$

*где $W_1$ — класс тех $\varphi$, которые удовлетворяют условию*

$$\int_a^b v_i \varphi \, d\mu \begin{cases} = e_i, & i \in A \cap \underline{D}, \\ \leqslant e_i, & i \in B \cap \underline{D}, \\ \geqslant e_i, & i \in C \cap \underline{D}, \end{cases}$$

*а $W_2$ определяется аналогично, только вместо $e_i$ используются $f_i$, а вместо $\underline{D}$ используется $\overline{D}$.*

Так же, как и в § 5 гл. III, предположим, что систему $\{u_i\}_0^{n+1}$ можно дополнить еще одной $\mu$-интегрируемой функцией $v_i$ так, чтобы получившееся в результате множество функций удовлетворяло следующему условию.

Условие 9.2.

(a) Системы функций $\{u_0, \ldots, u_i, v_i, u_{i+1}, \ldots, u_n\}$ и $\{u_0, \ldots, u_i; v_i, u_{i+1}, \ldots, u_{n+1}\}$ образуют $T$-системы на $(a, b)$. (b) Определители $(n+1)$-го порядка, образованные из каждого фиксированного упорядоченного множества из $n+1$ функций, выбранных из $\{u_0, \ldots, u_i, v_i, u_{i+1}, \ldots, u_n\}$, имеют один знак (ср. стр. 103).

В этих условиях справедлива

Теорема 9.2. *Если функции $u_0 \ldots u_i, v_i, u_{i+1}, \ldots, u_{n+1}$ удовлетворяют условиям* 9.1 *и* 9.2, *$V$, $\underline{\varphi}$, $\overline{\varphi}$ определены так же, как и в теореме* 9.1, *то*

$$\underline{\delta}_i = \inf_{\varphi \in V} (-1)^{n-i} \int_a^b v_i \varphi \, d\mu \quad \text{и} \quad \overline{\delta}_i = \sup_{\varphi \in V} (-1)^{n-i} \int_a^b v_i \varphi \, d\mu$$

*достигаются только на $\underline{\varphi}$ и $\overline{\varphi}$ соответственно*

Доказательство аналогично доказательству теоремы 5.2 гл. III.

## § 10. Обобщения

Теория, развитая в предыдущих параграфах, может быть обобщена следующим образом. Пусть $\nu$ и $\lambda$ — обобщенные меры ограниченной вариации, определенные на $(a, b)$ и удовлетворяющие условию

$$\nu > \lambda \tag{10.1}$$

(это означает, что $\nu(G) - \lambda(G) > 0$ для каждого открытого множества $G \subset (a, b)$). Если бы $\nu$ и $\lambda$ были положительными мерами, то из (10.1) следовало бы, что $\lambda$ абсолютно непрерывна относительно $\nu$. Вообще говоря, (10.1) есть более сильное ограничение, чем абсолютная непрерывность. Предположим в дальнейшем, что $\nu$—$\lambda$— положительная мера без атомов, т. е. $\int_a^t [d\nu(\tau) - d\lambda(\tau)]$ —непрерывная функция $t$. Пусть $\Sigma = \{\sigma \mid \sigma$ — обобщенная мера на $(a, b)$, для которой $\lambda \leq \sigma \leq \nu\}$.

Мы исследуем строение выпуклого множества

$$\Sigma_{n+1} = \left\{ (s_0, \dots, s_n) \mid s_i = \int_a^b u_i d\sigma, \quad i = 0, 1, \dots, n, \ \sigma \in \Sigma \right\}. \tag{10.2}$$

Предположим, что $u_i$, $i = 0, 1, \dots, n$, интегрируемы как по отношению к $|\nu|$. так и по отношению к $|\lambda|$.

Анализ $\Sigma_{n+1}$ можно свести к анализу $\Phi_{n+1}$. Действительно, определим $\mu = \nu - \lambda$ и для каждого $\sigma \in \Sigma$ рассмотрим $\theta = \sigma - \lambda$. Очевидно, $\theta$ абсолютно непрерывна относительно $\mu$, и поэтому ее можно представить в виде

$$\theta(U) = \int_U \varphi(t)\, d\mu(t), \tag{10.3}$$

где $\varphi(t)$ $\mu$-интегрируема и $0 \leq \varphi(t) \leq 1$. Обратно, для каждой $\varphi \in \Phi$ можно построить $\sigma = \theta + \lambda \in \Sigma$, где $\theta$ определено, как в (10.3).

Эти преобразования $\Sigma_{n+1}$ равносильны сдвигу этого множества. В частности, мы переходим от $\Sigma_{n+1}$ к $\Phi_{n+1}$ при помощи отображения

$$\mathbf{s} = (s_0, s_1, \dots, s_n) \to \mathbf{c} = (c_0, c_1, \dots, c_n),$$

$$\mathbf{s} \in \Sigma_{n+1}, \quad c_i = s_i + \alpha_i, \qquad i = 0, 1, \dots, n,$$

где

$$-\alpha_i = \int_a^b u_i(t)\, d\lambda(t), \qquad i = 0, 1, \dots, n. \tag{10.4}$$

Геометрию $\Phi_{n+1}$, введенную ранее, совершенно аналогично можно определить на $\Sigma_{n+1}$. Каноническое представление в $\Phi_{n+1}$ переходит в меру $\sigma$ (не обязательно положительную), которая на последовательных отрезках совпадает с $\nu$ и $\lambda$; нулевой (единичный) интервал для $\varphi$ соответствует интервалу, на котором $\sigma = \lambda$ ($\sigma = \nu$). Узловые точки $\varphi$ соответствуют тем точкам, на которых $\sigma$ переходит от совпадения с $\nu$ к совпадению с $\lambda$ или наоборот.

Таким образом, видно, что каждый результат, справедливый для $\Phi_{n+1}$, может быть сформулирован и для $\Sigma_{n+1}$. Мы оставляем читателю подробно сформулировать соответствующие теоремы. Мы только сформулируем аналог теоремы 2.3, так как он потребуется в следующем параграфе.

Теорема 10.1. *Необходимое и достаточное условие для того, чтобы* $\mathrm{c} = (c_0, \ldots, c_n)$ *принадлежала* $\Sigma_{n+1}$, *состоит в том, что*

$$\sum_{k=0}^{n} a_k c_k \leq \int_a^b u_+(t)\,d\nu(t) - \int_a^b u_-(t)\,d\lambda(t)$$

*для каждого многочлена*

$$u(t) = \sum_{k=0}^{n} a_k u_k(t),$$

*где* $u_+(t) = \max(0, u(t))$, *а* $u_-(t) = \max(0, -u(t))$.

Доказательство. Точка $(c_0, c_1, \ldots, c_n)$ содержится в $\Sigma_{n+1}$ тогда и только тогда когда $(c_0 + \alpha_0, \ldots, c_n + \alpha_n) \in \Phi_{n+1}(\nu - \lambda)$. В силу теоремы 2.3 это эквивалентно тому, что

$$\sum_{k=0}^{n} a_k (c_k + \alpha_k) \leq \int_a^b u_+(\nu - \lambda), \tag{10.5}$$

где $u = \sum_{k=0}^{n} a_k u_k$. Используя (10.4), мы легко приводим (10.5) к желаемому виду:

$$\sum_{k=0}^{n} a_k c_k \leq \int_a^b u_+ d\nu - \int_a^b u_- d\lambda.$$

## § 11. Задача минимизации

В этом параграфе элегантный анализ задачи о нахождении экстремума принадлежит Крейну [1951]. Он был обусловлен более ранними результатами Золотарева и других авторов, приведенными ниже, в примере 11.2. Более подробные ссылки можно найти у Ахиезера [1965].

Результаты и методы этого параграфа, по существу, основываются на геометрии, изложенной в предшествующих параграфах. Задача, рассматриваемая здесь, относится к $(\lambda, \nu)$-проблеме. Пусть нам даны вещественные числа

$$\gamma_0, \gamma_1, \ldots, \gamma_n, \text{ для которых } \sum_{k=0}^{n} \gamma_k^2 > 0$$

и две непрерывные функции $\alpha$ и $\beta$ ограниченной вариации на $(a, b)$, удовлетворяющие условию

$$d\alpha > 0, \; |d\beta| \leqslant d\alpha. \tag{11.1}$$

(Эти неравенства следует понимать в смысле (10.1).)

Пусть $U$ обозначает класс многочленов

$$u(t) = \sum_{k=0}^{n} a_k u_k(t),$$

коэффициенты которых удовлетворяют условию

$$\sum_{k=0}^{n} a_k \gamma_k = 1. \tag{11.2}$$

Требуется найти минимум на всем $\mathcal{U}$ функционала

$$\mathcal{F}(u) = \int_a^b |u(t)|\, d\alpha(t) + \int_a^b u(t)\, d\beta(t) \tag{11.3}$$

и охарактеризовать многочлен $u(t)$, для которого этот минимум достигается.

В силу (11.1) и (11.2) легко видеть, что $\inf\limits_{u \in \mathcal{U}} F(u) \geqslant 0$. Обозначим $\inf\limits_{u \in \mathcal{U}} \mathcal{F}(u)$ через $I$.

Пример 11.1. Представляет интерес частный случай, когда $\gamma_0 = \gamma_1 = \ldots = \gamma_{n-1} = 0$, $\gamma_n = 1$, $d\beta = 0$. В этом случае функционал (11.3) и задача поиска минимума принимает следующий вид: найти коэффициенты $a_k$ $(k = 0, 1, \ldots, n-1)$, для которых интеграл

$$\int_a^b |u_n(s) + a_0 u_0(t) + a_1 u_1(t) + \ldots + a_{n-1} u_{n-1}(t)|\, d\alpha(t)$$

принимает минимальное значение.

Продолжая рассуждения в общем случае, положим

$$\nu = \beta + \alpha, \quad \lambda = \beta - \alpha. \tag{11.4}$$

Тогда условие (11.1) примет вид

$$\lambda < 0 < \nu, \tag{11.5}$$

а (11.3) преобразуется в

$$F(u) = \int_a^b u_+(t)\, d\nu(t) - \int_a^b u_-(t)\, d\lambda(t), \tag{11.6}$$

где $u_+(t) = \max(u(t), 0)$, а $u_-(t) = \max(-u(t), 0)$.

Пусть $u(t) = \sum\limits_{k=0}^{n} a_k u_k(t)$ — произвольный многочлен, для которого $\sum\limits_{k=0}^{n} a_k \gamma_k > 0$.

Многочлен $u^*(t) = u(t) / \sum\limits_{k=0}^{n} a_k \gamma_k$ удовлетворяет условию (11.2), и поэтому

$$\mathcal{F}(u^*) = \frac{\mathcal{F}(u)}{\sum\limits_{k=0}^{n} a_k \gamma_k} \geqslant I$$

или

$$l\sum_{k=0}^{n} a_k\gamma_k \leqslant \int_a^b u_+(t)\,d\nu(t) - \int_a^b u_-(t)\,d\lambda(t), \tag{11.7}$$

если только $0 \leqslant l \leqslant I$.

Неравенство (11.7), очевидно, справедливо, если $\sum\limits_{k=0}^{n} a_k\gamma_k \leqslant 0$, и, таким образом (11.7) справедливо для любого многочлена $u$. Соотношение (11.7), вместе с теоремой 10.1, показывает, что

$$(l\gamma_0\ l\gamma_1, \ldots, l\gamma_n) \in \Sigma_{n+1} \tag{11.8}$$

для любого $l \in [0, I]$, где $\Sigma_{n+1} = \Sigma_{n+1}(\nu, \lambda)$ — множество (10.2) для мер $\nu$ и $\lambda$. Очевидно, (11.7) и (11.8) неверны, если $l > I$ при некотором выборе $u$.

Таким образом, величина $I$ совпадает с максимальным значением $l$, для которого справедливо (11.8). Геометрически $I$ можно найти следующим образом: из начала координат или из нулевого вектора $\mathbf{O}$ проводим луч через точку $\mathbf{Q} = (\gamma_0, \gamma_1, \ldots, \gamma_n)$ и находим точку $\mathbf{P}$ пересечения с границей $\Sigma_{n+1}$. Отношение $\mathbf{OP}/\mathbf{OQ}$ равно $I$. Это построение всегда возможно, так как из (11.5) следует, что нулевой вектор $\mathbf{O}$ является внутренней точкой $\Sigma_{n+1}$ (ср. теорему 2.2).

Перейдем к задаче характеризации минимизирующих многочленов. Рассмотрим опорную плоскость к $\Sigma_{n+1}$ в точке

$$(I\gamma_0, I\gamma_1, \ldots, I\gamma_n). \tag{11.9}$$

Таким образом, справедливо

$$\sum_{i=0}^{n} a_i c_i \leqslant d \tag{11.10}$$

для всех $\mathbf{c} = (c_0, c_1, \ldots, c_n) \in \Sigma_{n+1}$ и

$$I\sum_{i=0}^{n} a_i\gamma_i = d$$

для некоторого множества действительных чисел $a_0, a_1, \ldots\ a_n, d;\ \sum\limits_{i=0}^{n} a_i^2 > 0$.

Известно, что (11.9) представляется мерой $\sigma_0$. В частности,

$$I\gamma_i = \int_a^b u_i(t)\,d\sigma_0(t), \quad i = 0, 1, \ldots, n, \quad \lambda \leqslant \sigma_0 \leqslant \nu.$$

К тому же каждое $\mathbf{c}$, принадлежащее $\Sigma_{n+1}$, имеет представление

$$c_i = \int_a^b u_i(t)\,d\sigma(t), \quad i = 0, 1, \ldots, n, \quad \lambda \leqslant \sigma \leqslant \nu.$$

Соотношения (11.10) можно поэтому записать в виде

$$\int_a^b u^0(t)\,[d\sigma_0(t) - d\sigma(t)] \geqslant 0 \tag{11.11}$$

для любого $\sigma$ $(\lambda \leqslant \sigma \leqslant \nu)$, где $u^0 = \sum\limits_{i=0}^{n} a_i u_i(t)$.

Заметим, что (11.11) справедливо тодько для тех $\sigma_0$, для которых

$$d\sigma_0 = \begin{cases} d\nu, & \text{если } u^0 > 0, \\ d\lambda, & \text{если } u^0 < 0, \end{cases} \tag{11.12}$$

Более того, из (11.12) следует, что

$$l \sum_{i=0}^{n} a_0 \gamma_i = \int_a^b u^0(t)\, d\sigma_0(t) = \int_a^b u_+^0(t)\, d\nu(t) - \int_a^b u_-^0(t)\, d\lambda(t). \tag{11.13}$$

Это показывает, что многочлен $u^0(t) = \sum_{i=0}^{n} a_i u_i(t)$ является минимизирующим многочленом для функционала $\mathcal{J}$ Мы уже подчеркивали (см. следствие 2.2 (b)), что каждый многочлен порождает опорную плоскость к $\Sigma_{n+1}$. Мы утверждаем, что минимизирующий многочлен $u^0$ порождает опорную гиперплоскость в точке (11.9). Действительно, если (11 13) выполнено, то (11.12) тоже справедливо, и поэтому справедливы (11.11) и (11 10).

Можно также заметить, что узлы минимизирующего многочлена совпадают с узлами меры $\sigma_0$, которая представляет точку (11.9), и обратно. Фактически, если для многочлена $u^*$ справедливо (11.12), то также имеет место и (11.13) для $u^*(t) = \sum_{i=0}^{n} a_i^* u_i(t)$; последнее означает, что $u^*$ определяет опорную плоскость к $\Sigma_{n+1}$ в (11.9). Обратное утверждение очевидно.

Резюмируя сказанное, мы сформулируем следующую лемму.

Лемма 11.1. *Пусть $u^0 \in U$ — многочлен, для которого $\mathcal{J}(u)$ достигает минимума при ограничении* (11 2) *Если у $u^0$ имеется $n$ узловых корней на $(a, b)$, то он единствен. Если число узловых корней $u_0$ меньше $n$, то минимум также достигается для любого многочлена, нормированного условием* (11.2) *и имеющего те же самые узловые корни, что и $u^0$. Более того, $u^0$ — экстремальный многочлен тогда и только тогда, когда*

$$l\gamma_k = \int_a^b u_k(t) \frac{1 + \operatorname{sign} u^0(t)}{2}\, d\nu(t) + \int_a^b u_k(t) \frac{1 - \operatorname{sign} u^0(t)}{2}\, d\lambda(t)$$

*для некоторого $l > 0$ или, что эквивалентно,*

$$l\gamma_k = \int_a^b u_k(t)\, d\beta(t) + \int_a^b u_k(t) \operatorname{sign} u^0(t)\, d\alpha(t). \tag{11.14}$$

Нам потребуется следующая лемма.

Лемма 11.2. *Существуют две кусочно-постоянные функции $s(t)$, принимающие значения $+1$ или $-1$ и имеющие ровно $n+1$ точку разрыва на $(a, b)$, которые удовлетворяет условиям*

$$\int_a^b u_k(t)\, s(t)\, d\alpha(t) = \int_a^b u_k(t)\, d\beta(t), \qquad k = 0, 1, \ldots, n.$$

Доказательство Рассмотрим множество $\Sigma_{n+1}(\alpha, -\alpha)$, образованное согласно (10.1), где $\alpha$ играет роль $\nu$, а $-\alpha$ играет роль $\lambda$. Тогда в силу (11.1)

$-d\alpha < d\beta < d\alpha$ и, согласно теореме 2.2, точка с координатами

$$\int_a^b u_k(t)\,d\beta(t), \qquad k = 0, 1, \ldots, n, \tag{11.15}$$

является внутренней точкой $\Sigma_{n+1}(\alpha, -\alpha)$. Следовательно, она допускает два и только два главных представления (нижнее и верхнее). Пусть мера $\sigma$, удовлетворяющая условиям

$$\int_a^b u_k(t)\,d\beta(t) = \int_a^b u_k(t)\,d\sigma(t), \qquad k = 0, 1, \ldots, n, \tag{11.16}$$

соответствует одному из главных представлений в (11.15). Тогда $d\sigma(t) = s(t)\,d\alpha(t)$, где $s(t)$ удовлетворяет требуемым условиям.

Следовательно, если $s(t)$ удовлетворяет условиям леммы, то $\sigma$, определенная равенством $d\sigma(t) = s(t)\,d\alpha(t)$, является главным представлением моментной точки (11.15). Этим доказательство закончено.

Заметим, что $s_{\min}(b-) = -1$, $s_{\max}(b-) = +1$, где $s_{\min}(t)$ и $s_{\max}(t)$ соответствуют нижнему и верхнему главным представлениям (11.16).

Пусть $\Omega(t)$ — непрерывная функция такая, что каждая система функций $\{u_i\}_0^{n-1}$ и $\{u_0, u_1, \ldots, u_{n-1}, \Omega = u_n\}$ образует $T$-систему. Рассмотрим проблему минимизации (11.3) с $\gamma_0 = \gamma_1 = \ldots = \gamma_{n-1} = 0$, $\gamma_n = 1$. Это означает, что требуется найти многочлен $P(t) = \sum_{k=0}^{n-1} a_k u_k(t)$, на котором функционал

$$\mathcal{F}(\Omega - P) = \int_a^b |\Omega(t) - P(t)|\,d\alpha(t) + \int_a^b [\Omega(t) - P(t)]\,d\beta(t) \tag{11.17}$$

принимает минимальное значение.

Пусть $\theta_1, \theta_2, \ldots, \theta_n$ — точки разрыва функции $s_{\min}(t)$, соответствующей $T$-системе $\{u_k\}_0^{n-1}$ и заданным мерам $\alpha$ и $\beta$.

Теорема 11.1. *Минимальное значение* (11.17) *достигается при коэффициентах* $a_k$ $(k = 0, 1, \ldots, n-1)$, *определяемых из уравнений*

$$\sum_{k=0}^{n-1} a_k u_k(\theta_j) = \Omega(\theta_j); \qquad j = 1, \ldots, n. \tag{11.18}$$

Доказательство. Для любого многочлена $P(t) = \sum_{k=0}^{n-1} a_k u_k(t)$ справедливо соотношение

$$F(\Omega - P) \geqslant -\int_a^b \left[\Omega(t) - \sum_{k=0}^{n-1} a_k u_k(t)\right] s_{\min}(t)\,d\alpha(t) +$$
$$+ \int_a^b \left[\Omega(t) - \sum_{k=0}^{n-1} a_k u_k(t)\right] d\beta(t) = -\int_a^b \Omega(t)\,s_{\min}(t)\,d\alpha(t) + \int_a^b \Omega(t)\,d\beta(t).$$

Заметим, что последнее выражение в правой части не зависит от $P$. Знак равенств имеет место тогда и только тогда, когда

$$|\Omega(t) - P(t)| = -[\Omega(t) - P(t)]\cdot s_{\min}(t), \qquad a \leqslant t \leqslant b. \tag{11.19}$$

Для этого необходимо, чтобы $\Omega(t) - P(t)$ обращалось в нуль в точках, где $s_{\min}(t)$ изменяет свой знак, из чего следует (11.18).

С другой стороны, если справедливо (11.18), то также выполнено и (11.19), так как $(u_0, u_1, \ldots, u_{n-1}, \Omega)$ является $T$-системой. Более того, существуют коэффициенты $\{a_k\}_0^{n-1}$, являющиеся решением (11.18). Доказательство закончено.

Пример 11.2. Пусть $d\beta = 0$. Тогда точки разрыва $\theta_1, \ldots, \theta_n$ являются одновременно узлами для $s_{\max}(t)$ и $s_{\min}(t) = -s_{\min}(t)$ и однозначно определяются из уравнений

$$\int_a^{\theta_1} u_k(t)\,d\alpha(t) - \int_{\theta_1}^{\theta_2} u_k(t)\,d\alpha(t) + \ldots + (-1)^n \int_{\theta^n}^{b} u_k(t)\,d\alpha(t) = 0,$$

$$k = 0, 1, \ldots, n-1.$$

В частном случае $a = -1$, $b = 1$, $u_k(t) = t^k$ $(k = 0, 1, \ldots, n-1)$, $d\alpha(t) = dt$ мы имеем

$$\theta_j = -\cos\frac{j\pi}{n+1}, \qquad j = 1, \ldots, n.$$

Минимальным значением интеграла

$$\int_{-1}^{1}\left|\Omega(t) - \sum_{k=0}^{n-1} a_k t^k\right| dt$$

является

$$\int_{-1}^{1} \Omega(t)\,\operatorname{sign}\sin[(n+1)\arccos t]\,dt,$$

и оно достигается для многочлена $\sum\limits_{k=0}^{n-1} a_k t^k$ при условии, что справедливы равенства $\sum\limits_{k=0}^{n-1} a_k\theta_j^k = \Omega(\theta_j)$ $(j = 1, \ldots, n)$. Функция $\Omega(t)$, конечно, должна быть такой, чтобы $\{1, t, \ldots, t^{n-1}, \Omega(t)\}$ образовывали $T$-систему. В частности, если $\Omega(t) = t^n$, то

$$t^n - P(t) = \frac{\sin(n+1)\theta}{2^n \sin\theta}, \quad \text{где } \cos t = \theta.$$

## § 12. Теорема Ляпунова о множестве значений векторной меры

В предыдущих параграфах этой главы мы рассматривали моментные пространства $\Phi_{n+1}$, порожденные мерами, удовлетворяющими определенным ограничениям. В частности, пространство $\Phi_{n+1}$ состоит из всех векторов $\mathbf{c} = (c_0, \ldots, c_n)$ с компонентами вида

$c_i = \int\limits_a^b \varphi u_i d\mu$ $(i = 0, 1, \ldots, n)$, где $\{u_i\}_0^n$ — $T$-система, а функции $\varphi$ являются борелевскими измеримыми функциями, подчиняющимися ограничению $0 \leq \varphi \leq 1$. Было показано, что каждый элемент $\Phi_{n+1}$ порождается функцией $\varphi$, которая принимает только значения 0 и 1 таким образом, чтобы нулевые и единичные значения $\varphi$ чередовались самое большее, на $n+2$ последовательных интервалах. Если отбросить все ограничения на функции $u_i$, кроме того, что они неотрицательны, $\mu$-интегрируемы, и просто рассматривать конечные меры $d\mu_i = u_i d\mu$ на интервале $(a, b)$, то множество $\Phi_{n+1}$ можно интерпретировать как область значений некоторой векторной меры.

Важная теорема об области значений векторной меры принадлежит Ляпунову. Для того чтобы сформулировать эту теорему, нам потребуется ввести некоторую терминологию. Говорят, что конечная мера $\mu$, определенная на $\sigma$-поле $\mathcal{B}$ множеств абстрактного пространства $E$, является безатомной, если для любого измеримого множества $E_1$ положительной $\mu$-меры существует подмножество $E_0 \subset E_1$ такое, что $0 < \mu(E_0) < \mu(E_1)$. Теорема Ляпунова утверждает, что если $\mu_1, \ldots, \mu_n$ — конечные безатомные меры, то множество точек в $n$-мерном евклидовом пространстве, имеющих вид $(\mu_1(C), \ldots, \mu_n(C))$, при $C$, пробегающем все $\mathcal{B}$, является выпуклым и замкнутым. Следующая теорема является частным случаем этого результата.

Теорема 12.1. *Если $E$ — абстрактное пространство и $\mu_1, \ldots$ $\ldots, \mu_n$ — конечные безатомные меры, определенные на некотором $\sigma$-поле множеств из $E$, то каждая точка в множестве*

$$S(E) = \left\{ (c_1, \ldots, c_n) \mid c_i = \int\limits_E \varphi \, d\mu_i, \quad i = 1, \ldots, n, \quad 0 \leq \varphi \leq 1 \right\}$$

*может быть представлена в виде*

$$(c_1, \ldots, c_n) = (\mu_1(C), \ldots, \mu_n(C))$$

*для некоторого $C \in \mathcal{B}$.*

Замечание 12.1. Сформулированная выше теорема, естественно, обобщается на конечные безатомные обобщенные меры ограниченной вариации.

Доказательство. Сначала предположим, что $n = 1$. Мы хотим показать, что для каждого действительного $b \in (0, \mu_1(E))$ существует множество $C_0 \in \mathcal{B}$ такое, что $\mu_1(C_0) = b$. Чтобы доказать это, положим $\mathcal{C} = \{C \mid \mu_1(C) \leq b\}$ и определим частичную упорядоченность $\leq$ на множествах из $\mathcal{B}$ следующим образом: $A \leq B$ тогда и только тогда, когда существует нулевое множество $N$ такое, что $A \subset B \cup N$ ($N$ является нулевым множеством, если $\mu_1(N) = 0$).

Пусть $s = \sup_{\mathcal{C}'} \mu_1(C)$, где $\mathcal{C}'$ — любой линейно упорядоченный подкласс множества $\mathcal{C}$. Тогда существует последовательность $\{C_k\}_1^\infty$ множеств из $\mathcal{C}'$, которая возрастает относительно введенной частичной упорядоченности и для которой $\lim_{k\to\infty} \mu_1(C_k) = s$. В этом случае $\bigcup_k C_k \in \mathcal{C}$ и является максимальным элементом в $\mathcal{C}'$. В силу леммы Цорна класс $\mathcal{C}$ имеет максимальный элемент $C_0$, и мы утверждаем, что $\mu_1(C_0) = b$. Если бы $\mu_1(C_0) < b$, то мы смогли выбрать подмножество $A \subset E - C_0$, для которого $0 < \mu_1(A) < b - \mu_1(C_0)$, и, используя множество $C_0 \bigcup A$, прийти к противоречию с максимальностью $C_0$. Поэтому для $n = 1$ теорема справедлива.

Предположив, что теорема справедлива для $n - 1$, мы заметим, что, так же как и в теореме 2.1, множество $S(E)$ является выпуклым и компактным. Предположим, не умаляя общности, что $d\mu_i = f_i d\mu$, где $\mu$ — безатомная мера, так как мы могли бы положить $\mu = \sum_{i=1}^n \mu_i$. Пусть $\mathbf{c}^0 = (c_1^0, \ldots, c_n^0)$ — граничная точка $S(E)$.

Тогда существуют постоянные $a_0, a_1, \ldots, a_n$ такие, что $\sum_{i=1}^n a_i^2 > 0$,

$\sum_{i=1}^n a_i c_i + a_0 \geqslant 0$ и $\sum_{i=1}^n a_i c_i^0 + a_0 = 0$, так что $\sum_{i=1}^n a_i (c_i - c_i^0) \geqslant 0$ для всех $\mathbf{c} \in S(E)$ или, что эквивалентно,

$$\int_E \left( \sum_{i=1}^n a_i f_i \right) (\varphi - \varphi^0)\, d\mu \geqslant 0 \quad \text{для всех} \quad 0 \leq \varphi \leq 1.$$

Из этого неравенства мы выводим, что если $A = \left\{ t \mid t \in E,\ \sum_{i=1}^n a_i f_i(t) > 0 \right\}$ и $B = \left\{ t \mid t \in E,\ \sum_{i=1}^n a_i f_i(t) < 0 \right\}$, то $\varphi^0$ равно 0 и 1 на $A$ и $B$ соответственно почти наверное относительно $\mu$. Полагая $P = \left\{ t \mid t \in E,\ \sum_{i=1}^n a_i f_i(t) = 0 \right\}$, заметим, что если $\mu(P) = 0$, то $\varphi^0$ определена почти наверное и $\mathbf{c}^0$ имеет представление требуемого вида.

Предположим, что $\mu(P) > 0$. Будем считать для определенности, что $a_1 \neq 0$, так что на $P$ $f_1 = \sum_{i=2}^n a_i' f_i$. Пусть $\mathbf{d} = (d_2, \ldots, d_n)$ — точка с координатами $d_i = c_i^0 - \int_B f_i d\mu = \int_P \varphi^0 f_i d\mu, \quad i = 2, \ldots, n,$

содержащаяся в $S(P)$. Из предположения индукции следует, что $\int\limits_P \varphi^0 f_i d\mu = \int\limits_{P_1} f_i d\mu$, $i = 2, 3, \ldots, n$, для некоторого подмножества $P_1 \subset P$. Тогда, так как $f_i = \sum\limits_{l=2}^{n} a_l' f_l$ на $P$, то $\int\limits_P \varphi^0 f_1 d\mu = \int\limits_{P_1} f_1 d\mu$, так что

$$c_i^0 = \int\limits_{B \cup P_1} f_i d\mu, \qquad i = 1, 2, \ldots, n.$$

Доказательство будет полностью закончено, если мы покажем, что любая точка $(c_1^0, \ldots, c_n^0) \in \operatorname{Int} S(E)$ принадлежит границе $S(E_1)$ для некоторого подмножества $E_1 \subset E$. Предположив противное, положим $\mathcal{A} = \{A \mid \mathbf{c}^0 \notin S(A)\}$ и определим $\mathcal{A}'$ как произвольный линейно упорядоченный подкласс $\mathcal{A}$, где упорядоченность такая же, как и в случае $n = 1$. Тогда существует возрастающая подпоследовательность $\{A_k\}_0^\infty$ такая, что $\mu\left(\bigcup\limits_k A_k\right) = \sup\limits_{\mathcal{A}'} \mu(A')$, так что $\bigcup\limits_k A_k$ является максимальным элементом в $\mathcal{A}'$. Мы утверждаем, что либо $\mathbf{c}^0$ принадлежит границе $S\left(\bigcup\limits_k A_k\right)$, либо $\mathbf{c}^0 \notin S\left(\bigcup\limits_k A_k\right)$, т. е. $\bigcup\limits_k A_k \in \mathcal{A}$. Чтобы убедиться в этом, мы предположим, что

$$\mathbf{c}^0 \in \operatorname{Int} S\left(\bigcup_k A_k\right) = \operatorname{Int} S\left(\bigcup_k A_k'\right),$$

где $A_k'$ — непересекающиеся множества такие, что $A_m = \bigcup\limits_{l \leqslant m} A_l'$. Так как $\mu\left(\bigcup\limits_k A_k\right) < \infty$, то $\sum\limits_{k=m}^{\infty} \mu(A_k') \to 0$ при $m \to \infty$. Поэтому

$$\mathbf{c}^0 \in \operatorname{Int}\left[S\left(\bigcup_{k=1}^{m} A_k'\right) + S\left(\bigcup_{k=m+1}^{\infty} A_k'\right)\right]$$

и, следовательно, $\mathbf{c}^0 \in \operatorname{Int} S\left(\bigcup\limits_{k=1}^{m} A_k'\right) = \operatorname{Int} S(A_m)$ для достаточно больших $m$, что противоречит тому факту, что $A_m \in \mathcal{A}$. По предположению $\mathbf{c}^0 \notin \mathcal{D}S\left(\bigcup\limits_k A_k\right)$, следовательно, $\bigcup\limits_k A_k \in \mathcal{A}$. Поэтому каждое линейно упорядоченное подмножество $\mathcal{A}'$ обладает максимальным элементом в $\mathcal{A}$ и, следовательно, класс $\mathcal{A}$ имеет максимальный элемент $A_0$. Однако это невозможно, так как множество $S(A_0)$ компактно, а множество $E - A_0$ имеет подмножества произвольно малой положительной меры. Поэтому $\mathbf{c}^0 \in \mathcal{D}S(E_1)$ для некоторого подмножества $E_1 \subset E$. Доказательство закончено.

Теореме Ляпунова о множестве значений векторной меры уделено много внимания в различных работах. Другие доказательства теоремы Ляпунова см. у Халмоша [1948] и Блэкуэлла [1951]. Мы считаем, что доказательство, приведенное здесь, является новым и наиболее простым. Оно опирается на технику геометрии порожденных моментных пространств в духе данной книги. Теорема Ляпунова была обобщена и применена к различным статистическим задачам Блэкуэллом [1950]. Дворецким, Вальдом, Вольфовицем [1951] и другими. Она была использована Халкиным [1962] при установлении принципа максимума в оптимальном контроле. Дубинс и Спанье [1961] отмечали связь с топологией.

Замечание 12.2. Последняя часть доказательства, в которой доказывается, что каждая $c^0 \in S(E)$ содержится в границе $S(E_1)$ для некоторого подмножества $E_1 \subset E$, может быть выведена из следующего результата. Разобьем множество $\mathcal{B}$ на классы эквивалентности отношением $A \sim B$ тогда и только тогда, когда $\mu(A \Delta B)=0$ (где $A \Delta B$ обозначает симметрическую разность между $A$ и $B$); построим метрическое пространство $K$, состоящее из таких классов эквивалентности с метрикой $\rho(A, B) = \mu(A \Delta B)$; тогда пространство $K$ будет связным. Тот факт, что если $c^0$ содержится во внутренней части $S(E)$, то она принадлежит границе $S(E_1)$ для некоторого $E_1 \subset E$, теперь следует немедленно. Действительно, если предположить противное, то пространство $K$ можно было бы разбить на два непересекающихся открытых непустых множества $\mathcal{A} = \{A/c^0 \notin S(A)\}$ и $\mathcal{C} = \{A/c_0 \in \operatorname{Int} S(A)\}$, что противоречило бы тому, что $K$ связно. Приводимое ниже доказательство того, что $K$ линейно связно и следовательно, связно, сообщил нам К. Ито. Это доказательство мы разобъем на две предварительные леммы, первая из которых является утверждением теоремы Ляпунова при $n = 1$ (ср. Дубинс и Спанье [1961]).

Лемма 12.1. *Если $\mu$—безатомная мера, определенная на $\sigma$-поле множеств $\mathcal{B}$ абстрактного пространства $E$, и $A \in \mathcal{B}$ таково, что $\mu(A) > \alpha > 0$, то существует подмножество $A_1 \subset A$ такое, что $\mu(A_1) = \alpha$.*

Лемма 12.2. *Для каждого $A \in \mathcal{B}$ существует набор множеств $A_t$, $t \in [0, 1]$, такой, что:*

(i) $A_t \subset A_s$ для $t < s$ $(A_1 = A)$,

(ii) $\mu(A_t) = t_\mu(A)$.

Доказательство. Если $\mu(A)=0$, то лемма справедлива. Если $\mu(A>0)$, то применим лемму 12.1 и определим $A_{1/2}$ как подмножество $A = A_1$ такое, что $\mu(A_{1/2}) = \frac{1}{2}\mu(A)$. Затем определим $A_{1/2}$ как подмножество $A_{1/2}$ такое, что $\mu(A_{1/4}) = \frac{1}{2}\mu(A_{1/2})$ и $A_{3/4} = A_{1/2} \cup A'_{1/4}$, где $A'_{1/2} \subset A - A_{1/2}$ и $\mu(A'_{1/4}) = \frac{1}{2}\mu(A - A_{1/2})$. Продолжая этот процесс, мы можем определить $A_{k/2^n}$ так, чтобы $\mu(A_{k/2^n}) = \mu(A) \cdot k/2^n$. Для произвольного $t$ $A_t$ определяется соотношением

$$A_t = \bigcup_{k/2^n \leqslant t} A_{k/2^n}.$$

Легко проверить, что множества $A_t$ удовлетворяют лемме,

Теорема 12.2 *Метрическое пространство $K$ связно.*

Доказательство. Для любых двух заданных множеств $A$ и $B$ рассмотрим множества $A_t$ и $B_t$, определенные в лемме 12.2 и рассмотрим множества $C_t$

для $t \in [0, 1]$

$$C_t = \begin{cases} A_{-2t+1}, & t \in [0, 1/2], \\ B_{2t-1}, & t \in [1/2, 1]. \end{cases}$$

Множества $C$ соединяют $A$ и $B$ непрерывной линией. Поэтому пространство $K$ линейно связно и, следовательно, связно.

## § 13. Обобщения теоремы Ляпунова

В этом параграфе мы рассмотрим некоторые обобщения теоремы Ляпунова, приведенные у Карлина [1953]. Эти результаты мы применим в § 14 для того, чтобы решить различные задачи, такие, как «Задача о Ниле», «Задача о разрезании пирога», «Задача подобных областей».

В данном тексте приведены некоторые уточнения результатов Карлина [1953].

Теперь мы предположим, что $E$—абстрактное пространство, $\mathcal{B}$—$\sigma$-поле множеств из $E$, $\mu$—конечная безатомная мера, определенная на $\mathcal{B}$ и $A$ — ограниченное замкнутое выпуклое множество в $m$-мерном евклидовом пространстве. Пусть $L = L(\mu, \mathcal{B})$ обозначает пространство всех функций, интегрируемых относительно $\mu$, а $M = M(\mu, \mathcal{B})$—пространство всех существенно ограниченных измеримых функций, определенных на $E$. Если $L_m$ обозначает прямое произведение $L$ на себя, взятое $m$ раз, а $M_m$ — прямое произведение $M$ на себя, взятое $m$ раз, то при соответствующем выборе нормы $M_m$ является пространством, сопряженным к $L_m$.

Определим $M_A$ как множество всех функций из $M_m$ со значениями в $A$. Таким образом, элементами $M_A$ являются вектор-функции $\mathbf{x}(t) = (x_1(t), \ldots, x_m(t))$, каждая компонента которых $x_i(t) \in M_1$ и $\mathbf{x}(t) \in A$ почти наверное. Так как $A$ выпукло, то просто проверить, что $M_A$ выпукло, ограничено и слабо замкнуто.

Так как единичная сфера в $M_m$ компактна в слабой топологии, то мы получаем, что $M_A$ компактно. Из компактности и выпуклости, используя теорему Крейна—Мильмана, получаем, что $M_A$ порождается крайними точками. Существование крайних точек доказывается тривиально; любая функция в $M_A$, единственное значение которой является крайней точкой, в свою очередь является крайней точкой $M_A$.

Доказательства приведенных ниже теорем 13.2 и 13.3 опираются на следующую интуитивно очевидную теорему. (Трудности возникают при доказательстве измеримости построенной в этой теореме функции; для простоты изложения доказательство этой теоремы мы поместим после следствия 13.4.)

Теорема 13.1. *Пусть $A$ — компактное выпуклое множество в $m$-мерном евклидовом пространстве $R^m$. Тогда существует борелевская измеримая функция $\varphi$: $A \to R^m$ такая, что*:

(i) $\mathbf{x} \pm \varphi(x) \in A$ *для* $\mathbf{x} \in A$,

(ii) $\varphi(\mathbf{x}) \neq 0$, *если $x$ не является крайней точкой $A$.*

С помощью этой теоремы легко доказывается

Теорема 13.2. *Множество крайних точек $M_A$ в точности совпадает с $M_B$, где $B$ — множество крайних точек $A$.*

Доказательство. Пусть $\mathbf{x}_0(t)$ — крайняя точка $M_A$ и предположим противное, т. е. что $\mathbf{x}_0 \notin M_B$. Отсюда следует, что $x_0(t) \notin B$ на множестве положительной меры. Однако по теореме 13.1

$$\mathbf{x}_0(t) = \frac{\mathbf{x}_0(t) + \varphi(\mathbf{x}_0(t))}{2} + \frac{\mathbf{x}_0(t) - \varphi(\mathbf{x}_0(t))}{2},$$

но это противоречит тому, что $\mathbf{x}_0$ — крайняя точка $M_A$. Легко проверить, что $M_B$ состоит только из крайних точек $M_A$, что завершает доказательство.

Теперь пусть $M(A; \mu_1, \ldots, \mu_n) = \{\xi_i \mid i = 1, \ldots, n,$ где $\xi_i = \int \mathbf{x}\, d\mu_i$, $\mathbf{x} \in M_A$ и $\mu_j$, $j = 1, \ldots, n$, — меры, определенные на $\mathcal{B}\}$.

Теорема 13.3. *Пусть $\mu_j$ $(j = 1, \ldots, n)$ — конечная безатомная мера, и пусть $T$ обозначает множество всех $\mathbf{x} \in M_A$, удовлетворяющих условию $\mathbf{a}_i \leq \int\limits_E \mathbf{x}\, d\mu_i \leq \mathbf{b}_i$, где $\mathbf{a}_i$ и $\mathbf{b}_i$ $(i = 1, \ldots, n)$ — векторы из $R^m$. Тогда крайние точки $T$ содержатся в $M_B$. (Векторное неравенство означает, что неравенство справедливо для каждой компоненты).*

Замечание 13.1. Если $\mu = \mu_1 + \ldots + \mu_n$, то $\mu_i(U) = \int\limits_U f_i d\mu$ для всех $U \in \mathcal{B}$. Утверждение теоремы в терминах $f_i$ состоит в том, что крайние точки $T$ множества всех $\mathbf{x}(t)$ из $M_A$, для которых справедливо неравенство

$$\mathbf{a}_i \leq \int\limits_E \mathbf{x} f_i d\mu \leq \mathbf{b}_i,$$

содержатся в $M_B$, если $\mu$ — безатомная мера. Так как $A$ выпукло и замкнуто и линейные ограничения получены с помощью элементов из $L(\mu, \mathcal{B})$, то мы получаем, что $T$ компактно и выпукло и, следовательно, имеет крайние точки.

Доказательство теоремы 13.3. Пусть $\mathbf{x}_0$ — крайняя точка $T$, и предположим, что $\mathbf{x}_0 \notin M_B$. Если применить теорему Ляпунова (теорема 12.1) к $mn$ знакопеременным мерам, соответствующим координатам $\varphi(\mathbf{x}_0(t)) f_i(t) d\mu(t)$, $i = 1, \ldots, n$, то получим измеримое подмножество $E_0 \subset E$, удовлетворяющее условию

$$\int\limits_{E_0} \varphi(\mathbf{x}_0(t)) f_i(t) d\mu(t) = \frac{1}{2} \int\limits_E \varphi(\mathbf{x}_0(t)) f_i(t) d\mu(t) \tag{13.1}$$

для $i = 1, \ldots, n$. Определим теперь $\beta(\mathbf{x}_0(t)) = \varphi(\mathbf{x}_0(t))$ для $t \in E_0$, $\beta(\mathbf{x}_0(t)) = -\varphi(\mathbf{x}_0(t))$ для $t \notin E_0$ и $\psi(\mathbf{x}_0(t)) = -\beta(\mathbf{x}_0(t))$. Очевидно, что

$$\mathbf{x}_0(t) = \frac{\mathbf{x}_0(t) + \beta(\mathbf{x}_0(t))}{2} + \frac{\mathbf{x}_0(t) + \psi(\mathbf{x}_0(t))}{2},$$

а $x_0(t)+\beta(x_0(t))$ и $x_0(t)+\psi(x_0(t))$ принадлежат $M_A$. Простые вычисления с использованием (13.1) показывают, что

$$\int_E \beta(x_0(t)) f_i(t)\, d\mu(t) = \int_E \psi(x_0(t)) f_i(t)\, d\mu(t) = 0,$$

так что $x_0(t)+\psi(x_0(t))$ и $x_0(t)+\beta(x_0(t))$ удовлетворяют ограничениям, сформулированным в теореме, и поэтому принадлежат $T$.

Мы представили $x_0$ в виде выпуклой комбинации двух различных элементов из $T$, но это противоречит тому, что $x_0$ — крайняя точка, поэтому доказательство теоремы закончено.

Теорема 13.4. *Пусть $\mu_j$ ($j=1,\ldots,n$) — безатомные конечные меры. Тогда $M(A;\ \mu_1,\ldots,\mu_n)=M(B;\ \mu_1,\ldots,\mu_n)$.*

Замечание 13.2. Непосредственным следствием этой теоремы является тот факт, что $M(B;\ \mu_1,\ldots,\mu_n)$ выпукло и замкнуто, так как $M(A;\ \mu_1,\ldots,\mu_n)$ обладает этими свойствами.

Доказательство. Пусть $x_0$ — произвольная точка из $M_A$. Положим $\xi_i=\int x_0 d\mu_i$. Пусть $\Gamma$ — множество всех $x$ из $M_A$, для которых $\xi_i=\int x\, d\mu_i$, $i=1, 2, \ldots, n$. Это множество слабо замкнуто, выпукло и, следовательно, компактно. (См. замечание 13.1). Следовательно, существует крайняя точка $x_1$ в $\Gamma$. В силу утверждения теоремы 13.3 $x_1$ принадлежит $M_B$ и $\int x_1 d\mu_i=\xi_i=\int x_0 d\mu_i$, $i=1,2,\ldots$ $\ldots, n$. Таким образом, произвольная точка, взятая из $M(A;\ \mu_1,\ldots$ $\ldots,\mu_n)$, принадлежит $M(B;\ \mu_1,\ldots,\mu_n)$. Так как обратное включение очевидно, то теорема доказана.

Следствие 13.4. *Пусть $C$ — произвольное, замкнутое, ограниченное множество в $R^m$. Тогда, если $\mu_j$ ($j=1,\ldots,n$) — безатомные меры, то $M(C;\ \mu_1,\ldots$ $\ldots,\mu_n)$ выпукло и замкнуто.*

Доказательство. Пусть $A$ — выпуклая оболочка $C$. Очевидно, что $A$ замкнуто, и для множества $B$ его крайних точек справедливо включение $B\subset C$. Поэтому

$$M(B;\ \mu_1,\ldots,\mu_n)\subset M(C;\ \mu_1,\ldots,\mu_n)\subset M(A;\ \mu_1,\ldots\ \mu_n).$$

В силу теоремы 13.4 первое и последнее из этих множеств совпадают, откуда следует, что $M(C;\mu_1,\ldots,\mu_n)$ замкнуто и выпукло

Доказательство теоремы 13.1. (Приводимое доказательство принадлежит К. Ито.) Не умаляя общности, мы предположим, что множество $A$ $m$-мерно. Утверждение теоремы эквивалентно существованию борелевской измеримой функции $f: A\to A$ такой, что $f(x)$ и $2x-f(x)$ принадлежат $A$ и $f(x)\neq x$, если $x$ не является крайней точкой $A$.

Для каждого $x\in A$ положим

$$A(x)=\{y\mid y\in A, 2x-y\in A\}.$$

Легко видеть, что это множество компактно, выпукло, симметрично относительно $x$ и сводится к единственной точке только в том случае, если $x$ — крайняя точка $A$. Пусть $c^0, c^1,\ldots,c^m$ — $m+1$ точек в $R^m$, не содержащихся в $A$ и не лежащих ни на какой $(m-1)$-мерной гиперплоскости. Определим $f_i(x)\in A(x)$ как точку в $A(x)$, лежащую ближе всего к точке $c_i$. Такая точка существует и однозначно определена; существование следует из того, что расстояние $\|y-c^i\|$ — непрерывная функция точки $y\in A(x)$, а множество $A(x)$ компактно. Единственность следует из выпуклости $A(x)$.

Теперь мы покажем, что если $A(\mathbf{x}) \neq \{\mathbf{x}\}$, то для некоторого $i$ $f_i(\mathbf{x}) \neq \mathbf{x}$. Действительно, $f_i(\mathbf{x}) \neq \mathbf{x}$ для любого $i$, если $\mathbf{x}$ — внутренняя точка $A$. Если $\mathbf{x}$ лежит на границе $A$ и $f_0(\mathbf{x}) = \ldots = f_m(\mathbf{x}) = \mathbf{x}$, то каждая из $\mathbf{c}^i$ принадлежит гиперплоскости $H'$, проходящей через $\mathbf{x}$, и являющейся дополнительной к гиперплоскости наименьшей размерности, содержащей $A(\mathbf{x})$. Но если $A(\mathbf{x}) \neq \{\mathbf{x}\}$, то $\dim H' \leqslant m - 1$, так что $m + 1$ независимых точек $c^0, \ldots, c^m$ не могут лежать в $H'$. Поэтому $f_i(\mathbf{x}) \neq \mathbf{x}$ для некоторого $i$.

Функцию $f(\mathbf{x})$ определим равной $\mathbf{x}$, если $\mathbf{x}$ — крайняя точка, а в противном случае с помощью формулы

$$f(\mathbf{x}) = \begin{cases} f_0(\mathbf{x}), & \text{если } f_0(\mathbf{x}) \neq \mathbf{x}, \\ f_1(\mathbf{x}), & \text{если } f_0(\mathbf{x}) = \mathbf{x}, \quad f_1(\mathbf{x}) \neq \mathbf{x}, \\ \cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots \\ f_m(\mathbf{x}), & \text{если } f_0(\mathbf{x}) = \ldots = f_{m-1}(\mathbf{x}) = \mathbf{x}, \quad f_m(\mathbf{x}) \neq \mathbf{x}. \end{cases}$$

Очевидно, что $f(\mathbf{x})$ будет удовлетворять всем требованиям, если мы докажем что она является борелевской измеримой функцией. Чтобы установить это, обозначим через $U_\varepsilon(\mathbf{x})$ объединение множеств $A(\mathbf{x}')$ по всем $\mathbf{x}' \in A$, для которых $\|\mathbf{x} - \mathbf{x}'\| \leqslant \varepsilon$. Множество $U_\varepsilon(\mathbf{x})$ $m$-мерно, выпукло и компактно, так как оно является проекцией на последние $m$ координат $2m$-мерного выпуклого, компактного множества

$$K(\varepsilon, \mathbf{x}) = \widetilde{A} \cap [\{\mathbf{x}' \mid \|\mathbf{x}' - \mathbf{x}\| \leqslant \varepsilon\} \times R^m],$$

где $\widetilde{A} = \{(\mathbf{x}, y) \mid \mathbf{x} \in A, y \in A(\mathbf{x})\}$.

Множество $K(\varepsilon, \mathbf{x})$ выпукло и компактно, потому что, очевидно, $\widetilde{A}$ обладает этими свойствами. Определим теперь $m$-мерную точку

$$\mathbf{M}(U_\varepsilon(\mathbf{x}), \mathbf{c}^i, k) = \frac{\int\limits_{U_\varepsilon(\mathbf{x})} \mathbf{y} \cdot \|\mathbf{y} - \mathbf{c}^i\|^{-k} \mu(d\mathbf{y})}{\int\limits_{U_\varepsilon(\mathbf{x})} \|\mathbf{y} - \mathbf{c}^i\|^{-k} \mu(d\mathbf{y})},$$

где $\mu(d\mathbf{y})$ — $m$-мерная лебегова мера. Ближайшая к $\mathbf{c}^i$ точка $\mathbf{N}(U_\varepsilon(\mathbf{x}, \mathbf{c}^i))$ в $U_\varepsilon(\mathbf{x})$ однозначно определена, так как $U_\varepsilon(\mathbf{x})$ компактно и выпукло. Заметим, что

$$\lim_{k \to \infty} \mathbf{M}(U_\varepsilon(\mathbf{x}), \mathbf{c}^i, k) = \mathbf{N}(U_\varepsilon(\mathbf{x}), \mathbf{c}^i),$$

так как мера

$$\frac{\|\mathbf{y} - \mathbf{c}^i\|^{-k} \mu(d\mathbf{y})}{\int\limits_{U_\varepsilon(\mathbf{x})} \|\mathbf{y} - \mathbf{c}^i\|^{-k} \mu(d\mathbf{y})}$$

сходится к $\delta$-мере в точке $\mathbf{N}(U_\varepsilon(\mathbf{x}), \mathbf{c}^i)$.

Функции $\mathbf{M}(U_\varepsilon(\mathbf{x}), \mathbf{c}^i, k)$ непрерывны по $\mathbf{x}$ при фиксированных $\varepsilon$ и $k$. Для того чтобы доказать это, мы сначала покажем, что $\mathbf{M}(U_\varepsilon(\mathbf{x}), \mathbf{c}^i, k)$ непрерывны по $\varepsilon$ при фиксированных $\mathbf{x}$ и $k$. Так как $K(\varepsilon, \mathbf{x})\ .\ K(\varepsilon_0, \mathbf{x})$ при $\varepsilon \downarrow \varepsilon_0$ и $\operatorname{Int} K(\varepsilon, \mathbf{x}) \uparrow \operatorname{Int} K(\varepsilon_0, \mathbf{x})$ при $\varepsilon \uparrow \varepsilon_0$, то $U_\varepsilon(\mathbf{x}) \downarrow U_{\varepsilon_0}(\mathbf{x})$ при $\varepsilon \downarrow \varepsilon_0$ и

$\operatorname{Int} U_\varepsilon(\mathbf{x}) \uparrow \operatorname{Int} U_{\varepsilon_0}(x)$ при $\varepsilon \uparrow \varepsilon_0$. Так как граница $U_\varepsilon(\mathbf{x})$ имеет нулевую меру, то

$$\int\limits_{U_\varepsilon(\mathbf{x})} g(\mathbf{y})\,\mu(d\mathbf{y}) \to \int\limits_{U_{\varepsilon_0}(x)} g(\mathbf{y})\,\mu(d\mathbf{y}),$$

если $\varepsilon \downarrow 0$ или $\varepsilon \uparrow \varepsilon_0$, а $g(\mathbf{y})$ — ограниченная неотрицательная функция, определенная на достаточно широкой области. Далее, так как $U_{\varepsilon+\delta}(\mathbf{x}) \supset U_\varepsilon(\mathbf{z}) \supset \supset U_{\varepsilon-\delta}(\mathbf{x})$, $\|\mathbf{x}-\mathbf{z}\| < \delta$, то

$$\int\limits_{U_\varepsilon(\mathbf{x})} g(\mathbf{y})\,\mu(d\mathbf{y}) = \lim_{\delta\to 0} \int\limits_{U_{\varepsilon+\delta}(\mathbf{x})} g(\mathbf{y})\,\mu(d\mathbf{y}) \geqslant \overline{\lim_{\mathbf{z}\to\mathbf{x}}} \int\limits_{U_\varepsilon(\mathbf{z})} g(\mathbf{y})\,\mu(d\mathbf{y}) \geqslant$$

$$\geqslant \lim_{\mathbf{z}\to\mathbf{x}} \int\limits_{U_\varepsilon(\mathbf{z})} g(\mathbf{y})\,\mu(d\mathbf{y}) \geqslant \lim_{\delta\to 0} \int\limits_{U_{\varepsilon-\delta}(\mathbf{x})} g(\mathbf{y})\,\mu(d\mathbf{y}) = \int\limits_{U_\varepsilon(\mathbf{x})} g(\mathbf{y})\,\mu(d\mathbf{y}).$$

Отсюда следует непрерывность $\mathbf{M}(U_\varepsilon(\mathbf{x}), \mathbf{c}^i, k)$ по $\mathbf{x}$, и можно сделать вывод что $\mathbf{N}(U_\varepsilon(\mathbf{x}), \mathbf{c}^i)$ — борелевская измеримая по $\mathbf{x}$ функция, так как

$$\lim_{k\to\infty} \mathbf{M}(U_\varepsilon(\mathbf{x}, \mathbf{c}^i, k) = \mathbf{N}(U_\varepsilon(\mathbf{x}), \mathbf{c}^i).$$

Для завершения доказательства теперь достаточно показать, что

$$\lim_{\varepsilon\to 0} \mathbf{N}(U_\varepsilon(\mathbf{x}), \mathbf{c}^i) = f_i(\mathbf{x}) = \mathbf{N}(A(\mathbf{x}), \mathbf{c}^i).$$

Положим $\mathbf{p}^k = \mathbf{N}(U_{\varepsilon_k}(\mathbf{x}), \mathbf{c}^i)$, где $\varepsilon_k \to 0$ при $k \to \infty$. Тогда, так как $\mathbf{p}^k \in U_{\varepsilon_1}(\mathbf{x})$ для всех $k$ и $U_{\varepsilon_1}(\mathbf{x})$ компактно, то существует подпоследовательность $\mathbf{p}^{k'}$, сходящаяся к некоторой точке $\mathbf{p}$. Точка $\mathbf{p}$ содержится в $A(\mathbf{x})$, так как $\mathbf{p} \in U_{\varepsilon_{k'}}(\mathbf{x})$ для каждого $k'$ и $\bigcap\limits_{k'} U_{\varepsilon_{k'}}(\mathbf{x}) = A(\mathbf{x})$. Более того,

$$\|\mathbf{c}^i - \mathbf{p}\| \geqslant \|\mathbf{c}^i - \mathbf{N}(A(\mathbf{x}), \mathbf{c}^i)\|,$$

$$\|\mathbf{c}^i - \mathbf{p}\| \leqslant \lim_{k'\to\infty} \|\mathbf{c}^i - \mathbf{p}^{k'}\| \leqslant \|\mathbf{c}^i - \mathbf{N}(A(\mathbf{x}), \mathbf{c}^i)\|,$$

так что $\mathbf{p} = \mathbf{N}(A(\mathbf{x}), \mathbf{c}^i)$. Поэтому

$$\lim_{\varepsilon\to 0} \mathbf{N}(U_\varepsilon(\mathbf{x}), \mathbf{c}^i) = f_i(\mathbf{x}).$$

Доказательство теоремы закончено.

## § 14. Применения теоремы Ляпунова

Результаты теоремы 13.4 можно применить к трем задачам, упомянутым в начале § 13. «Задача о разрезании пирога» формулируется следующим образом. Задан некоторый предмет, например пирог, который требуется разделить на $m$ человек таким образом, чтобы каждый человек считал, что он получил справедливую долю или, по его собственной оценке, по крайней мере $1/m$ часть пирога. Решение этой задачи, которую мы сформулируем ниже математически, состоит в том, что нож медленно разрезает пирог до тех пор, пока один из $m$ человек не заявит, что кусок, определенный таким образом, составляет не меньше чем $1/m$ часть целого пирога. Ему дают этот кусок, а каждый из оставшихся $m - 1$ человек считает, что оставшаяся часть составляет не менее $(m-1)/m$ от всего пирога, так что процесс можно продолжать до тех пор, пока каждый не получит свою порцию.

Если имеется $m$ рациональных чисел $\alpha_i$, $i = 1, \ldots, m$, $\sum_{i=1}^{m} \alpha_i = 1$, то описанный выше метод можно применить для того, чтобы устроить разбиение таким образом, чтобы $i$-й человек считал, что он получил по крайней мере $\alpha_i$-ю часть пирога. Чтобы пояснить это, положим $\alpha_i = p_i M^{-1}$, где $p_i$ и $M$ — целые. Предположим теперь, что у нас $M$ человек, и пусть $i$-й человек выбирает $p_i$ раз. Проводя процесс разрезания пирога, как это описано выше, мы получим, что $i$-й человек будет считать, что он получил $p_i$ кусков, каждый из которых составляет не менее $1/M$ часть всего пирога.

Можно было бы также задаться вопросом, возможно ли разрезать пирог на $m$ кусков так, чтобы каждый человек считал, что каждый кусок составляет $1/m$ часть целого пирога и все были бы довольны тем, что каждый из них получил справедливую долю. Это была бы немного более бескорыстная группа людей, чем группа, рассмотренная в первоначальной формулировке, приведенной выше.

Еще один вариант этой задачи состоит в том, чтобы делить пирог таким образом, чтобы каждый из $m$ лиц считал, что он на самом деле получил более чем $1/m$ часть пирога.

Можно показать, что эта задача имеет решение при условии, что по крайней мере двое из участников различаются в своей оценке некоторой части пирога.

Следует заметить, что во втором варианте задачи о разрезании пирога (случай бескорыстной группы лиц) имеется $m$ мер, определенных на некоторых подмножествах данного множества, и мы оцениваем $m$ мер на $m$ множествах в данном разбиении. Если мы изменим число мер и оставим неизменным число множеств в разбиении, то мы придем к тому, что называется задачей о Ниле. Эта задача описывается следующим образом. Каждый год на реке Нил бывает наводнение. При каждом уровне наводнения и при каждом разбиении на фиксированное число участков земли, окружающей Нил, каждый участок оценивается определенным образом по своей пригодности для сельского хозяйства и для других целей. Каждый уровень наводнения соответствует цене на окружающей Нил земле и можно задаться вопросом, возможно ли разделить землю на $m$

участков таким образом, чтобы стоимость каждого участка составляла $1/m$ от общей стоимости независимо от уровня наводнения. Мы покажем, что для любого конечного числа уровней наводнения эта задача имеет решение. Для бесконечного числа уровней нет необходимости решать эту задачу.

Частным случаем этой задачи является «задача подобных областей», в которой, вместо разбиения множества, мы хотим найти для каждого $\alpha$ между нулем и единицей единственное подмножество $S$ такое, что для каждой из $n$ мер доля меры $S$ к общей мере равнялась бы $\alpha$. Случай $\alpha = 1/2$ называется «задачей бисекции».

Решение приведенных выше задач опирается на частный случай теоремы 13.4. Пусть $A$ — $(m-1)$-мерный симплекс точек $(y_1, \dots, y_m)$ таких, что $y_i \geqslant 0$, $i = 1, \dots, m$, и $\sum_{i=1}^{m} y_i = 1$. Крайние точки этого множества являются $m$-мерными векторами с $i$-й компонентой, равной единице, и с нулевыми остальными компонентами. Легко видеть, что множество $M(B; \mu_1, \dots \mu_n)$ в теореме 13.4 можно отождествить с множеством $(n \times m)$-матриц $R$, определенным следующим образом. Рассмотрим разбиение $(A_1, \dots, A_m)$ множества $E$, где $A_i \in \mathcal{B}$, $i = 1, \dots, m$, и пусть $j$-я строка типичной матрицы в $k$ равняется $(\mu_j(A_1), \dots, \mu_j(A_m))$. Множество $R$ получается, если рассматривать такие матрицы при всех разбиениях пространства $E$. Из теоремы 13.4 мы непосредственно получаем следующую.

Теорема 14.1. *Если $\mu_1, \dots, \mu_n$ — конечные безатомные меры, а множества $A$ и $R$ определены, как указано выше, то множество матриц $R$ компактно и выпукло, причем $R = M(A; \mu_1, \dots, \mu_n)$.*

Следствие 14.1. *Если $\mu_1, \dots, \mu_n$ — базатомные вероятностные меры, то для каждого набора из $m$ неотрицательных чисел $\alpha_i$, $\sum_{i=1}^{m} \alpha_i = 1$, существует разбиение $A_1, \dots, A_m$ пространства $E$ такое, что $\mu_i(A_j) = \alpha_j$ при всех $i = 1, \dots, n$ и $j = 1, \dots, m$.*

Доказательство. Матрица, каждая строка которой равна $(\alpha_1, \dots, \alpha_m)$, принадлежит $M(A; \mu_1, \dots, \mu_n)$. Следствие вытекает из теоремы 14.1.

Непосредственно выбирая меры из следствия 14.1, можно показать, что каждая из обсуждавшихся выше задач имеет решение, за исключением задачи о разделении пирога так, чтобы каждый считал, что он получил более чем справедливую часть. Решение этой задачи содержится в следующей лемме.

Лемма 14.1. *Если $\mu_1, \dots, \mu_n$ — безатомные вероятностные меры и $\mu_i \neq \mu_j$ для некоторых $i \neq j$, то для любых $m$ заданных действительных чисел $\alpha_i > 0$, $i = 1, 2, \dots, m$, $\sum_{i=1}^{m} \alpha_i = 1$, существует разбиение $A_1, \dots, A_m$ пространства $E$ такое, что $\mu_i(A_i) > \alpha_i$ для $i = 1, \dots, m$.*

Доказательство. В силу теоремы 14.1 для доказательства леммы достаточно показать, что существует функция $\beta = (\beta_1, \ldots, \beta_m)$, определенная на $E$, такая, что $\beta_i(t) \geqslant 0$, $\sum_{i=1}^{m} \beta_i(t) = 1$ и $\int_E \beta_i d\mu_i > \alpha_i$ для $i = 1, \ldots, m$ или $\int_E (\beta_i - \alpha_i)\, d\mu_i > 0$, $i = 1, \ldots, m$.

Мы будем записывать меры в виде $d\mu_i = f_i d\mu$. Предположим противное, т. е. что $\beta$ с вышеперечисленными свойствами не существует. Мы рассмотрим множество $S$ всех векторов $(b_1, \ldots, b_m)$, где $b_i = \int_E (\beta_i - \alpha_i) f_i d\mu$, полученное при изменении вектор-функции $\beta$. Это множество замкнуто, выпукло и содержит точку $(0, 0, \ldots, 0)$. (Достаточно выбрать $\beta_i(t) \equiv \alpha_i$). Более того, $S$ не пересекает положительный ортант. Следовательно, существуют постоянные $\xi_i$, $i = 1, \ldots, m$, не все из которых равны нулю и такие, что

$$\sum_{i=1}^{m} \xi_i \int_E (\beta_i - \alpha_i) f_i\, d\mu \leqslant 0$$

при всех $\beta = (\beta_1, \ldots, \beta_m)$.

Мы предположим, что $\xi_m \neq 0$. Учитывая, что $\sum_{i=1}^{m} \beta_i = 1$ и $\sum_{i=1}^{m} \alpha_i = 1$, получим, что последнее неравенство можно переписать в виде

$$\int_E \sum_{i=1}^{m-1} (\xi_i f_i - \xi_m f_m)(\beta_i - \alpha_i)\, d\mu \leqslant 0. \qquad (14.1)$$

Мы утверждаем, что из (14.1) следует, что $\xi_i f_i - \xi_m f_m = 0$ почти наверное относительно $\mu$. Если бы это было не так, то мы бы положили

$$A_i = \{t \mid \xi_i f_i - \xi_m f_m = 0\},$$

$$B_i = \{t \mid \xi_i f_i - \xi_m f_m < 0\},$$

$$C_i = \{t \mid \xi_i f_i - \xi_m f_m > 0\}$$

и определили бы

$$\beta_i^0(t) = \begin{cases} 0, & t \in A_i, \\ \alpha_i - \alpha/(m-1), & t \in B_i, \qquad i = 1, \ldots m-1, \\ \alpha_i + \beta/(m-1), & t \in C_i, \end{cases}$$

где $\alpha = \min\{\alpha_1, \dots, \alpha_{m-1}\}$, а $\beta = 1 - \sum_{i=1}^{m-1} \alpha_i$ и

$$\beta_m^0(t) = \begin{cases} 1, & t \in A_i, \\ \beta + \alpha, & t \in B_i, \\ 0, & t \in C_i. \end{cases}$$

При этом выборе $\beta^0$ нарушилось бы неравенство (14.1), за исключением того случая, когда $\xi_i f_i = \xi_m f_m$ почти наверное для $i = 1, \dots, m-1$. Но так как $\int f_i d\mu = 1$ для каждого $i$, то отсюда следует, что $\xi_i = \xi_m$ для $i = 1, \dots, m-1$. В этом случае все меры $\mu_i$ совпадают, что противоречит предположениям леммы.

Примеры и применения этого параграфа взяты у Дубинса и Спанье [1961]. Доказательства, приведенные выше, являются немного измененными и улучшенными доказательствами этой работы.

Глава IX

# МИНИМАКСНАЯ АППРОКСИМАЦИЯ, НЕРАВЕНСТВО МАРКОВА — БЕРНШТЕЙНА И СВЯЗАННЫЕ С НИМИ ЗАДАЧИ

Основным инструментом при получении результатов этой главы является теорема представления (теорема 10.1 гл. II), выражающая положительный многочлен в виде единственной комбинации двух других многочленов, обладающих специальными осцилляционными свойствами. Многие экстремальные задачи легко решаются при подходящем использовании теоремы представления. К этой категории задач мы относим ряд задач наилучшей минимаксной аппроксимации, неравенство Маркова — Бернштейна для многочленов и различные экстремальные задачи типа тех, что рассматривались Бернштейном [1926].

Пусть $\{u_i\}_0^n$ — чебышевская система на конечном интервале $[a, b]$. Пусть далее $p(t)$ и $q(t)$ — непрерывные функции на $[a, b]$, удовлетворяющие неравенствам $p(t) > u(t) > q(t)$ для некоторого многочлена $u(t)$. Теорема 10.2 гл. II утверждает существование двух многочленов $\bar{u}$ и $\underline{u}$ со специальными осцилляционными свойствами. Точнее, $\bar{u}$ и $\underline{u}$ лежат между $p$ и $q$ и касаются функций $p$ и $q$ поочередно в $n+1$ точках на $[a, b]$.

Решения многих экстремальных задач в теории аппроксимации функций многочленами связаны с специальными многочленами $\bar{u}$ и $\underline{u}$ при подходящем выборе $p$ и $q$. Этот способ основывается на простом подсчете корней, учитывающем единственность $\underline{u}$ и $\bar{u}$ с указанными осцилляционными свойствами. Действительно, многочлены $\underline{u}$ и $\bar{u}$ покрывают промежуток между $p(t)$ и $q(t)$ не менее $n$ раз при $t$, пробегающем $[a, b]$. Поэтому, если $u(t)$ — многочлен, лежащий между $p$ и $q$, то $\bar{u}-u$ и $\underline{u}-u$ имеют по крайней мере $n$ корней при условии, что неузловые корни считаются дважды. Очевидно, далее, что $\bar{u}-u$ и $\underline{u}-u$ не могут иметь $n+1$

корней, если они не обращаются в нуль тождественно. Отсюда следует, что $\bar{u} - u$ и $u - \underline{u}$ удовлетворяют некоторым неравенствам. Такие неравенства могут рассматриваться как экстремальные характеристики $\bar{u}$ и $\underline{u}$, соответствующие различным вариационным задачам.

## § 1. Многочлены наилучшего приближения

Имеется обширная литература по конструктивной теории наилучшей аппроксимации. Значительная часть книги Натансона [1949] посвящена обзору многочисленных ответвлений этой теории. Гл. II у Ахиезера [1965] также хорошо отражает результаты по наилучшей аппроксимации в смысле минимаксной нормы. Первыми важный вклад в эту область внесли Чебышев, Хаар, Ахиезер, Золотарев, Валле-Пуссен, Бернштейн и другие. Отправной точкой для многих дальнейших исследований является знаменитый трактат Бернштейна [1926], в котором рассматриваются многие экстремальные задачи.

Во многих работах содержатся исследования по наилучшей аппроксимации для функций многих переменных и в пространствах более общего вида. В этой связи мы сошлемся, например, на Ривлина и Шапиро [1961], Ньюмана и Шапиро [1963], Райса [1963], Начбина [1965] и на библиографию в этих работах. Интенсивно изучался и вопрос об единственности наилучшей аппроксимации. Мы не будем перечислять большое число подходящих ссылок.

Пусть $C[a, b]$ обозначает пространство всех непрерывных функций, определенных на $[a, b]$, и пусть $\|f\|$ означает норму

$$\|f\| = \max_{a \leqslant t \leqslant b} |f(t)|, \qquad f \in C[a, b]. \tag{1.1}$$

Рассмотрим задачу аппроксимации (по норме (1.1)) функции $f \in C[a, b]$ $u$-многочленами по $T$-системе $\{u_i\}_0^n$ на интервале $[a, b]$. Именно, требуется определить или по крайней мере охарактеризовать многочлен $u = \sum_{i=0}^{n} a_i u_i$, для которого величина

$$\left\| f - \sum_{i=0}^{n} a_i u_i \right\| \tag{1.2}$$

достигает минимума.

При изучении наилучшей аппроксимации функции $f$ многочленами часто используют также $L_r$-нормы

$$\left\| f - \sum_{i=0}^{n} a_i u_i \right\|_r = \left( \int_a^b \left| f - \sum_{i=0}^{n} a_i u_i \right|^r d\mu \right)^{1/r}, \qquad r > 0, \tag{1.3}$$

где $\mu$ — фиксированная мера.

Случай, когда $r = 2$, а $u_i(t) = t^i$ $(i = 0, 1, \dots, n)$, приводит к рассмотрению ортогональных многочленов.

Случай $r = 1$ обсуждался в § 11 гл. VIII. Норма (1.1) есть предельный случай (1.3) при $r \to \infty$.

Данный параграф посвящен раскрытию внутренней связи между решением минимаксной задачи

$$\min_{a_0, \dots, a_n} \max_{a \leqslant t \leqslant b} \left| f(t) - \sum_{i=0}^{n} a_i u_i(t) \right| \tag{1.4}$$

и многочленами $\bar{u}$ и $\underline{u}$ из теоремы 10.2 гл. II при подходящем выборе $p(t)$ и $q(t)$.

Начнем с классической характеризации решения (1.4) (ср. Бернштейн [1926]).

Теорема 1.1. (1) *Многочлен, минимизирующий* (1.4), *определяется единственным образом для любой функции* $f \in C[a, b]$ *в том и только том случае, когда функции* $\{\varepsilon u_0, u_1, \dots, u_n\}$ *образуют* $T$*-систему при* $\varepsilon = +1$ *или* $-1$.

(2) *Если* $\{u_i\}_0^n$ — $T$*-система, то для каждой функции* $f \in C[a, b]$ *единственный многочлен* $v_* = \sum_{i=0}^{n} a_i u_i$, *минимизирующий* (1.4), *характеризуется тем, что существуют* $n+2$ *точки* $\{t_i\}_1^{n+2}$ $(a \leqq t_1 < t_2 < \dots < t_{n+2} \leqq b)$ *такие, что*

$$(-1)^i \delta \{f(t_i) - v_*(t_i)\} = \max_{a \leqslant t \leqslant b} |f(t) - v_*(t)|, \qquad i = 1, 2, \dots, n+2, \tag{1.5}$$

*где* $\delta = +1$ *или* $-1$.

Для удобства изложения доказательство теоремы 1.1 отнесено в § 3. Полезно указать, что если многочлен $v_* = \sum_{i=0}^{n} a_i u_i$ удовлетворяет условию (1.5), то $v_*$ — многочлен, наименее уклоняющийся от функции $f(t)$. Это утверждение легко доказать с помощью метода подсчета корней, описанного во введении. В самом деле, пусть $v_*$ удовлетворяет условию (1.5). Это означает, в частности, что $[a, b]$ можно разбить по крайней мере на $n+1$ отрезков таких, что на каждом отрезке $v_*$ изменяется от $f(x) + m$ до $f(x) - m$ $(m = \| f - v_* \|)$. В таком случае любой многочлен, удовлетворяющий условию $|f(t) - u(t)| \leqq m$, $t \in [a, b]$, очевидно, обладает тем свойством, что $v_* - u$ имеет не менее $n+1$ нулей (не менее одного на каждом отрезке), если неузловые нули считаются дважды. Обращаясь к теореме 4.2 гл. I, мы заключаем, что $u(t) \equiv v_*(t)$. Проведенное рассуждение доказывает, что многочлен $v_*$ наименее уклоняется от $f(x)$ и определяется свойством (1.5) единственным образом.

Пример 1.1. Одно из приложений теоремы 1.1 проистекает из следующей задачи: среди всех многочленов степени $n$ с главным коэффициентом, равным единице, найти многочлен, абсолютный максимум которого минимален, т. е. определить значения коэффициентов $a_1, a_2, \ldots, a_n$, для которых величина

$$\max_{-1 \leqslant t \leqslant 1} |t^n + a_1 t^{n-1} + \ldots + a_n| \tag{1.6}$$

достигает минимума.

На языке теоремы 1.1 это означает, что нужно найти наилучшую аппроксимацию функции $f(t) = t^n$ многочленом степени $n - 1$. Эта задача в геометрической постановке была изучена ранее как часть теоремы 4.2 гл. IV. Более прямое доказательство использует критерий теоремы 1.1 и проводится следующим образом.

Пусть

$$R_n(t) = \frac{1}{2^{n-1}}, \quad T_n(t) = \frac{1}{2^{n-1}} \cos n\theta, \qquad t = \cos\theta,$$

где $T_n(t)$ обозначает $n$-й многочлен Чебышева первого рода.

Очевидно, главный коэффициент $R_n(t)$ равен единице и

$$\max_{-1 \leqslant t \leqslant 1} |R_n(t)| = \frac{1}{2^{n-1}}.$$

Кроме того, в точках

$$t_k = -\cos\frac{k\pi}{n}, \qquad k = 0, 1, \ldots, n,$$

имеем

$$R_n(t_k) = \frac{(-1)^{n-k}}{2^{n-1}}, \qquad k = 0, 1, \ldots, n.$$

Следовательно, многочлен $R_n(t)$ принимает свое максимальное абсолютное значение $1/2^{n-1}$ с чередующимися знаками в точках $t_0, t_1, \ldots$ $\ldots, t_n$. Из теоремы 1.1 мы заключаем, что наименее уклоняющийся от нуля многочлен с главным коэффициентом единица равен $2^{-n+1} T_n(t)$.

Другие, более сложные примеры применения теоремы 1.1 представлены в § 2. Наилучшее приближение к $t^n$ на конечном интервале многочленом степени $n - 1$ было впервые изучено Чебышевым. В книге Бернштейна [1950] содержатся многочисленные примеры и характеристики решений соответствующих задач для $T$-систем (см. также гл. 2 у Ахиезера [1965]). Наилучшее минимальное приближение многочленами $n - 2$ степени к функции $t^n + At^{n-1}$ (где $A$ задано) получено Золотаревым (см. Ахиезер [1965], стр. 280). Эта задача связана с изучением эллиптических функций.

Периодический аналог теоремы 1.1 дается следующим результатом.

Теорема 1.2. *Если $\{u_i\}_0^{2m}$ — $T$-система периодических функций на $[a, b)$, то многочлен, наименее уклоняющийся от заданной непрерывной функции $f$ на $[a, b]$, характеризуется единственным* *образом тем, что существуют* $2m+2$ *точек* $\{t_i\}_1^{2m+2}$ $(t_1 < t_2 < \dots$ $\dots < t_{m+2} < t_1 + b - a)$ *таких, что*

$$(-1)^i \delta \{f(t_i) - u(t_i)\} = \max_{a \leqslant t \leqslant b} |f(t) - u(t)|, \qquad \delta = +1 \text{ или } -1$$

*при* $i = 1, 2, \dots, 2m+2$.

## § 2. Многочлены наилучшего приближения (продолжение)

Полезно теперь вывести соотношения между экстремальным решением $v_*$ задачи (1.4) и многочленами $\underline{u}$ и $\bar{u}$ из теоремы 10.2 гл. II. Многочлены $\underline{u}$ и $\bar{u}$ играют основную роль в случае, когда расширенная система $\{u_0, u_1, \dots, u_n, f\}$ представляет собой $T$-систему дополнительно к обычному предположению о том, что $\{u_i\}_0^n$ — $T$-система. Пусть это требование выполняется. Мы хотим определить экстремальный многочлен, на котором достигается

$$\min_{a_0, \dots, a_n} \max_{a \leqslant t \leqslant b} \left| f(t) - \sum_{i=0}^{n} a_i u_i(t) \right|. \tag{2.1}$$

Построим многочлены $\bar{u}$ и $\underline{u}$, существование которых гарантируется теоремой 10.2 гл. II для системы $\{u_0, u_1, \dots, u_n, f\}$, связанной с функциями $p(t) \equiv 1$ и $q(t) \equiv -1$. Именно, многочлен

$$\bar{u}(t) = \sum_{i=0}^{n} b_i u_i(t) + d f(t) \tag{2.2}$$

характеризуется условиями:

(1) $-1 \leq \bar{u}(t) \leq 1$;

(2) существуют $n+2$ точек $\{s_i\}_1^{n+2}$ $(a \leq s_1 < s_2 < \dots < s_{n+2} \leq b)$ таких, что $\bar{u}(s_i) = (-1)^{n+1-i}$, $i = 1, 2, \dots, n+2$.

Подчеркнем, что функция $f(t)$ входит в $T$-систему, используемую при построении многочлена $\bar{u}$. Отметим также, что в рассматриваемом случае при $p(t) = +1$ и $q(t) = -1$, имеем $\underline{u}(t) = -\bar{u}(t)$.

Коэффициент $d$ при $f(t)$ в формуле (2.2) не может быть равен нулю. В противном случае многочлен $\sum_{i=0}^{n} b_i u_i$ обращается в нуль в $n+1$ точках и, следовательно, тождественно равен нулю.

Используя свойства (1) и (2), мы выводим, что

$$\left| f(t) - \left( -\frac{1}{d} \sum_{i=0}^{n} b_i u_i(t) \right) \right| \leq \frac{1}{|d|},$$

и равенство достигается при условиях, данных в теореме 1.1.

Таким образом, справедлива

Теорема 2.1. *Если функции* $u_0, u_1, \ldots, u_n$ *образуют* $T$-*систему, то многочлен* $-\frac{1}{d}\sum_{i=0}^{n} b_i u_i(t)$, *определенный посредством* (2.2), *является наименее уклоняющимся от функции* $f(t)$, *и минимальное уклонение равно* $1/|d|$.

Сделаем несколько замечаний, относящихся к случаю, когда расширенное множество функций $u_0, u_1, \ldots, u_n, f$ не образует $T$-систему. Мы исключаем тривиальный случай, когда функция $f$ есть $u$-многочлен.

I. Пусть $C$ обозначает множество чисел $c > 0$, для которых существует многочлен, лежащий строго между $f(t) - c$ и $f(t) + c$. По теореме 10.2 гл. II для каждого $c \in C$ существует многочлен (по системе $u_0, u_1, \ldots, u_n$), имеющий максимальное колебание между $f(t) + c$ и $f(t) - c$. Именно, $u(t; c)$ лежит между $f(t) + c$ и $f(t) - c$ и существуют $n + 1$ точек $\{s_i\}_1^{n+1}$ $(a \leq s_1 < s_2 < \ldots$ $\ldots < s_{n+1} \leq b)$ таких, что

$$u(s_i; c) = f(s_i) + \delta(-1)^i c, \qquad i = 1, 2, \ldots, n + 1, \tag{2.3}$$

где $\delta$ равно $+1$ или $-1$. Многочлен $u(t; c)$ из теоремы 10.2 гл. II может быть выбран равным либо $\bar{u}$, либо $\underline{u}$.

Пусть $c_0 = \inf\{c \mid c \in C\}$. Выберем последовательность $c_k \to c_0$, $c_k \in C$ так, что $u(t; c_k)$ сходится к некоторому многочлену $v(t; c_0)$. Тогда, разумеется,

$$f(t) - c_0 \leq v(t; c_0) \leq f(t) + c_0. \tag{2.4}$$

Если теперь

$$c_0' = \min_{\{a_i\}} \max_{t} \left| f(t) - \sum_{i=0}^{n} a_i u_i(t) \right|, \tag{2.5}$$

то мы легко выводим из определения $c_0$ и $c_0'$, что $c_0' = c_0$. Следовательно, $v(t; c_0) = v_*(t)$ — экстремальный многочлен для (2.1) и равенство должно достигаться попеременно слева и справа в (2.4) не менее $n + 2$ раз.

Заметим далее, что предельный многочлен $v(t; c_0)$. единствен и не зависит от аппроксимирующей последовательности $u(t; c_k)$ в силу единственности многочлена наименьшего уклонения. Итак,

$$\lim_{c \to c_0} u(t; c) = v_*(t) = v(t; c_0), \tag{2.6}$$

и, следовательно, при приближении $c$ к $c_0$ многочлены $\bar{u}(t; c)$ и $\underline{u}(t; c)$ из теоремы 10.2 гл. II стремятся к общему пределу.

II. Если $\{u_i\}_0^{2m}$ — периодическая $T$-система на $[a, b]$ и $f(t)$ — непрерывная периодическая функция на $[a, b]$, то для больших $c$ существует однопараметрическое семейство многочленов $u(t; t_0, c)$, удовлетворяющих

$$f(t) - c \leq u(t; t_0, c) \leq f(t) + c, \tag{2.7}$$

которые покрывают расстояние между $f(t) + c$ и $f(t) - c$ не менее $2m$ раз, в то время как $t$ пробегает $[a, b]$. Допустим теперь, что $c_0$ является нижней гранью тех значений $c$, для которых существует многочлен, лежащий строго между $f(t) - c$ и $f(t) + c$, и пусть $v_*(t)$ — единственный наименее уклоняющийся от $f(t)$ многочлен, полученный в теореме 1.2. Используя те же соображения, что и в пункте I, мы заключаем, что

$$\lim_{c \to c_0} u(t; t_0, c) = v_*(t) \quad \text{для всех } t_0.$$

## § 3. Доказательство теоремы 1.1

Следующее доказательство теоремы 1.1 стандартно и его можно найти во многих работах по теории аппроксимации (см., например, Ахиезер [1965]).

Доказательство. Предположим сначала, что $\{\varepsilon u_0, u_1, \ldots, u_n\}$ не является $T$-системой. Тогда существуют $n + 1$ различных точек $t_0 < t_1 < \ldots < t_n$ $(t_i \in [a, b))$ таких, что

$$\det [\| u_i(t_j) \|_{i,j=0}^n] = 0. \tag{3.1}$$

Следовательно, существуют вещественные константы $c_0, c_1, \ldots, c_n$, $\sum_{i=0}^n c_i > 0$, удовлетворяющие уравнениям $\sum_{i=0}^n c_i u_k(t_i) = 0$, $k = 0, 1, \ldots, n$, и такие, что для любого многочлена $u$ выполняется условие

$$\sum_{i=0}^n c_i u(t_i) = 0. \tag{3.2}$$

Далее из (3.1) следует также существование ненулевого многочлена $\tilde{u} = \sum_{i=0}^n b_i u_i(t)$, обращающегося в нуль в точках $t_0, t_1, \ldots, t_n$.

Пусть функция $g$, принадлежащая классу $C[a, b]$, такова, что $\| g(t) \| \leq 1$ и

$$g(t_i) = c_i / |c_i|, \quad \text{если } c_i \neq 0.$$

Если $\lambda > 0$ выбирается так, чтобы выполнялось $\| \lambda\tilde{u} \| < 1$, то $f(t) = g(t)[1 - |\lambda\tilde{u}(t)|]$ имеет тот же знак, что и $g(t)$ в точках $t_i$ при $c_i \neq 0$.

Построим теперь бесконечный ряд многочленов с одним и тем же минимальным уклонением от функции $f(t)$. Если $\left\| f - \sum_{i=0}^n a_i u_i \right\| < 1$

для некоторых $a_0, a_1, \ldots, a_n$, то справедливо неравенство

$$-1 < g(t_j)[1 - |\lambda\tilde{u}(t_j)|] - \sum_{i=0}^{n} a_i u_i(t_j) < 1, \qquad j = 0, 1, \ldots, n,$$

которое сводится к неравенству

$$-1 < g(t_j) - \sum_{i=0}^{n} a_i u_i(t_j) < 1, \qquad j = 0, 1, \ldots, n.$$

Поэтому, если $c_j \neq 0$, то величина $\sum_{i=0}^{n} a_i u_i(t_j)$ имеет тот же знак, что и $c_j$, а $\sum_{j=0}^{n} c_j \sum_{i=0}^{n} a_i u_i(t_j) \neq 0$. Но это противоречит условию (3.2). Следовательно,

$$\left\| f - \sum_{i=0}^{n} a_i u_i \right\| \geqslant 1.$$

Если теперь $|\delta| \leq 1$, то

$$|f(t) - \delta\lambda\tilde{u}(t)| \leq |f(t) + \delta\lambda\tilde{u}(t)| \leq |g(t)|(1 - |\lambda\tilde{u}(t)|) +$$
$$+ |\delta\lambda\tilde{u}(t)| \leq 1 - (1 - |\delta|)|\lambda\tilde{u}(t)| \leq 1,$$

поэтому $\delta\lambda\tilde{u}(t)$ — экстремальный многочлен для $|\delta| \leq 1$.

Положим теперь, что $\{\varepsilon u_0, u_1, \ldots, u_n\}$ — $T$-система и установим справедливость утверждения (2) теоремы. Отсюда будет легко следовать утверждение пункта (1) об единственности.

Если $\{\varepsilon u_0, u_1, \ldots, u_n\}$ — $T$-система, то легко доказать существование минимизирующего многочлена. Установим справедливость второй части теоремы доказательством от противного. Предположим, что $v = \sum_{i=0}^{n} b_i u_i$ удовлетворяет условию

$$\|f - v\| = m = \min_{a_0, \ldots, a_n} \left\| f - \sum_{i=0}^{n} a_i u_i \right\|$$

и что $f(t) - v(t)$ принимает поочередно значения $\pm m$ только в $k \leq n + 1$ точках. Предположим, для определенности, что в первой из этих точек $f(t) - v(t)$ принимает значение $+m$. Тогда существует $k - 1$ точек $y_1, y_2, \ldots, y_{k-1}$ $(a < y_1 < y_2 < \ldots < y_{k-1} < b)$ таких, что

$$f(y_i) - v(y_i) = 0, \qquad i = 1, 2, \ldots, k - 1,$$

и для некоторого $\mu > 0$.

$$m \geqslant f(t) - v(t) \geqslant -m + \mu, \qquad t \in [a, y_1] \cup [y_2, y_3] \cup \ldots,$$
$$m - \mu \geqslant f(t) - v(t) \geqslant -m, \qquad t \in [y_1, y_2] \cup [y_3, y_4] \cup \ldots$$

По теореме 5.2 гл. I существует многочлен $w$, все нули которого на открытом интервале $(a, b)$ суть узловые нули в точках $y_1, y_2, \ldots$ $\ldots, y_{k-1}$ и $w(t) \leq 0$, $a \leq t \leq y_1$. Если теперь $\delta > 0$ выбрано так, что $|\delta w(t)| \leq \mu/2$, то для $t \in (a, b)$ выполняется неравенство

$$|f(t) - v(t) + \delta w(t)| < m. \tag{3.3}$$

Равенство для конечной точки $a$ в (3.3) возможно только в случае, когда $f(a) - v(a) = m$ и $w(a) = 0$, а для точки $b$ — только в случае, если $|f(b) - v(b)| = m$ и $w(b) = 0$. Чтобы получить неравенство в точках $a$ и $b$, положим, что $w_1(t)$ — любой многочлен, для которого $w_1(t)[f(t) - v(t)] > 0$ при $t = a, b$. Тогда при достаточно малых $\eta$ имеем

$$|f(t) - v(t) + \delta w(t) - \eta w_1(t)| < m$$

для всех $t \in [a, b]$. Следовательно,

$$\min \left\| f - \sum_{i=0}^{n} a_i u_i \right\| < m;$$

это противоречит тому, что $m$ — минимальное уклонение.

Таким образом, всякий многочлен наименьшего уклонения удовлетворяет условию (2).

Утверждение единственности в части (2) было доказано ранее.

## § 4. Примеры применения теоремы 1.1

Для полноты изложения приведем три простых примера применения теоремы 1.1, на которые будем в дальнейшем ссылаться. Представления будут в каждом случае явно указаны.

Пример 4.1 (ср. Ахиезер [1965], стр. 58). Мы хотим найти многочлен вида $\sum_{i=0}^{n} a_i t^i$, наименее уклоняющийся от функции $1/(t-c)$ $(c > 1)$ на интервале $[-1, 1]$.

Отметим попутно, что система $\{1, t, t^2, \ldots, t^n, -1/(t-c)\}$ есть $T$-система на интервале $[-1, 1]$. Это немедленно следует из того, что функция

$$\sum_{i=0}^{n} a_i t^i + \frac{a}{t-c} \qquad \left(\sum_{i=0}^{n} a_i^2 + a^2 > 0\right)$$

имеет, очевидно, не более $n+1$ нулей на $[-1, 1]$. Знак определителя (1.1 гл. I) для рассматриваемой системы, как легко видеть, положителен.

Во всех примерах, которые здесь рассматриваются, метод построения наименее уклоняющегося многочлена заключается в нахождении многочлена, обладающего осцилляционными свойствами, указанными в теореме 1.1. Этот многочлен и дает наилучшую аппроксимацию.

Для данной задачи рассмотрим функцию

$$\psi(t) = \frac{K}{2}\left[z^n \frac{\beta - z}{1-\beta z} + z^{-n}\frac{1-\beta z}{\beta - z}\right], \tag{4.1}$$

где $t = 2^{-1}(z+z^{-1})$, $c = 2^{-1}(\beta+\beta^{-1})$ или $\beta = c - \sqrt{c^2-1} < 1$, $\beta > 0$ и $K = 4\beta^{n+2}(1-\beta^2)^{-2}$.

Поскольку $\psi(t)$ — рациональная функция от $z$ и $\psi(t(z)) = \psi(t(z^{-1}))$, то имеем

$$\psi(t) = \frac{P(z)}{Q(z)} = \frac{P(z^{-1})}{Q(z^{-1})} = \frac{P(z)+P(z^{-1})}{Q(z)+Q(z^{-1})} = \frac{P_1(z+z^{-1})}{Q(z+z^{-1})} = \frac{P_1(2t)}{Q_1(2t)},$$

где $P$, $Q$, $P_1$ и $Q_1$ — многочлены. Следовательно, $\psi(t)$ — рациональная функция от $t$, у которой единственный конечный полюс расположен в точке $t = c$. Кроме того, из соотношения

$$\lim_{t\to c}(t-c)\psi(t) = \frac{K(1-\beta^2)^2}{4\beta^{n+2}} = 1$$

вытекает, что

$$\psi(t) = \frac{1}{t-c} - P_n(t), \tag{4.2}$$

где $P_n(t)$ — многочлен степени не более $n$.

Теперь при $t$, пробегающем справа налево отрезок $[-1,1]$, из соотношения

$$t = \frac{1}{2}\left(z+\frac{1}{z}\right) = \frac{e^{i\theta}+e^{-i\theta}}{2} = \cos\theta$$

следует, что $z$ пробегает верхнюю половину окружности $\{e^{i\theta} \mid 0 \leq \theta \leq \pi\}$. Легко проверить, что $|\psi(t)| \leq K$, так как при $|z| = 1$ мы имеем $\left|z^n \frac{\beta-z}{1-\beta z}\right| = 1$. Кроме того, $\psi(t)$ принимает значение $K$ или $-K$, если

$$\arg\left\{z^n \frac{\beta - z}{1-\beta z}\right\} \tag{4.3}$$

равен 0 или $\pi \pmod{2\pi}$. Далее, так как функция $z^n \frac{\beta-z}{1-\beta z}$ имеет $n+1$ нуль при $|z| < 1$, мы заключаем с помощью принципа аргумента, что величина (4.3) возрастает от $\pi$ до $(n+2)\pi$, когда $z$ описывает верхнюю половину единичной окружности.

Поэтому функция $\psi(t)$ принимает свое максимальное абсолютное значение $K$ с чередующимся знаком в $n+2$ точках $[-1, 1]$ и многочлен $P_n(t)$, определенный посредством (4.2), удовлетворяет требованиям теоремы 1.1. Следовательно, $P_n(x)$ есть многочлен, наименее уклоняющийся от функции $(t-c)^{-1}$ на $[-1, 1]$. Значение наименьшего уклонения дается выражением $K = (c-\sqrt{c^2-1})^n/(c^2-1)$.

Уместно сравнить «многочлен» $\psi(t)$ с парой осцилляционных многочленов $\bar{u}$ и $\underline{u}$ теоремы 10. 2 гл. II. Мы видим из (4.2), что $\psi(t)$ — многочлен по отношению к $T$-системе $\{1, t, \ldots, t^n, -1/(t-c)\}$. Многочлены $\bar{u}$ и $\underline{u}$, колеблющиеся между $p(t) \equiv K$ и $q(t) \equiv -K$, выражаются в явном виде как $\bar{u}(t) = \psi(t)$ и $\bar{u}(t) = -\psi(t)$. Определение знака возможно, так как $\psi(t) \to -\infty$ при $t \uparrow c$.

Пример 4.2. Сформулируем без доказательства следующее обобщение примера 1 (см. Ахиезер [1965], стр. 249).

Пусть

$$w(t) = \prod_{i=1}^{2q}\left(1 - \frac{1}{c_i}\right), \qquad c_i > 1, \qquad i = 1, 2, \ldots, 2q,$$

$$t = 2^{-1}(z + z^{-1}), \qquad |z| \leq 1, \qquad c_k = 2^{-1}(\beta_k + \beta_k^{-1}),$$

$$|\beta_k| < 1, \qquad k = 1, 2, \ldots, 2q,$$

$$\Omega(z) = \prod_{k=1}^{2q} \sqrt{z - \beta_k},$$

и

$$\mathcal{L}_m = \begin{cases} \dfrac{1}{2^{m-1}} \displaystyle\prod_{k=1}^{2q} \sqrt{1 + \beta_k^2}, & m > q, \\[2ex] \dfrac{1}{2^{q-1}} \dfrac{1}{1 + \beta_1\beta_2 \ldots \beta_{2q}} \displaystyle\prod_{k=1}^{2q} \sqrt{1 + \beta_k^2}, & m = q. \end{cases}$$

Тогда

$$T_m(t; w) = \frac{\mathcal{L}_m}{2}\left\{\frac{z^{2q-m}\Omega(z^{-1})}{\Omega(z)} + z^{m-2q}\frac{\Omega(z)}{\Omega(z^{-1})}\right\}\sqrt{w(t)}$$

— многочлен степени $m$ относительно переменной $t$ с единичным главным коэффициентом и

$$\min_{\{b_k\}} \max_{-1 \leq t \leq 1} \frac{|t^m + b_1 t^{m-1} + \ldots + b_m|}{\sqrt{w(t)}} = \mathcal{L}_m. \tag{4.4}$$

Многочлен $T_m(t; w)$ — экстремальный.

Доказательство (4.4) получается так же, как в примере 4.1, с использованием функции $z^{2q-m}\dfrac{\Omega(z^{-1})}{\Omega(z)}$.

Утверждение в примере 4.1 может быть выведено из утверждения примера 4.2.

Можно дать ряд различных интерпретаций для связи многочлена $T_m(t; w)$ с функциями $\bar{u}$ и $\underline{u}$ теоремы 10.2 гл. II. Если положить

$p(t) = \mathcal{L}_m \sqrt{w(t)}$ и $q(t) = -\mathcal{L}_m \sqrt{w(t)}$, то $T_m(t; w)$ есть многочлен $\underline{u}(t)$ по системе $\{t^i\}_0^m$ для этих $p$ и $q$. А если отнести множитель $\sqrt{w(t)}$ к степенным функциям, то $T_m(t, w)/\sqrt{w(t)}$ — многочлен $\underline{u}(t)$ по $T$-системе $\{t^i/\sqrt{w(t)}\}_0^m$, который максимально колеблется между константами $p(t) \equiv \mathcal{L}_m$ и $q(t) \equiv -\mathcal{L}_m$.

Пример 4.3. Получим теперь аналог примера 4.2 для системы периодических функций $1, \cos\theta, \sin\theta, \ldots, \cos m\theta, \sin m\theta$. А именно, наша задача заключается в минимизации выражения

$$\max_{0 \leqslant \theta \leqslant 2\pi} |G(\theta)|, \tag{4.5}$$

где

$$G(\theta) = \frac{H(\theta)}{h(\theta)}. \tag{4.6}$$

Функция $h(\theta)$ есть фиксированный положительный многочлен степени $l$, а $H(\theta)$ — произвольный многочлен вида

$$H(\theta) = A\cos m\theta + B\sin m\theta + \sum_{k=0}^{m-1}(a_k \cos k\theta + b_k \sin k\theta), \tag{4.7}$$

где $m \geqslant l$, $A$ и $B$ $(A^2 + B^2 > 0)$ заданы.

Следуя правилу теоремы 1.2, мы должны найти многочлен вида (4.7) с заданным главным коэффициентом, обладающий тем свойством, что $G(\theta)$ принимает свое максимальное абсолютное значение $2m$ раз с чередующимися знаками или, что эквивалентно, нам нужно определить многочлен по отношению к периодической $T$-системе

$$\frac{1}{h(\theta)}, \quad \frac{\cos\theta}{h(\theta)}, \quad \frac{\sin\theta}{h(\theta)}, \ldots, \frac{\cos(m-1)\theta}{h(\theta)}, \quad \frac{\sin(m-1)\theta}{h(\theta)},$$

который наименее уклоняется от функции

$$\frac{1}{h(\theta)}(A\cos m\theta + B\sin m\theta).$$

Для того чтобы построить такой многочлен, используем классическую теорему факторизации для положительных тригонометрических многочленов и определим

$$h(\theta) = |g(z)|^2,$$

где $z = e^{i\theta}$ и

$$g(z) = \gamma\prod_{\nu=1}^{l}(z - z_\nu) = \gamma z^l + b_1 z^{l-1} + \ldots + \gamma',$$

$$\gamma > 0, \quad |z_\nu| < 1, \quad \nu = 1, 2, \ldots, l.$$

Рассмотрим функцию

$$\gamma^2 G_0(\theta) = \mathcal{R}\left\{c z^{m-1}\frac{g(z)}{g^*(z)}\right\}, \quad z = e^{i\theta}, \quad m \geqslant l,$$

где

$$c = \begin{cases} A - iB, & m \geqslant l+1, \\ \dfrac{A - iB + \bar{k}(A+iB)}{1-|k|^2}, & m = l, \quad k = (\gamma'/\gamma)^2, \end{cases}$$

и $g^*(z)$ — обратный многочлен, определенный выражением $g^*(z) = z^l\bar{g}(z^{-1}) = \gamma' z^l + \ldots + \bar{b}_1 z + \gamma = \gamma \prod_{\nu=1}^{l} (1 - \bar{z}_\nu z)$. Заметим, что $|\gamma'/\gamma| = \left|\prod_{\nu=1}^{l} z_\nu\right| < 1$, поэтому $c$ определено корректно. Так как

$$\gamma^2 G_0(\theta) = \frac{1}{2}\left\{\frac{cz^{m-2l}g(z)}{\overline{g(z)}} + \frac{\bar{c}z^{2l-m}\overline{g(z)}}{g(z)}\right\} = \frac{1}{2}\left\{\frac{cz^{m-2l}(g(z))^2 + \bar{c}z^{2l-m}\overline{(g(z))^2}}{|g(z)|^2}\right\},$$

то можно легко проверить, что числитель $G_0(\theta)$ имеет требуемый вид. Те же рассуждения, что и в примере 4.1, показывают, что аргумент функции

$$\omega = cz^{m-l}\frac{g(z)}{g^*(z)} = cz^{m-l}\prod_{\nu=1}^{l}\frac{(z - z_\nu)}{1 - \bar{z}_\nu z}, \qquad z = e^{i\theta}$$

возрастает, пробегая интервал длины $2m\pi$ при $\theta$, возрастающем от 0 до $2\pi$, и $G_0(\theta)$ принимает свое максимальное абсолютное значение

$$\max_\theta |G_0(\theta)| = \gamma^{-2}|c| \tag{4.8}$$

с чередующимися знаками не менее $2m$ раз. Следовательно, функция $G_0(\theta)$ минимизирует (4.5) над классом функций, определенных посредством (4.6) и (4.7).

Для того чтобы интерпретировать функцию $G_0(\theta)$ в терминах теорем представления § 6 гл. VI, воспроизведем следующий результат теоремы 6.2 гл. VI. Пусть $\{u_i\}_0^{2m}$ — $T$-система периодических функций на $[0, 2\pi]$, а $p(\theta)$ и $q(\theta)$ — непрерывные периодические функции на $[0, 2\pi]$ такие, что для некоторого многочлена $v$ выполняется $p(\theta) > v(\theta) > q(\theta)$. Теорема 6.2 гл. VI утверждает, что в этом случае существует однопараметрическое семейство многочленов $u(\theta; \theta_0)$, обладающих свойствами

(1) $p(\theta) \geqslant u(\theta; \theta_0) \geqslant q(\theta)$,

(2) $v(\theta) - u(\theta; \theta_0)$ обращается в нуль в $2m$ различных точках, одна из которых $\theta_0$,

(3) существует $2m$ точек $s_1 < s_2 < \ldots < s_{2m}$, которые перемежаются с нулями $v(\theta) - u(\theta; \theta_0)$ и

$$u(s_i; \theta_0) = \begin{cases} p(s_i), & i \text{ — нечетное}, \\ q(s_i), & i \text{ — четное}. \end{cases}$$

Полный класс многочленов $u(\theta; \theta_0)$ порождается значениями параметра $\theta_0$ из полуоткрытого интервала, ограниченного двумя смежными корнями фиксированного многочлена $u(\theta; \theta')$. Более того, класс многочленов $u(\theta; \theta_0)$, полученный таким образом, не зависит от конкретного вида функции $v(\theta)$.

Далее из (4.5) и (4.8) получаем

$$|G_0(\theta)| \leq \gamma^{-2}|c|$$

или

$$|\gamma^{-2}\mathcal{R}\left[\frac{cz^{m-2l}(g(z))^2}{|g(z)|^2}\right]| \leqslant \gamma^{-2}|c|, \qquad z = e^{i\theta}.$$

Последнее неравенство можно также записать в виде

$$-\gamma^{-2}h(\theta) \leqslant \gamma^{-2}\mathcal{R}\{e^{i\varphi}z^{m-2l}(g(z))^2\} \leqslant \gamma^{-2}h(\theta), \qquad z = e^{i\theta}, \quad (4.9)$$

где $e^{i\varphi} = \frac{c}{|c|}$. Так как равенство возникает попеременно слева и справа в (4.9) не менее $2m$ раз, то средний член в (4.9) есть необходимо один из многочленов $u(\theta; \theta_0)$, построенных в теореме 6.2 гл. VI. При $\varphi$, пробегающем $[0, 2\pi)$, эти многочлены образуют полный класс многочленов $u(\theta; \theta_0)$. Таким образом, класс многочленов

$$\gamma^{-2}\mathcal{R}\{e^{i\varphi}z^{m-2l}(g(z))^2\}, \qquad z = e^{i\theta}, \qquad e^{i\varphi} = \frac{c}{|c|},$$

параметризованный посредством $\varphi \in [0, 2\pi]$, совпадает с классом многочленов $u(\theta; \theta_0)$, $\theta_0 \in [0, 2\pi]$. Чтобы доказать это утверждение, мы замечаем, что всякий многочлен $u(\theta; \theta_0)$ имеет главные члены $A\cos m\theta + B\sin m\theta$, где $A^2 + B^2 > 0$. Поэтому мы можем определить $\varphi$ так, что $\gamma^{-2}\mathcal{R}\{e^{i\varphi}z^{m-2l}(g(z))^2\}$ имеет те же самые главные члены, и, таким образом, разность рассматриваемых многочленов тождественно равна нулю, так как имеет не менее $2m$ нулей.

Пример с тригонометрическими функциями из этого параграфа изучен Сеге [1964]. Бернштейн рассматривал менее полный вариант аналогичной задачи для обычных многочленов с полиномиальной весовой функцией. Полное решение задачи приводится у Ахиезера [1965] (стр. 256).

## § 5. Обобщение неравенства Маркова — Бернштейна

Две знаменитые теоремы, принадлежащие Бернштейну и Маркову, дают экстремальные характеризации многочленов Чебышева. Происхождение этих результатов связано со следующими задачами.

Задача 5.1. Найти многочлен $m-1$ степени, который доставляет максимум выражению

$$\max_{-1 \leqslant t \leqslant 1} |P_{m-1}(t)| \qquad (5.1)$$

среди всех многочленов степени $m-1$, удовлетворяющих условиям

$$(1-t^2)\,P_{m-1}^2(t) \leqslant 1, \qquad -1 \leqslant t \leqslant 1. \tag{5.2}$$

Задача 5.2. Пусть $P_m(t)$ обозначает произвольный многочлен, удовлетворяющий ограничениям

$$\max_{-1 \leqslant t \leqslant 1} |P_m(t)| \leqslant 1. \tag{5.3}$$

Найти верхнюю границу для $\max\limits_{-1 \leqslant t \leqslant 1} |P'_m(t)|$.

Экстремальный многочлен в обоих случаях оказывается классическим многочленом Чебышева. Решение задач 5.1 и 5.2 приводит к результату, известному под названием неравенств Маркова — Бернштейна (см. § 8 ниже).

Неравенство Маркова — Бернштейна является важным классическим результатом и часто плодотворно применяется в теории интерполяции и аппроксимации.

Обычный метод решения задачи 5.2 заключается в сведении ее к задаче 5.1. Формулируют сначала тригонометрический вариант задачи 5.2, который легко решается. Затем можно соединить результат в тригонометрическом случае с решением задачи 5.1 и, таким образом, получить решение задачи 5.2 (более детально об этом методе см. в § 8).

Экстремальная характеризация и доказательство единственности при решении задачи 5.1 основываются на существовании многочлена, имеющего максимальное колебание при ограничениях (5.2). В классическом варианте задачи 5.1 этот многочлен есть $U_{m-1}(t)$ — многочлен Чебышева 2-го рода. Основная идея доказательства заключается в том, что всякий другой многочлен $P_{m-1}(t)$, удовлетворяющий ограничениям (5.2), не может доставлять большее значение в (5.1), ибо в этом случае разность $P_{m-1}^2(t) - U_{m-1}^2(t)$ имеет слишком много нулей.

Развитие и обобщение сформулированных задач происходит в нескольких направлениях. В большинстве случаев обобщается только задача 5.2. Например:

(1) Задача 5.2 обобщалась на некоторые случаи целых функций порядка $\rho$, см. Боас [1954] (гл. 8), Ахиезер [1956] (стр. 140) и Бернштейн [1923 b].

(2) Систематически изучалась задача 5.2 с $P_m(z)$, определенным на комплексной плоскости (см. Сеге [1925], Бернштейн [1937a] и Дочев [1962]).

(3) Обобщение на многомерный случай, формулируемое в терминах гармонических многочленов. При этом производная в задаче 5.2 заменяется на градиент (см., например, Сеге [1928]). У Хормандера [1954] эти исследования связаны с некоторыми понятиями теории гиперболических конусов.

Мы предлагаем новый ряд обобщений, отправной точкой для которых является задача 5.1. Эти обобщения состоят в замене классических многочленов, формируемых из $T$-системы $\{t^i\}_0^n$, $u$-многочленами, получаемыми из произвольной $T$-системы $\{u_i\}_0^n$.

Мы подготовлены теперь к формулировке нашего варианта обобщенного неравенства Бернштейна. Положим для начала, что $\{u_i\}_0^n$ — $ET$-система второго порядка на конечном интервале $[a, b]$. Подчеркнем, что это предположение подразумевает существование непрерывных производных от функций $u_0, u_1, \ldots, u_n$. Свойство $ET$ второго порядка необходимо для доказательства единственности экстремальных функций. В этой связи будет полезна следующая лемма. Обозначим через $Z_2^*(u)$ число нулей многочлена $u$, где нули кратности 2 и больше засчитываются дважды, а все остальные нули — по одному разу.

Лемма 5.1. *Если* $\{u_i\}_0^n$ — *ET-система второго порядка, то* $Z_2^*(u) \leqslant n$ *для любого ненулевого многочлена* $u = \sum_{i=0}^{n} a_i u_i$ (см. теорему 4.3 гл. I).

Пусть $p(t)$ — непрерывная на $[a, b]$ функция и $q(t) = \sum_{i=0}^{n} a_i u_i(t)$ — фиксированный многочлен, удовлетворяющий неравенству $p(t) > q(t)$, $t \in [a, b]$. Рассмотрим класс многочленов

$$\mathcal{U} = \{u \mid q(t) \leqslant u(t) \leqslant p(t), \quad t \in [a, b]; \quad u(a) = q(a)\}. \qquad (5.4)$$

Теорема 10.2 гл. II и ее следствие утверждают существование единственного многочлена $v_* \in \mathcal{U}$, характеризуемого следующим свойством.

Свойство $A$. $v_* - q$ обращается в нуль на множестве индекса $n/2$, а $p - v_*$ обращается в нуль по крайней мере один раз между любыми двумя нулями $v_* - q$ и между наибольшим внутренним нулем и конечной точкой $b$. Специальный многочлен $v_*$ есть в действительности многочлен $\bar{u}$, когда $n$ четно и $\underline{u}$, когда $n$ нечетно.

В предыдущих параграфах были указаны соотношения между многочленами $\underline{u}$ и $\bar{u}$ и экстремальным многочленом в задаче наилучшей аппроксимации. Многочлен $v_*$ обладает еще некоторыми замечательными экстремальными свойствами, которые формулируются ниже в теоремах 5.1 и 5.2.

Теорема 5.1. *Пусть* $\mathcal{U}$ *определено посредством* (5.4), *и пусть* $v_*$ — *единственный многочлен, характеризуемый свойством* $A$. *Тогда*

$$\max_{u \in \mathcal{U}} u'(a) \qquad (5.5)$$

*доставляется единственным многочленом* $v_*$.

Доказательство. Так как многочлены в $\mathcal{U}$ равномерно ограничены, то такими же будут и коэффициенты всех многочленов из $\mathcal{U}$. Следовательно, максимум в (5.5) достигается. Если многочлен $w$ дает этот максимум, то $w'(a) \geqslant v_*'(a)$.

Пусть $s_1$ — наименьший из нулей $p(t) - v_*(t)$. Ясно, что $w(t) - v_*(t)$ имеет не менее $n-1$ нулей на $[s_1, b]$ при условии, что неузловые нули считаются дважды. Если $w'(a) > v_*(a)$, то $w(t) - v_*(t)$ имеет нуль в $(a, s_1]$, что вместе с конечной точкой $a$ и $n-1$ нулями в $[s_1, b]$ составляет $n+1$ нулей. Если $w'(a) = v_*'(a)$, то $w(t) - v_*(t)$ имеет нуль в точке $t = a$, кратность которого не меньше двух. Итак, во всех случаях $w(t) - v_*(t)$ имеет $n+1$ нулей, где кратные нули считаются дважды. Из леммы 5.1 следует, что $w(t) \equiv v_*(t)$.

Введем теперь класс многочленов несколько более узкий, чем (5.4), а именно

$$\mathcal{U}_0 = \begin{cases} \mathcal{U}, & n \text{ — нечетное}, \\ \{u \mid u \in \mathcal{U},\ u(b) = q(b)\}, & n \text{ — четное}. \end{cases} \tag{5.6}$$

Допустим, что $p$ принадлежит классу $C^1$ и что:

(1) для четных $n$ функция

$$f(t) = \frac{p(t) - q(t)}{(b-t)(t-a)}$$

строго убывает на $(a, t_0)$ и строго возрастает на $(t_0, b)$ при некотором $t_0 \in (a, b)$;

(2) для нечетного $n$ функция

$$g(t) = \frac{p(t) - q(t)}{t-a}$$

строго убывает на $[a, b]$.

Будем считать, что функции в (5.7) и (5.8) доопределены по непрерывности всякий раз, когда знаменатель обращается в нуль.

Т е о р е м а 5.2 (обобщенное неравенство Маркова — Бернштейна). *Пусть выполняются предположения* (1) *и* (2). *Многочлен* $v_* \in \mathcal{U}_0$, *построенный в теореме* 10.2 *гл.* II *и в точности характеризуемый свойством* $A$, *является единственным многочленом, который доставляет*

$$\max_{u \in \mathcal{U}_0} \max_{t \in [a,b]} \frac{u(t) - q(t)}{(b-t)(t-a)} \tag{5.7}$$

*при четном* $n$ *и*

$$\max_{u \in \mathcal{U}_0} \max_{t \in [a,b]} \frac{u(t) - q(t)}{t-a} \tag{5.8}$$

*при нечетном* $n$.

З а м е ч а н и е 5.1. В § 8 будет точно указано соотношение между этой теоремой и задачей 5.1, а также будет дан ряд других применений теоремы.

Д о к а з а т е л ь с т в о. Мы докажем только (5.7). Отметим, что многочлен $v_*$ обладает тем свойством, что

$$\frac{v_*(t) - q(t)}{(b-t)(t-a)} \tag{5.9}$$

колеблется между 0 и $f(t)$ $(a<t<b)$, совпадая с $f(t)$ в $m$ $(n=2m)$ точках $s_1, \ldots, s_m$ и обращаясь в нуль в $m-1$ точках $t_1, \ldots, t_{m-1}$; $a<s_1<t_1<s_2<t_2<\ldots<t_{m-1}<s_m<b$. Из свойств монотонности $f(t)$ (по предположению (1)) следует, что максимум в (5.7) достигается для некоторого $t \in [a, s_1] \cup [s_m, b]$. Если теперь $w \in \mathcal{U}_0$ является экстремальным для (5.7) и $w \neq v_*$, то на основе теоремы 5.1 получим, что

$$\frac{w'(a)-q'(a)}{b-a} < \frac{v_*'(a)-q'(a)}{b-a}.$$

Для интервала $[a, b]$ из теоремы 5.1 вытекает, что

$$\frac{w'(b)-q'(b)}{b-a} > \frac{v_*'(b)-q'(b)}{b-a}.$$

Существует $x_0 \in (a, s_1] \cup [s_m, b)$, при котором выполняется неравенство

$$\frac{w(x_0)-q(x_0)}{(b-x_0)(x_0-a)} \geqslant \frac{v_*(x_0)-q(x_0)}{(b-x_0)(x_0-a)}.$$

Однако в таком случае $w-v_*$ имеет $n-1$ нуль на $[a, b]$, считая неузловые нули дважды, поэтому $w-v_*$ имеет $n+1$ нуль на $[a, b]$, и, следовательно, $w \equiv v_*$.

Доказательство того, что $v_*$ — единственный экстремальный многочлен для (5.8), достигается подобными рассуждениями. Доказательство теоремы закончено.

В теореме 5.1 было показано, что значение производной от многочлена $v_*$, характеризуемого свойством $A$, в конечной точке $a$ максимально в классе $\mathcal{U}$. То же самое свойство для $v_*$ выполняется в некотором смысле, который мы сейчас опишем, для произвольной точки $z \in [a, b]$.

В случае нечетного $n$, скажем, $n=2m+1$, существуют точки $a=\tilde{t}_0<\tilde{s}_1<\tilde{t}_1<\tilde{s}_2<\ldots<\tilde{t}_m<\tilde{s}_{m+1} \leqslant b$, в которых многочлен $v_*$ удовлетворяет соотношениям

$$v_*(\tilde{t}_k)=q(\tilde{t}_k), \qquad k=0, 1, \ldots, m,$$

$$v_*(\tilde{s}_k)=p(\tilde{s}_k), \qquad k=1, 2, \ldots, m+1.$$

Для случая $n=2m+2$ единственное изменение состоит в том, что $\tilde{s}_{m+1}<b$ и $v_*(b)=q(b)$. При $t$, пробегающем значения из интервала типа $(\tilde{t}_i, \tilde{s}_{i+1})$, многочлен $v_*$ возрастает от значения $q(\tilde{t}_i)$ до значения $p(\tilde{s}_{i+1})$, а на интервале типа $(\tilde{s}_i, \tilde{t}_i)$ значения $v_*(t)$ убывают от $p(\tilde{s}_i)$ до $q(\tilde{t}_i)$. Таким образом, $v_*$ возрастает на $(\tilde{t}_i, \tilde{s}_{i+1})$ и убывает на $(\tilde{s}_i, \tilde{t}_i)$.

Учитывая это, определим множества $A$ и $B$ следующим образом:

$$A = (a, \tilde{s}_1) \bigcup (\tilde{t}_1, \tilde{s}_2) \bigcup \ldots \bigcup (\tilde{t}_m, \tilde{s}_{m+1}),$$

$$B = \begin{cases} (\tilde{s}_1, \tilde{t}_1) \bigcup (\tilde{s}_2, \tilde{t}_2) \bigcup \ldots \bigcup (\tilde{s}_m, \tilde{t}_m), & n = 2m+1, \\ (\tilde{s}_1, \tilde{t}_1) \bigcup (\tilde{s}_2, \tilde{t}_2) \bigcup \ldots \bigcup (\tilde{s}_m, \tilde{t}_m) \bigcup (\tilde{s}_{m+1}, b), & n = 2m+2. \end{cases}$$

Пусть для каждого $z \in [a, b]$ $\mathcal{U}(z)$ — класс многочленов

$$\mathcal{U}(z) = \{u \mid q(t) \leqslant u(t) \leqslant p(t), \quad t \in [a, b], \quad u(z) = v_*(z)\}.$$

Теорема 5.3 *Экстремумы*

$$\max_{u \in \mathcal{U}(z)} u'(z), \qquad z \in A, \tag{5.10}$$

$$\min_{u \in \mathcal{U}(z)} u'(z), \qquad z \in B \tag{5.11}$$

*достигаются на единственном многочлене* $v_*$.

Доказательство. Пусть $z \in A$ и $w$ — многочлен, для которого в (5.10) достигается экстремальное значение. Тогда $w'(z) \geqslant v'_*(z)$. Если $w'(z) > v'_*(z)$, то $w - v_*$ имеет не менее $n$ нулей, не считая точки $z$, при условии, что неузловые нули считаются дважды. Итак, $w(z) = v_*(z)$ и, следовательно, $w = v_*$. Если $w'(z) = v'_*(z)$, то $w - v$ имеет снова $n+1$ нуль, где нуль в точке $z$ считается дважды. В случае (5.11) доказательство точно такое же.

Замечание 5.2. Еще раз подчеркнем, что если $\{u_i\}_0^n$ не есть $ET$-система второго порядка, а просто $T$-система, то теорема остается в силе, за исключением утверждения об единственности.

## § 6. Обобщенные неравенства Маркова — Бернштейна для бесконечных интервалов

Утверждение теоремы 5.2 легко распространяется на полубесконечный интервал $[0, \infty)$. Предположим, что $\{u_i\}_0^n$ — $ET$-система второго порядка на $[0, \infty)$ и дополнительно что выполняются требования:

(1) $u_n(t) > 0$, $t \geqslant \hat{t}$, для некоторого $\hat{t} > 0$;
(2) $\lim\limits_{t \to \infty} u_i(t)/u_n(t) = 0$, $i = 0, \ldots, n-1$;
(3) $\{u_i\}_0^{n-1}$ — $T$-система на $[0, \infty)$.

Рассмотрим непрерывно дифференцируемую на $[0, \infty)$ функцию $f(t) > 0$ такую, что предел

$$\lim_{t \to \infty} \frac{f(t)}{u_n(t)} \tag{6.1}$$

существует, и его значение положительно или равно $+\infty$.

Отметим, что указанным предположением удовлетворяют функции $u_i(t) = t^i$, $i = 0, 1, \ldots, n$, и $f(t) = e^t$.

Мы хотим найти многочлен $\widetilde{u}$, доставляющий максимум

$$\max_{u \in \Lambda} \max_{t \in [0, \infty)} \frac{u(t)}{tf(t)}, \tag{6.2}$$

где $\Lambda$ — класс многочленов, определенный условиями

$$\Lambda = \{u \mid 0 \leqslant u(t) \leqslant f(t),\ u(0) = 0\},$$

и значение $t^{-1}u(t)$ полагается равным $u'(0)$ в точке $t = 0$.

Чтобы решить эту задачу, найдем сначала многочлен (как в теореме 5.1), на котором достигается максимум

$$\max_{u \in \Lambda} u'(0). \tag{6.3}$$

Как и в случае конечного интервала, экстремальные многочлены, на которых достигаются максимумы (6.2) и (6.3), совпадают.

Пусть $w(t)$ — строго положительная на $[0, \infty)$ функция такая, что $w(t) = u_n(t)$, $t \geqslant \hat{t}$, и пусть

$$v_k(x) = \begin{cases} \dfrac{u_k(\operatorname{tg} x)}{w(\operatorname{tg} x)}, & x \in [0, \pi/2), \\ \delta_{kn}, & x = \pi/2, \end{cases}$$

$$\bar{f}(x) = \frac{f(\operatorname{tg} x)}{w(\operatorname{tg} x)}, \qquad x \in [0, \pi/2].$$

Система $\{v_k\}_0^n$ — $T$-система на $[0, \pi/2]$ и $\bar{f}(x) > 0$ на $[0, \pi/2)$, причем значение $\bar{f}(x)$ в точке $\pi/2$ либо положительно, либо равно бесконечности в силу условия, наложенного на (6.1).

Предположим сначала, что $\bar{f}(\pi/2)$ конечно. Пусть $v_*$ — многочлен, обладающий свойством $A$ из § 5. Если $\bar{f}(\pi/2)$ бесконечно, то построим $v_*$ для

$$\bar{f}_N(x) = \begin{cases} \bar{f}(x), & f(x) \leqslant N, \\ f(x), & f(x) > N, \end{cases}$$

причем для всех достаточно больших $N$ получается один и тот же многочлен $v_*$.

Преобразуем теперь многочлен $v_*$ следующим образом:

$$u_*(t) = w(\operatorname{tg}^{-1} t)\, v_*(\operatorname{tg}^{-1} t), \quad t \in [0, \infty). \tag{6.4}$$

При $n = 2m$ многочлен $u_*(t)$ обращается в нуль в точках

$$0 = \widetilde{t}_0 < \widetilde{t}_1 < \ldots < \widetilde{t}_{m-1}, \tag{6.5}$$

а в точках $\widetilde{s}_1, \widetilde{s}_2, \ldots, \widetilde{s}_m$ совпадает с $f(t)$. Эти точки перемежаются следующим образом:

$$\widetilde{t}_0 < \widetilde{s}_1 < \widetilde{t}_1 < \ldots < \widetilde{t}_{m-1} < \widetilde{s}_m < \infty. \tag{6.6}$$

В случае $n = 2m + 1$ соответствующие точки удовлетворяют соотношениям

$$0 = \tilde{t}_0 < \tilde{s}_1 < \tilde{t}_1 < \tilde{s}_2 < \ldots < \tilde{s}_m < \tilde{t}_m < \tilde{s}_{m+1} < \infty. \qquad (6.7)$$

Отметим, что в четном случае мы имеем $2m - 1$ нулей, считая неузловые нули дважды. В $+\infty$ имеется еще один нуль и, следовательно, $v_*(t) = \alpha_0 v_0(t) + \ldots + \alpha_{n-1} v_{n-1}(t)$, последний коэффициент равен нулю, так как $v_*$ равно нулю в точке $\pi/2$.

Теорема 6.1. *Экстремальные значения*

$$\max_{u \in \Lambda} u'(0), \qquad (6.8)$$

$$\max_{u \in \Lambda} \max_{t \in [0, \infty]} \frac{u(t)}{tf(t)}, \qquad (6.9)$$

*где* $\Lambda = \{u \mid 0 \leq u(t) \leq f(t),\ u(0) = 0\}$ *достигаются на единственном многочлене* $u_*$, *определенном в* (6.4).

Доказательство этой теоремы аналогично предшествующим рассуждениям. Отметим, что $u(t)/f(t) \leq 1$, а функция $1/t$ убывает, и поэтому максимум в (6.9) достигается на интервале $(0, \tilde{s}_1)$. Далее доказательство такое же, как в теореме 5.2.

## § 7. Обобщенные неравенства Маркова — Бернштейна для периодических функций

В этом параграфе мы изложим некоторые аналоги обобщенных неравенств Маркова — Бернштейна для периодических функций. В этом случае теория в некоторых отношениях богаче, как будет видно из дальнейшего. Например, утверждение теоремы 7.3, формулируемой ниже, не имеет аналога в непериодическом случае. Дополнительная структура возникает в результате того, что мы теперь имеем однопараметрическое семейство осциллирующих многочленов вместо двух специальных многочленов $\underline{u}$ и $\bar{u}$.

Естественно, что в этом случае мы полагаем $n$ четным, скажем, $n = 2m$, и предполагаем, что $\{u_i\}_0^{2m}$ образуют $ET$-систему порядка 2. Иногда мы будем требовать, чтобы $\{u_i\}_0^{2m}$ являлось $ET$-системой порядка 3. Обобщенные чебышевские условия второго и третьего порядков накладываются для того, чтобы обеспечить единственность решения различных экстремальных задач. Единственность в каждом случае вытекает из того факта, что если $\{u_i\}_0^{2m}$ — $ET$-система порядка 2, то каждый ненулевой многочлен имеет не более $n$ нулей, где нули второй и большей кратности считаются дважды. Для $ET$-систем порядка 3 нули учитываются вплоть до кратности 3.

Пусть $p(t)$ — положительная, периодическая и непрерывная на $[a, b)$ функция. При фиксированном $t_0 \in [a, b]$ обозначим через $v(t; t_0)$ многочлен, описанный в теореме 6.1 гл. VI и обладающий следующими свойствами:

(1) $p(t) \geqslant v(t;\, t_0) \geqslant 0$;

(2) $v(t;\, t_0)$ имеет $m$ различных нулей, один из которых $t_0$;

(3) $p(t) - v(t;\, t_0)$ обращается в нуль не менее одного раза между любыми двумя нулями $v(t;\, t_0)$ (рассматриваемыми в циклическом порядке).

Рассмотрим класс многочленов

$$\mathcal{W}(t_0) = \{u \mid 0 \leq u(t) \leq p(t),\ t \in [a, b),\ u(t_0) = 0\}.$$

Экстремальная характеризация $v(t;\, t_0)$ дается следующей теоремой (ср. теорему 5.1).

Теорема 7.1. *Если $\{u_i\}_0^{2m}$ — ET-система третьего порядка, то*

$$\max_{u \in \mathcal{W}(t_0)} u''(t_0) \tag{7.1}$$

*достигается на единственном многочлене $v(t;\, t_0)$.*

Доказательство. Пусть $w$ — произвольный многочлен из $\mathcal{W}(t_0)$, для которого

$$w''(t_0) = \max_{u \in \mathcal{W}(t_0)} u''(t_0).$$

Тогда $w''(t_0) \geqslant v''(t_0;\, t_0)$. Если $w''(t_0) > v''(t_0;\, t_0)$, то функция $w(t) - v(t;\, t_0)$ имеет не менее $2m+1$ нулей, считая кратные нули дважды. Если $w''(t_0) = v''(t_0;\, t_0)$, то $w(t) - v(t;\, t_0)$ имеет не менее $2m+1$ нулей, причем нуль в точке $t_0$ имеет порядок 3. В обоих случаях приходим к противоречию с условиями, наложенными на систему $\{u_i\}_0^{2m}$, если не выполняется $w(t) \equiv v(t;\, t_0)$.

Наша следующая задача заключается в формировании аналога для теоремы 5.3. Пусть $p(t)$ и $q(t)$ — непрерывные периодические на $[a, b)$ функции, для которых $p(t) > q(t)$, и предположим, что существует многочлен $\tilde{u}$, для которого $p(t) > \tilde{u}(t) > q(t)$. При каждом $t_0 \in [a, b)$ теорема 6.2 гл. VI утверждает существование единственного многочлена $u(t;\, t_0)$, обладающего свойствами:

(1) $p(t) \geqslant u(t;\, t_0) \geqslant q(t)$;

(2) $\tilde{u}(t_0) = u(t_0;\, t_0)$ и существуют $n$ точек $\{s_i\}_1^n$ таких, что $t_0 < s_1 < s_2 < \ldots < s_n < t_0 + b - a$, а $u(t;\, t_0)$ равно $q(t)$ и $p(t)$ поочередно в точках $s_1, s_2, \ldots, s_n$.

Однопараметрическое семейство $u(t;\, t_0)$ (с параметром $t_0$) изучалось в § 2, где осцилляционные свойства $u(t;\, t_0)$ использовались для иллюстрации соотношения между этими многочленами и экстремальными многочленами в задаче наилучшей аппроксимации. В § 4 был дан пример класса многочленов $u(t;\, t_0)$ для тригонометрических многочленов в случае, когда $p(t) = -q(t) = h(t)$, где $h(t)$ — положительный многочлен степени не более $m$.

На осцилляционных свойствах основано решение некоторых экстремальных задач, что мы покажем далее, в теореме 7.2. Для формулировки теоремы 7.2 нам потребуются некоторые предварительные замечания.

Пусть точка $z_0 \in [a, b)$ и величина $c$ таковы, что $q(z_0) < c < p(z_0)$. Так как коэффициенты $u(t; t_0)$ — непрерывные функции параметра $t_0$, то существуют два многочлена $\bar{u}_{z_0}$ и $\underline{u}_{z_0}$, оба равные $c$ в точке $z_0$, удовлетворяющие условиям (1) и (2), за исключением того, что $\bar{u}_{z_0}$ и $\underline{u}_{z_0}$ изменяются в противоположных направлениях. Именно, пусть $\bar{u}_{z_0}(t) = u(t; \bar{t}_0)$ и $\underline{u}_{z_0}(t) = u(t; \underline{t}_0)$. Если $\{\bar{s}_i\}_1^n$ есть значения $s$, соответствующие $\bar{t}_0$ и $\bar{s}_{i_0} < z_0 < \bar{s}_{i_0+1}$, то $u(\bar{s}_{i_0}; \bar{t}_0) = q(\bar{s}_{i_0})$ и $u(\bar{s}_{i_0+1}; \bar{t}_0) = p(\bar{s}_{i_0+1})$. Точно так же, если $\{s_i\}_1^n$ — значения $s$, соответствующие $\underline{t}_0$ и $\underline{s}_{j_0} < z_0 < \underline{s}_{j_0+1}$, то $u(\underline{s}_{j_0}; \underline{t}_0) = p(\underline{s}_{j_0})$ и $u(\underline{s}_{j_0+1}; \underline{t}_0) = q(\underline{s}_{j_0+1})$.

Рассмотрим класс многочленов $\mathcal{V}(z_0)$, определенный выражением

$$\mathcal{V}(z_0) = \{u \mid q(t) \leq u(t) \leqslant p(t),\ t \in [a, b],\ u(z_0) = c\},$$

где $q(z_0) < c < p(z_0)$.

Теорема 7.2. *Пусть $\{u_i\}_0^{2m}$ — ET-система порядка 2, и пусть при фиксированном $z_0 \in [a, b)$ $\bar{u}_{z_0}$ и $\underline{u}_{z_0}$ такие, как определено выше. Тогда*

$$\max_{u \in \mathcal{V}(z_0)} u'(z_0) = \bar{u}'_{z_0}(z_0), \tag{7.2}$$

$$\min_{u \in \mathcal{V}(z_0)} u'(z_0) = \underline{u}'_{z_0}(z_0). \tag{7.3}$$

*В каждом случае экстремальный многочлен единствен.*

Доказательство снова осуществляется соответственным подсчетом нулей и использованием того факта, что $\bar{u}_{z_0}$ и $\underline{u}_{z_0}$ осциллируют между $p(t)$ и $q(t)$ максимальное число раз. Детали мы опускаем.

Теперь мы подготовлены к доказательству главной теоремы настоящего параграфа. Пусть допустимый класс многочленов состоит из

$$\mathcal{V}_h = \{u \mid |u(t)| \leqslant h(t)\},$$

где $h(t)$ — строго положительная непрерывная периодическая на $[a, b)$ функция.

Теорема 7.3. *Пусть $\{u_i\}_0^{2m}$ — периодическая ET-система порядка 2. Экстремальное значение*

$$\max_{u \in \mathcal{V}_h} \max_{t \in [a, b)} |u'(t)| \tag{7.4}$$

*достигается на многочлене $u(t; t_0)$ при некотором $t_0$. Иными словами, при вычислении максимума (7.4) достаточно ограничить внимание однопараметрическим семейством многочленов $u(t; t_0)$.*

Доказательство. Рассмотрим произвольный многочлен $u \in \mathcal{V}_h$ и фиксируем точку $t^* \in [a, b)$. Мы различаем два случая: $|u(t^*)| < h(t^*)$ и $|u(t^*)| = h(t^*)$. В первом случае используем теорему 7.2, которая утверждает существование многочлена $u(t; t_0) \in \mathcal{V}_h$ для некоторого $t_0$ со свойствами

$$u(t^*;\, t_0) = u(t^*), \quad |u'(t^*;\, t_0)| \geqslant |u'(t^*)|. \tag{7.5}$$

Во втором случае увеличим функцию $h$ до $h_\varepsilon$ так, что $|u(t^*)| < h_\varepsilon(t^*)$. Многочлены $u_\varepsilon(t;\, t_0)$ и их производные $u'_\varepsilon(t;\, t_0)$ зависят от $\varepsilon$ равномерно непрерывно по $t$. (Это так, поскольку коэффициенты этих многочленов — непрерывные по $\varepsilon$ и $t_0$ функции.) Мы можем теперь обратиться к предыдущему случаю и вывести снова справедливость (7.5) с $u(t^*;\, t_0)$, замененными на $u_\varepsilon(t^*;\, t_0)$. (Здесь $t_0$ также может зависеть от $\varepsilon$.) Производя предельный переход по $\varepsilon$, получаем справедливость (7.5).

Второе соотношение в (7.5) можно записать в виде

$$\sup_{t_0} |u'(t^*;\, t_0)| \geqslant \sup_{u \in \mathcal{V}^h} |u'(t^*)|.$$

Так как $t^*$ — произвольная точка $[a, b)$, то предложение теоремы установлено.

Пример 7.1. Рассмотрим $T$-систему тригонометрических функций

$$1,\ \cos\theta,\ \sin\theta,\ \cos 2\theta,\ \sin 2\theta,\ \ldots,\ \cos m\theta,\ \sin m\theta. \tag{7.6}$$

Многочлены $u(\theta;\, \theta_0)$, колеблющиеся между $p(\theta) \equiv 1$ и $q(\theta) \equiv -1$, как установлено в теореме 6.2 гл. VI, имеют вид $\sin m(\theta + \theta_0)$. Применение теоремы 7.3 дает следующий результат. Если $g(\theta)$ — тригонометрический многочлен $m$-й степени и $|g(\theta)| \leqslant 1$, то $|g'(\theta)| \leqslant m$, причем равенство достигается в том и только в том случае, когда $g(\theta) = \sin m(\theta + \theta_0)$. Единственность следует из теоремы 7.2.

Пример 7.2. Рассматриваемая $T$-система есть снова (7.6). Пусть $h(\theta)$ — фиксированный положительный тригонометрический многочлен степени $l \leqslant n$, и пусть

$$h(\theta) = |g(z)|^2,$$

где $z = e^{i\theta}$ и

$$g(z) = \gamma \prod_{\nu=1}^{l} (z - z_\nu), \qquad \gamma > 0, \qquad |z_\nu| < 1, \qquad \nu = 1, 2, \ldots, l.$$

Класс многочленов с максимальным осциллированием $u(\theta;\, \theta_0)$, $0 \leq \theta_0 \leqslant 2\pi$, лежащих между $h(\theta)$ и $-h(\theta)$, совпадает с классом многочленов

$$\mathcal{R}\,[e^{i\varphi} z^{m-2l} (g(z))^2], \qquad z = e^{i\theta}, \qquad m \geqslant l \tag{7.7}$$

(см. пример 4.3), где $\varphi$ — вещественный параметр.

Предположим, что $P(\theta)$ — тригонометрический многочлен порядка не более $m$, удовлетворяющий неравенству

$$|P(\theta)| \leqslant \gamma^{-2} h(\theta).$$

Тогда экстремальное значение

$$\max_P \max_\theta |P'(\theta)| \tag{7.8}$$

достигается для многочлена вида

$$\gamma^{-2} R\,[e^{i\varphi} z^{m-2l} (g(z))^2], \qquad z = e^{i\theta},$$

где $\varphi$ — параметр.

В особом случае, когда $h(\theta) = 1 - 2r\cos\theta + r^2$ ($|r| < 1$), элементарные вычисления показывают, что значение (7.8) равно $(1 + |r|)^2 (m + |r|(m-2))^2$.

## § 8. Примеры и приложения

Данный параграф представляет ряд конкретных иллюстраций обобщенных неравенств Маркова—Бернштейна, исследованных в § 5.

А. Все результаты этого раздела являются классическими, однако полезно получить их с помощью методов, предложенных в предшествующих параграфах.

Мы начнем с представления

$$1 = T_m^2(t) + (1 - t^2) U_{m-1}^2(t), \qquad -1 \leqslant t \leqslant 1, \tag{8.1}$$

где

$$T_m(t) = \cos m\theta \qquad (t = \cos\theta),$$
$$U_m(t) = \frac{1}{m+1} T'_{m+1}(t) = \frac{\sin(m+1)\theta}{\sin\theta} \tag{8.2}$$

— многочлены Чебышева первого и второго родов соответственно. В обозначениях теоремы 10.1 гл. II уравнение (8.1) выражает представление многочлена $u(t) \equiv 1$ в терминах экстремальных многочленов $\underline{u}(t) = T_m^2(t)$ и $\overline{u}(t) = (1 - t^2) U_{m-1}^2(t)$.

(1) Пусть в теореме 5.1 $q(t) \equiv 0$, $p(t) \equiv 1$. Тогда класс многочленов $\mathcal{U}$, определенный в (5.4) при $n = 2m$ — это класс многочленов $P_{2m}(t) = \sum_{i=0}^{n} a_i t^i$, удовлетворяющих соотношениям

$$0 \leqslant P_{2m}(t) \leqslant 1, \quad -1 \leqslant t \leqslant 1 \text{ и } P_{2m}(-1) = 0. \tag{8.3}$$

Применяя теорему 5.1, получим

$$P'_{2m}(-1) \leqslant \frac{d}{dt}(1 - t^2) U_{m-1}^2(t)\Big|_{t=-1}, \tag{8.4}$$

и равенство возникает только в том случае, когда

$$P_{2m}(t) = v_*(t) = (1 - t^2) U_{m-1}^2(t).$$

(2) Задача 5.1, поставленная в начале § 5, заключалась в вычислении максимума

$$\max_{-1 \leqslant t \leqslant 1} |P_{m-1}(t)| \tag{8.5}$$

над множеством многочленов степени $m-1$, удовлетворяющих условию

$$(1-t^2)\, P_{m-1}^2(t) \leqslant 1, \qquad -1 \leq t \leq 1. \tag{8.6}$$

Решение этой задачи заключено в следующем, несколько более общем результате. Рассмотрим класс $\mathcal{U}_0$ (см. (5.6)) многочленов, удовлетворяющих (8.3) и дополнительному условию $P_{2m}(+1) = 0$. Применяя теорему 5.2, имеем для $P_{2m} \in \mathcal{U}_0$ соотношение

$$\max_{-1 \leqslant t \leqslant 1} \frac{P_{2m}(t)}{1-t^2} \leqslant \max_{-1 \leqslant t \leqslant 1} U_{m-1}^2(t) = m^2, \tag{8.7}$$

причем равенство возникает только в том случае, когда $P_{2m}(t) = (1-t^2)\, U_{m-1}^2(t)$.

Кроме того, рассуждение по индукции с использованием соотношения

$$U_{m-1}(t) = \frac{\sin m\theta}{\sin \theta} = \cos (m-1)\theta + \cos \theta \, \frac{\sin (m-1)\theta}{\sin \theta}, \qquad m \geqslant 1, \tag{8.8}$$

показывает, что равенство в правой части (8.7) достигается только при $t = \pm 1$.

Решение задачи 5.1, которое мы только что описали, может быть выражено в следующем виде: если $P_{m-1}(t)$ — многочлен степени $m-1$ на $[-1, 1]$ и

$$\sqrt{1-t^2}\, |P_{m-1}(t)| \leq 1, \tag{8.9}$$

то

$$|P_{m-1}(t)| \leqslant m, \tag{8.10}$$

где равенство выполняется только в случае $P_{m-1}(t) = \gamma U_{m-1}(t)$, $(|\gamma| = 1)$ и $t = \pm 1$.

(3) Продолжая иллюстрацию использования результатов § 5, рассмотрим теорему 5.3. Многочлен $(1-t^2)\, U_{m-1}^2(t)$ обращается в нуль в точках

$$t_k = -\cos \frac{k\pi}{m}, \qquad k = 0, 1, \ldots, m, \tag{8.11}$$

и равен единице в нулях $T_m(t)$, т. е. в точках

$$s_k = \cos \frac{(2k-1)\pi}{2m}, \qquad k = 1, 2, \ldots, m. \tag{8.12}$$

Теорема 5.3 утверждает, что при $z \in (-1, s_1) \cup (t_1, s_2) \cup \ldots \cup (t_{m-1}, s_m)$ многочлен $(1-t^2)\, U_{m-1}^2(t)$ имеет максимальное значение производной в точке $z$ среди всех многочленов $P_{2m}$, удовлетворяющих соотношениям $0 \leqslant P_{2m}(t) \leqq 1$ и $P_{2m}(z) = (1 - t^2) U_{m-1}^2(z)$.

Следующие три результата охватывают классическое доказательство, использованное при выводе решения задачи 5.2 § 5.

(4) Положим $S_m(\theta) = \mu_1 \sin\theta + \mu_2 \sin 2\theta + \ldots + \mu_m \sin m\theta$. Записывая $P_{m-1}(t) = P_{m-1}(\cos\theta) = S_m(\theta)/\sin\theta$, мы можем легко вывести из (8.9) и (8.10), что если многочлен $S_m(\theta)$ удовлетворяет неравенству

$$|S_m(\theta)| \leqq 1, \tag{8.13}$$

то

$$\left| \frac{S_m(\theta)}{\sin\theta} \right| \leqslant m, \tag{8.14}$$

и равенство достигается только при $S_m(\theta) = \pm \sin m\theta$.

(5) В примере 7.1 мы показали, что если многочлен

$$g_m(\theta) = \lambda_0 + \lambda_1 \cos\theta + \mu_1 \sin\theta + \ldots + \lambda_m \cos m\theta + \mu_m \sin m\theta$$

удовлетворяет $|g_m(\theta)| \leqq 1$, то $|g_m'(\theta)| \leqslant m$. Доказательство этого результата использует (4) и осуществляется следующим образом. Если

$$S(\theta_0; \theta) = \frac{g_m(\theta + \theta_0) - g_m(\theta - \theta_0)}{2},$$

то $S(\theta_0; \theta)$ — синус-многочлен от $\theta_0$, удовлетворяющий (8.13). В таком случае

$$\left| \frac{S(\theta_0; \theta)}{\sin\theta_0} \right| \leqq m.$$

Устремляя $\theta_0$ к нулю, получим, что $|g_m'(\theta)| \leqq m$.

(6) Мы готовы теперь к решению задачи 5.2 § 5. Утверждается, что если $P_m(t)$ — многочлен степени $m$ такой, что $|P_m(t)| \leqq 1$ при $-1 \leqq t \leqq 1$, то $|P_m'(t)| \leqq m^2$, и равенство имеет место только при $P_m(t) = \gamma T_m(t)$, $|\gamma| = 1$ и $t = \pm 1$. Чтобы проверить этот результат, положим $P_m(\cos\theta) = g_m(\theta)$. Из (5) следует, что $|(\sin\theta) P_m'(\cos\theta)| \leqslant m$, откуда видно, что $P_m'(t)/m$ удовлетворяет условиям (8.9). Следовательно,

$$\sqrt{1 - t^2} \left| \frac{P_m'(t)}{m} \right| \leqq 1,$$

и из (8.10) следует $|P_m'(t)| \leqq m^2$. То, что равенство возникает только в указанных случаях, выводится из (2).

B. Во всех иллюстрациях, представленных в части A, предполагалось, что функции $p$ и $q$ (см. § 5) есть в точности $p(t) \equiv 1$ и $q(t) \equiv 0$. При другом определении этих функций получаются новые неравенства. Например, мы можем снова положить, что $q(t) \equiv 0$ и определить $p(t)$ равенством

$$m(m+\alpha+\beta+1)\,p(t)=m(m+\alpha+\beta+1)\{P_m^{(\alpha,\beta)}(t)\}^2+$$
$$+(1-t^2)\left\{\frac{d}{dt}P_m^{(\alpha,\beta)}(t)\right\}^2, \qquad (8.15)$$

где $P_m^{(\alpha,\beta)}(t)$ — многочлены Якоби, ортогональные на отрезке $[-1, 1]$ по отношению к весовой функции $\omega(t)=(1-t)^\alpha(1+t)^\beta$, $(\beta>-1, \alpha>-1)$ и нормированные условием $P_m^{(\alpha,\beta)}(1)=\binom{m+\alpha}{m}$.

Для того чтобы установить свойства монотонности $p(t)$, необходимые для применения теоремы 5.2, мы поступим следующим образом. Известно, что $y=P_m^{(\alpha,\beta)}(t)$ удовлетворяет дифференциальному уравнению (см. Сеге [1959], гл. 4)

$$(1-t^2)\,y''+[\beta-\alpha-(\alpha+\beta+2)\,t]\,y'+m(m+\alpha+\beta+1)\,y=0.$$

Используя это соотношение, получим

$$m(m+\alpha+\beta+1)\,p'(t)=2(\alpha-\beta+(\alpha+\beta+1)\,t)\left\{\frac{d}{dt}P_m^{(\alpha,\beta)}(t)\right\}^2, \qquad (8.16)$$

откуда ясно, что $p'(t)$ изменяет знак один раз в точке $t_0=(\beta-\alpha)/(\alpha+\beta+1)$. Отметим, что $-1<t_0<1$ в том и только в том случае, когда $(\alpha+1/2)(\beta+1/2)>0$.

При $\alpha=\beta=\lambda-1/2$ ультрасферические многочлены $P_m^{(\lambda)}(t)$ определяются выражением

$$P_m^{(\lambda)}(t)=\frac{\Gamma(\lambda+1/2)\,\Gamma(m+2\lambda)}{\Gamma(2\lambda)\,\Gamma(m+\lambda+1/2)}P_m^{(\lambda-1/2,\lambda-1/2)}(t), \qquad \lambda>-\frac{1}{2}.$$

В этом случае (8.16) сводится к

$$m(m+2\lambda)\,p'(t)=2\lambda t\left\{\frac{d}{dt}P_m^{(\lambda)}(t)\right\}^2.$$

Если $\lambda>0$, то $p'(t)>0$ при $0<t<1$ и $p'(t)<0$ при $-1<t<0$. Следовательно, функция $f(t)=p(t)/(1-t^2)$ удовлетворяет условиям монотонности в теореме 5.2.

Применяя теорему 5.2 с функцией

$$m(m+2\lambda)\,p(t)=m(m+2\lambda)\{P_m^{(\lambda)}(t\}^2+(1-t^2)\left\{\frac{d}{dt}P_m^{(\lambda)}(t)\right\}^2,$$

мы получаем, что всякий многочлен $Q_{2m}(t)$ степени, не большей чем $2m$, подчиненный ограничениям $0\leq Q_{2m}(t)\leq m(m+2\lambda)\,p(t)$ и

$Q_{2m}(\pm 1)=0$, удовлетворяет также неравенству

$$\max_{-1\leqslant t\leqslant 1}\left\{\frac{Q_{2m}(t)}{1-t^2}\right\}\leqslant \max_{-1\leqslant t\leqslant 1}\left\{\frac{d}{dt}P_m^{(\lambda)}(t)\right\}^2 \qquad (8.17)$$

и равенство достигается только при $Q_{2m}(t)=(1-t^2)\{(d/dt)P_m^{(\lambda)}(t)\}^2$.

Для функции $p(t)$, определенной в (8.15), теоремы 5.1 и 5.3 применимы при всех $\alpha>-1$, $\beta>-1$.

С. В этой части описываются некоторые дальнейшие экстремальные свойства многочленов $\bar{u}$ и $\underline{u}$, связанные с теоремой представления (теорема 10.2 гл. II). Доказательства снова опираются на осцилляционные свойства $\bar{u}$ и $\underline{u}$ и обычный метод подсчета соответствующих нулей.

Полученные результаты в применении к случаю $u_i(t)=t^i$ $(i=0, 1, \ldots, n)$ дают дополнительную экстремальную характеризацию классических многочленов Чебышева.

Предположим, что $\{u_i\}_0^n$ — $T$-система на интервале $[\alpha, \beta]$. Пусть $p(t)$ и $q(t)$ — непрерывные на замкнутом подынтервале $[a, b]\subset[\alpha, \beta]$ функции такие, что для некоторого многочлена $u(t)$ имеем $p(t)>u(t)>q(t)$, $t\in[a, b]$. Будем рассматривать систему $\{u_i\}_0^n$ как $T$-систему на меньшем интервале $[a, b]$, и пусть $\underline{u}$ и $\bar{u}$ — специальные функции, описанные в теореме 10.2 гл. II и осциллирующие между $p(t)$ и $q(t)$ на интервале $[a, b]$. Следует подчеркнуть, что $\underline{u}$ и $\bar{u}$ зависят от подынтервала $[a, b]$.

Теорема 8.1. *Пусть $\underline{u}$ и $\bar{u}$ обозначают многочлены, описанные выше. Всякий многочлен $u(t)$, отличный от $\underline{u}$ и $\bar{u}$ и удовлетворяющий неравенствам*

$$q(t)\leqq u(t)\leqq p(t), \qquad t\in[a, b],$$

*удовлетворяет также неравенствам*

$$\bar{u}(t)<u(t)<\underline{u}(t), \qquad t\in[b, \beta]. \qquad (8.18)$$

*Если $t\in[\alpha, a)$, то* (8.18) *выполняется для четных $n$, а для нечетных $n$ неравенства заменяются на противоположные.*

*Другими словами, всякий многочлен, лежащий между $p$ и $q$ на меньшем интервале $[a, b]$, заключен строго между $\bar{u}$ и $\underline{u}$ вне этого интервала.*

Доказательство. Из осцилляционных свойств $\underline{u}$ и $\bar{u}$ заключаем, что многочлены $\bar{u}(t)-u(t)$ и $\underline{u}(t)-u(t)$ имеют каждый по $n$ нулей на интервале $[a, b]$ при соглашении, что неузловые нули считаются дважды. Предположим, что $u(t_0)\geqslant\underline{u}(t_0)$ при некотором $t_0\in[b, \beta]$; тогда разность $\underline{u}(t)-u(t)$ имеет $n+1$ нуль на $[\alpha, \beta]$, где неузловые нули учтены дважды. Но $\{u_i\}_0^n$ — $T$-система на $[\alpha, \beta]$ и следовательно, $\underline{u}(t) \equiv u(t)$. Остальные случаи аналогичны.

Применяя предшествующие рассуждения для случая, когда $p(t) = 1$, $q(t) = -1$, $[a, b] = [-1, 1]$, $u_i(t) = t^i$, $i=0, 1, \ldots, n$, и $\underline{u}(t) = T_n(t)$, получим следующие два простых утверждения:

(1) Если $P_n(t)$ — многочлен степени $n$ такой, что

$$P_n(t) \leq 1 \text{ при } -1 \leqslant t \leq 1,$$

то

$$|P_n(t_0)| \leqslant |T_n(t_0)| \text{ при } |t_0| > 1.$$

(2) Если $|P_n(t)| \leqslant 1$ при $t \in [-1, 1]$, то коэффициент $a_n$ при $t^n$ многочлена $P_n(t)$ подчиняется неравенству

$$|a_n| \leq 2^{n-1},$$

причем равенство возникает в том и только в том случае, когда $P_n(t) = T_n(t)$.

Тем же методом, что и в теореме 8.1, можно провести следующее исследование. Пусть $\mathcal{P}_n$ обозначает класс многочленов $P_n$ степени, не большей $n$, удовлетворяющих условиям

$$|P_n(t)| \leqslant 1 \text{ при } t \in [-1, 1],$$

$$P_n(-1) = a, \text{ где } 0 \leq a < 1.$$

Мы хотим найти многочлен $P_n \in \mathcal{P}_n$, для которого достигается

$$\max |P_n(t_0)|, \text{ где } t_0 > 1.$$

Чтобы решить эту задачу, положим, что $s_1$ — наименьшее из $t \in (-1, 1)$, для которого $T_n(s_1) = \pm 1$. Так как $T_n(t)$ пробегает интервал $[-1, 1]$ при $t$, изменяющемся от $-1$ до $s_1$, то существует точка $s_0 \in (-1, s_1)$ такая, что

$$T_n(s_0) = \begin{cases} a, & n = 2m, \\ -a, & n = 2m + 1. \end{cases}$$

Отметим, что $T_n(t)$ убывает или возрастает на $[-1, s_1]$ соответственно при $n$ четном и нечетном. Далее линейное преобразование

$$y(t) = \frac{2t - 1 - s_0}{1 - s_0} \quad \left(t(y) = \frac{(1 - s_0)y + 1 + s_0}{2}\right)$$

удовлетворяет равенствам $y(s_0) = -1$ и $y(1) = 1$, и поэтому многочлен

$$S_n(y) = T_n\left(\frac{(1 - s_0)y + 1 + s_0}{2}\right)$$

при $y \in [(-3 - s_0)/(1 - s_0), 1]$ имеет те же самые осцилляционные свойства, что и $T_n(t)$ на $[-1, 1]$. В случае $n = 2m$, $T_n(s_1) = -1$

и $T_n(s_0) = a$. Если теперь $|P_n(t_0)| \geqslant |S_n(t_0)|$, то $P_n(t_0) \geqslant S_n(t_0)$ или $-P_n(t_0) \geqslant S_n(t_0)$. В первом случае $P_n(-1) = a = T_n(s_0)$, и поэтому $P_n(t) - S_n(t)$ имеет $n$ нулей на $[y(s_1), t_0]$ и еще один нуль в точке $-1$. Поэтому $P_n(t) \equiv S_n(t)$. Если $-P_n(t_0) \geqslant S_n(t_0)$, то $P_n(t) + S_n(t)$ имеет $n$ нулей на $[y(s_1), t_0]$ и еще один нуль на $(-1, y(s_1))$ благодаря тому, что $-P_n(-1) = -a$, откуда вытекает, что график $P_n(t)$ пересекает график $S_n(t)$ для некоторого $t \in \in (-1, y(s_1))$. В случае $n = 2m + 1$ многочлен $T_n(t)$ возрастает при $t \in [-1, s_1]$, а $s_0$ определено так, что $T_n(s_0) = -a$ и доказательство такое же, как в четном случае. Таким образом, доказано следующее утверждение.

Теорема 8.2. *Пусть $P_n$ — многочлен степени не более $n$ такой, что*

$$|P_n(t)| \leq 1 \text{ при } |t| \leq 1 \text{ и } P_n(-1) = a \qquad (0 \leq a \leq 1).$$

*Тогда при $t_0 > 1$*

$$|P_n(t_0)| \leq \left| T_n\left(\frac{1-s_0}{2} t_0 + \frac{1+s_0}{2}\right) \right|, \tag{8.19}$$

*где $s_0$ — наименьшее значение $t$ в $(-1, 1)$, для которого*

$$T_n(s_0) = \begin{cases} a, & n = 2m, \\ -a, & n = 2m+1. \end{cases}$$

*Равенство возникает в (8.19) в том и только в том случае, когда*

$$P_n(t) = \begin{cases} T_n\left(\frac{1-s_0}{2} t + \frac{1+s_0}{2}\right), & n = 2m, \\ -T_n\left(\frac{1-s_0}{2} t + \frac{1+s_0}{2}\right), & n = 2m+1. \end{cases}$$

Теорема 8.2 была впервые получена Ривлиным и Шапиро [1961].

## § 9. Дополнения

В этом параграфе мы опишем некоторые дальнейшие приложения теорем представления для положительных многочленов.

Предположим, что $\{u_i\}_0^n$ — $ET$-система на $[a, b]$. Пусть $L$ и $L_1$ обозначают линейные функционалы, определенные на классе $u$-многочленов. Будем предполагать далее, что $L_1$ положителен, т. е. $L_1(u) > 0$ для всякого неотрицательного нетривиального многочлена $u(t)$. Требуется найти максимум и минимум выражения

$$R(u) = \frac{L(u)}{L_1(u)} \tag{9.1}$$

над классом $\mathcal{P}$ неотрицательных нетривиальных многочленов.

Теорема 9.1. *Если* $\{u_i\}_0^n$ — *ET-система, то экстремальные значения* (9.1) *над классом* $\mathcal{P}$ *достигаются для многочленов, имеющих* $n$ *нулей с учетом кратности.*

Доказательство. Отметим сначала, что экстремальные значения (9.1) конечны и действительно достигаются. В оставшейся части доказательства будем рассматривать только случай максимума (9.1). Очевидно, что задача максимизации $R(u)$ над классом $\mathcal{P}$ эквивалентна задаче максимизации $L(u)$ по отношению к множеству неотрицательных многочленов, подчиненных условию нормировки $L_1(u) = 1$.

Достаточно поэтому установить, что для всякого многочлена $u$, удовлетворяющего равенству $L_1(u) = 1$, существует многочлен $\tilde{u}$, имеющий $n$ нулей с учетом кратности, такой, что $L_1(\tilde{u}) = 1$ и $L(\tilde{u}) \geqslant L(u)$.

Предположим, что $L_1(u) = 1$ и что многочлен $u$ имеет не более $n - 1$ нулей с учетом кратности. В этом случае теорема 10.3 гл. II утверждает существование представления

$$u(t) = \underline{u}(t) + \bar{u}(t),$$

где многочлены $\underline{u}$ и $\bar{u}$ различны, неотрицательны и имеют каждый по $n$ нулей с учетом кратности. Пусть $\alpha = L_1(\underline{u})$ и $\beta = L_1(\bar{u})$. Тогда $\alpha > 0$, $\beta > 0$ и $\alpha + \beta = 1$, так как $L_1(u) = 1$ по предположению. Очевидно, что мы можем записать

$$u(t) = \alpha \frac{\underline{u}(t)}{\alpha} + \beta \frac{\bar{u}(t)}{\beta}.$$

Следовательно,

$$L(u) \leq \max\left\{\frac{L(\underline{u})}{\alpha}, \quad \frac{L(\bar{u})}{\beta}\right\},$$

откуда следует требуемый результат.

Получим более точные результаты для случая, когда $u_i(t) = t^i$, $i = 0, 1, \ldots, n$, а $L$ и $L_1$ имеют вид

$$L(P) = \int_a^b P(t)\, t\, d\sigma(t), \quad L_1(P) = \int_a^b P(t)\, d\sigma(t).$$

Мера $\sigma$ порождает моментную точку

$$c_i = \int_a^b t^i d\sigma, \qquad i = 0, 1, \ldots, n + 1, \tag{9.2}$$

относительно которой предположим, что она содержится во внутренности пространства моментов, связанного с $T$-системой $\{t^i\}_0^{n+1}$. (Обратим внимание на то, что к системе добавлена функция $t^{n+1}$.)

Если $n = 2m$, то теорема 9.1 утверждает, что максимум выражения

$$R(P) = \frac{\int_a^b P(t)\, t\, d\sigma}{\int_a^b P(t)\, d\sigma},$$

взятый по отношению к классу нетривиал ных неотрицательных многочленов $P(t) = \sum_{i=0}^{n} a_i t^i$, достигается для многочлена вида $P_m^2(t)$ или $(t-a)(b-t)P_{m-1}^2(t)$. Здесь $P_k$ обозначает многочлен степени $k$.

Пусть $\bar{\sigma}$ и $\underline{\sigma}$ — верхнее и нижнее главные представления моментной точки (9.2), веса и корни которых о означаются соответственно $\{s_k^*, \lambda_k\}_1^{m+2}$ и $\{t_k^*, \mu_k\}_1^{m+1}$, где $a = s_1^* < t_1^* < s_2^* < \ldots < s_{m+1}^* < t_{m+1}^* < s_{m+2}^* = b$. Тогда

$$\frac{\int_a^b P_m^2(t)\, t\, d\sigma(t)}{\int_a^b P_m^2(t)\, d\sigma} = \frac{\int_a^b t P_m^2(t)\, d\underline{\sigma}(t)}{\int_a^b P_m^2(t)\, d\underline{\sigma}(t)} = \frac{\sum_{k=1}^{m+1} t_k^* P_m^2(t_k^*)\mu_k}{\sum_{k=1}^{m+1} P_m^2(t_k^*)\mu_k} \leqslant t_{m+1}^*.$$

Точно так же, используя $\bar{\sigma}$, получаем

$$\frac{\int_a^b t(t-a)(b-t)P_{m-1}^2(t)\, d\sigma(t)}{\int_a^b (t-a)(b-t)P_{m-1}^2(t)\, d\sigma(t)} \leqslant s_{m-1}^*.$$

Так как $t_{m+1}^* > s_{m+1}^*$, то имеем

$$\frac{\int_a^b P(t)\, t\, d\sigma(t)}{\int_a^b P(t)\, d\sigma(t)} \leqslant t_{m+1}^*$$

для всякого неотрицательного нетривиального многочлена степени не более $n = 2m$.

Аналогично можно показать, что

$$\frac{\int_a^b P(t)\, t\, d\sigma(t)}{\int_a^b P(t)\, d\sigma(t)} \geqslant t_1^*.$$

Если $n = 2m + 1$, то обозначим корни и веса верхнего и нижнего главных представлений $\bar{\sigma}$ и $\underline{\sigma}$ для $(c_0, c_1, \ldots, c_{2m+2})$ через $\{s_k^*, \lambda_k\}_1^{m+2}$ и $\{t_k^*, \mu_k\}_1^{m+2}$, где $a = t_1^* < s_1^* < t_2^* < \ldots < t_{m+2}^* < s_{m+2}^* = b$. В этом случае получим

$$s_1^* \leqslant \frac{\int\limits_a^b P(t)\, t\, d\sigma(t)}{\int\limits_a^b P(t)\, d\sigma(t)} \leqslant t_{m+2}^*.$$

В следующей теореме $\{P_k\}$, $\{\overline{Q}_k\}$ и $\{\underline{Q}_k\}$ — многочлены, ортонормальные по отношению к $d\sigma$, $(b-t)\,d\sigma$ и $(t-a)\,d\sigma$ соответственно.

Теорема 9.2. *Пусть $P(t)$ — неотрицательный многочлен на $[a, b]$ степени не более $m$ $(P(t) \not\equiv 0)$, и пусть $c$ — положительное число, а $d$ вещественно.*

(1) *Если $n = 2m$, то*

$$ct_1^* + d \leqslant \frac{\int\limits_a^b P(t)(ct+d)\, d\sigma(t)}{\int\limits_a^b P(t)\, d\sigma(t)} \leqslant ct_{m+1}^* + d, \tag{9.3}$$

*где $t_1^*$ и $t_{m+1}^*$ — наименьший и наибольший корни $\underline{P}_{m+1}$.*

(2) *Если $n = 2m + 1$, то*

$$cs_1^* + d \leqslant \frac{\int\limits_a^b P(t)(ct+d)\, d\sigma(t)}{\int\limits_a^b P(t)\, d\sigma(t)} \leqslant ct_{m+2}^* + d, \tag{9.4}$$

*где $s_1^*$ — наименьший корень $\overline{Q}_{m+1}$, а $t_{m+2}^*$ — наибольший корень $\underline{Q}_{m+1}$. Границы, указанные в (9.3) и (9.4), точны.*

Доказательство. Единственная часть утверждения теоремы, которая еще не доказана, состоит в отождествлении значений $t_1^*$, $t_{m+1}^*$, $s_1^*$ и $t_{m+2}^*$. По этому поводу мы отсылаем читателя к лемме 2.2 гл. IV.

Для доказательства точности границ (9.3) и (9.4) достаточно определить $P(t)$ как многочлен, обращающийся в нуль в подходящих корнях мер $\underline{\sigma}$ или $\bar{\sigma}$.

Точно так же доказываются аналоги теоремы 9.2 для других особых случаев. Например, положим, что $a = 0$, $b > 0$, $L(u) = \int\limits_a^b u(t)\, t^2 d\sigma(t)$ и $n = 2m$. Корни и веса верхнего и нижнего главных представлений для

$$c_k = \int_a^b t^k d\sigma(t), \qquad k = 0, \dots, 2m + 2,$$

обозначаются через $\{s_k^*, \lambda_k\}_1^{m+2}$ и $\{t_k^*, \mu_k\}_1^{m+2}$, где $0 = t_1^* < s_1^* < t_2^* < \dots < t_{m+2}^* < s_{m+2}^* = b$. Тогда

$$\frac{\int_0^b t^2 P_m^2(t)\, d\sigma(t)}{\int_0^b P_m^2(t)\, d\sigma(t)} \leqslant [t_{m+2}^*]^2.$$

Кроме того,

$$\frac{\int_0^b t(b-t)\, P_{m-1}^2(t)\, t^2 d\sigma(t)}{\int_0^b t(b-t)\, P_{m-1}^2(t)\, d\sigma(t)} \leqslant [s_{m+1}^*]^2.$$

Следовательно,

$$\frac{\int_0^b P(t)\, t^2 d\sigma(t)}{\int_0^b P(t)\, d\sigma(t)} \leqslant [t_{m+2}^*]^2, \tag{9.5}$$

где $t_{m+2}^*$ — наибольший корень $Q_{m+1}$. Здесь $\{Q_k\}$ — многочлены, ортонормальные на $[0, b]$ по отношению к $t d\sigma(t)$.

Граница в (9.5) не обязательно точна.

Приложение. В качестве приложения теоремы 9.2 рассмотрим следующую задачу. Пусть $P(t)$ — произвольный неотрицательный многочлен степени $N$ на $[-1, 1]$. Запишем

$$P(t) = \sum_0^N a_n P_n(t),$$

где $P_n(t)$ — многочлены Лежандра на $[-1, 1]$, нормированные так, что $P_n(1) = 1$. Мы хотим определить границы для $a_2/a_0$. Так как $P_2(t) = (3t^2 - 1)/2$ и $g(\tau) = g(t^2) = 2^{-1}(P(t) + P(-t))$, $\tau = t^2$, — многочлен от $\tau$ степени $[N/2]$, то имеем

$$a_0 = \frac{1}{2}\int_{-1}^1 P(t)\, dt = \frac{1}{2}\int_0^1 g(\tau)\, \tau^{-1/2} d\tau,$$

$$a_2 = \frac{5}{2}\int_{-1}^1 P(t)\, P_2(t)\, dt = \frac{5}{2}\int_0^1 g(\tau)\, \frac{3\tau - 1}{2}\, \tau^{-1/2} d\tau.$$

Следовательно,

$$\frac{a_2}{a_0} = 5\int_0^1 g(\tau)\frac{3\tau-1}{2}\tau^{-1/2}d\tau \Bigg/ \int_0^1 g(\tau)\tau^{-1/2}d\tau,$$

где $g(\tau)$ — произвольный неотрицательный многочлен на $[0, 1]$ степени $[N/2]$.

Поэтому, если $[N/2] = 2m$, то теорема 9.2 утверждает, что

$$\frac{3t_1^*-1}{2} \leqslant \frac{a_2}{5a_0} \leqslant \frac{3t_{m+1}^*-1}{2},$$

где $t_1^*$ и $t_{m+1}^*$ — наибольший и наименьший корни многочлена $\underline{P}_{m+1}$, а если $[N/2] = 2m+1$, то

$$\frac{3s_1^*-1}{2} \leqslant \frac{a_2}{5a_0} \leqslant \frac{3t_{m+2}^*-1}{2},$$

где $t_{m+2}^*$ — наибольший корень $\underline{Q}_{m+1}$ и $s_1^*$ — наименьший корень $\overline{Q}_{m+1}$. Здесь $\underline{P}_k, \underline{Q}_k, \overline{Q}_k$ обозначают многочлены, ортогональные на $[0, 1]$ по отношению к $\tau^{-1/2}d\tau$, $\tau^{1/2}d\tau$ и $(1-\tau)\tau^{-1/2}d\tau$.

Переходя на интервал $[-1, 1]$ и учитывая Сеге [1959] (стр. 58), найдем, что

$$\underline{P}_{m+1}(t^2) = P_{m+1}^{(0,-1/2)}(2t^2-1) = P_{2m+2}(t),$$

$$\underline{Q}_{m+1}(t^2) = P_{m+1}^{(0,1/2)}(2t^2-1) = t^{-1}P_{2m+3}(t),$$

$$\overline{Q}_{m+1}(t^2) = P_{m+1}^{(1,-1/2)}(2t^2-1) = \frac{m+1}{2m+1}P_{2m+2}^{(1,1)}(t),$$

где $P_n^{(\alpha,\beta)}$ — многочлены Якоби, ортогональные на $[-1, 1]$ по отношению к $(1-t)^\alpha(1+t)^\beta$ и нормированные условием

$$P_n^{(\alpha+\beta)}(1) = \binom{n+\alpha}{n} \quad \text{и} \quad P_n(t) = P_n^{(2,2)}(t).$$

Примеры этого параграфа взяты у Сеге [1959] (стр. 186) и [1962].

## § 10. Периодические $ET$-системы и экстремальные задачи с двумя ограничениями

Пусть $L(u)$ — линейный функционал, где $u$ пробегает множество всех неотрицательных многочленов, подчиненных единственному линейному ограничению $L_1(u) = c_1$ (где $L_1$ — положительный линейный функционал). Мы указывали в § 9, что для обычной $ET$-системы на интервале $[a, b]$ экстремальное значение $L(u)$ достигается для многочлена, все нули которого вещественны. В дальнейшем мы будем часто писать $Lu$ вместо $L(u)$ и $L_1u$ вместо $L_1(u)$ для простоты обозначений.

Если $\{u_i\}_0^{2m}$ — периодическая система, то мы покажем, что тот же самый результат верен при наложении двух линейных ограничений. Это есть обобщение теоремы Бернштейна [1950] (стр. 472), которая устанавливает этот факт в специальном случае тригонометрических многочленов (см. также Григорьева [1962]). Главным при доказательстве теоремы 10.1 является применимость теоремы представления (теорем 6.3 гл.VI) с однопараметрическим семейством специальных многочленов, имеющих полное множество нулей. Следует подчеркнуть, что при наложении более чем двух ограничений могут не существовать экстремальные многочлены, все нули которых вещественны. В случае обычных $ET$-систем на интервале рассматриваемое утверждение, вообще говоря, неверно даже при двух ограничениях.

Будем рассматривать периодическую $ET$-систему $\{u_i\}_0^{2m}$ на $[a, b)$. Обозначим через $\mathcal{R}(c_1, c_2)$ класс неотрицательных многочленов $u = \sum_{i=0}^{2m} a_i u_i$, удовлетворяющих линейным ограничениям $L_1(u) = c_1$ и $L_2(u) = c_2$. Далее (теорема 10.1) мы покажем при некоторых слабых ограничениях на $L_1, L_2$ и $L$, что экстремальные значения $L(u)$, $u \in \mathcal{R}(c_1, c_2)$, достигаются для многочленов, имеющих $2m$ нулей.

Если наложено одно ограничение $L_1(u) = c_1$, то, как ранее отмечалось, метод теоремы 9.1 применим и в периодическом случае. Именно, если $\{u_i\}_0^{2m}$ — периодическая $ET$-система, то минимум над классом неотрицательных многочленов, удовлетворяющих $L_1(u) = c_1$ (для положительных $L_1$), достигается для многочлена, имеющего $n = 2m$ нулей с учетом кратности. В случае периодических $ET$-систем мы имеем следующее обобщение теоремы 9.1.

Теорема 10.1. *Пусть $\{u_i\}_0^{2m}$ — периодическая $ET$-система на $[a, b)$, а $L_1$, $L_2$ и $L$ — линейные функционалы, определенные при $u \in \mathcal{P}_{2m+1}$ (класс всех неотрицательных многочленов). Пусть $\mathcal{R}(c_1, c_2)$ обозначает класс многочленов $u \in \mathcal{P}_{2m+1}$, удовлетворяющих условиям $L_1 u = c_1$, $L_2 u = c_2$, и пусть $\mathcal{C}(c_1, c_2)$ состоит из многочленов из $\mathcal{R}(c_1, c_2)$, имеющих $2m$ нулей с учетом кратности.*

(а) *Если $L_1$ или $L_2$ — положительный линейный функционал, то*

$$\inf_{\mathcal{C}(c_1,c_2)} Lu \leqslant \inf_{\mathcal{R}(c_1,c_2)} Lu \leqslant \sup_{\mathcal{R}(c_1,c_2)} Lu \leqslant \sup_{\mathcal{C}(c_1,c_2)} Lu.$$

(б) *Если $L$ — положительный функционал, то*

$$\inf_{\mathcal{C}(c_1,c_2)} Lu \leqslant \inf_{\mathcal{R}(c_1,c_2)} Lu.$$

Доказательство. Для того чтобы установить предложение (а), достаточно показать, что для всякого многочлена $u \in \mathcal{R}(c_1, c_2)$ такого, что $u \notin \mathcal{C}(c_1, c_2)$, существуют два многочлена $v$ и $w$ в $\mathcal{C}(c_1, c_2)$, для которых $Lv \leq Lu \leq Lw$. Положим для определенности, что $L_1$ положительно.

Пусть $u$ содержится в $\mathcal{R}(c_1, c_2)$, но не в $\mathcal{C}(c_1, c_2)$. Для всякого $t^* \in [a, b)$ теорема 6.3 гл. VI утверждает существование единственного многочлена $u^*(t; t^*)$ со свойствами:

(1) $-u(t) \leqslant u^*(t; t^*) \leqslant u(t)$;

(2) $u^*(t^*; t^*) = 0$;

(3) $u(t) - u^*(t; t^*)$ и $u(t) + u^*(t; t^*)$ имеют каждый по $2m$ нулей, которые строго чередуются, если не рассматривать нули $u(t)$.

Из единственности $u^*(t; t^*)$ легко установить, что коэффициенты $u^*(t; t^*)$ — непрерывные функции $t^*$. Рассмотрим замкнутую кривую $C$ в двумерном пространстве, порожденном парами

$$\{L_1(u(t) + u^*(t; t^*)),\ L_2(u(t) + u^*(t; t^*))\}$$

при $t$, пробегающем $[a, b]$. Вследствие симметричного характера границ в (1) имеем для каждого $t^*$ число $t_1^*$ такое, что $u^*(t; t_1^*) = -u^*(t; t^*)$. Отсюда следует, что $C$ содержит точку $(c_1, c_2)$. Поэтому, если $(c_1, c_2) \neq (0, 0)$, то существует $t_0^*$, для которого

$$L_i(u(t) + u^*(t; t_0^*)) = (1 + \delta) c_i, \qquad i = 1, 2, \qquad \delta \geqslant 0. \quad (10.1)$$

Так как $-u(t) \leqslant u(t; t_0^*) \leq u(t)$, то положительность функционала $L_1$ влечет $-c_1 < \delta c_1 < c_1$ и, следовательно, $|\delta| < 1$. Утверждение (а) может быть легко получено применением $L$ к выпуклой комбинации

$$u(t) = \frac{(1-\delta)}{2}\left[\frac{u(t) - u^*(t; t_0^*)}{1-\delta}\right] + \frac{(1+\delta)}{2}\left[\frac{u(t) + u^*(t; t_0^*)}{1+\delta}\right]. \quad (10.2)$$

В самом деле, мы заключаем, на основе (1) и (3), что оба взятых в скобки многочлена принадлежат $\mathcal{C}(c_1, c_2)$, а значение $Lu$, очевидно, лежит между значениями $L[u(t) + u^*(t; t_0^*)]/(1 + \delta)$ и $L[u(t) - u^*(t; t_0^*)]/(1 - \delta)$.

В ситуации, когда $(c_1, c_2) = (0, 0)$, условие (1) влечет, как легко видеть, что $L_1(u(t; t^*)) = 0$ при всех $t^*$. Варьируя $t^*$, получим, что $L_2(u(t; t_0^*)) = 0$ для некоторого $t_0^*$. Полагая $\delta = 0$ в (10.2), получим снова требуемый результат. Этим завершается доказательство (а).

При проверке утверждения (б) предположим, что $(c_1, c_2) \neq (0, 0)$, так как в противном случае $v(t) \equiv 0$ принадлежит $\mathcal{C}(c_1, c_2)$ и $0 = Lv \leqslant Lu$. Действуя так же, как при доказательстве (а), получим (10.1), а затем рассмотрим два случая: $0 \leqslant \delta < 1$ и $\delta \geqslant 1$. Если $0 \leqslant \delta < 1$, то используем (10.2), как и ранее. Остается рассмотреть случай $\delta \geqslant 1$. Однако условие (1) эквивалентно $0 \leqslant u(t) + u^*(t; t_0^*) \leq 2u(t)$, и поэтому при $\delta \geqslant 1$ имеем

$$0 \leq \frac{u(t) + u^*(t; t_0^*)}{1+\delta} \leqslant \frac{2u}{1+\delta} \leq u.$$

В этом случае $v(t) = (1 + \delta)^{-1}[u(t) + u^*(t; t_0^*)]$ принадлежит $\mathcal{C}(c_1, c_2)$ и, очевидно, $Lv \leq Lu$. Это завершает доказательство теоремы.

Глава X

# НЕКОТОРЫЕ ЗАДАЧИ НАИЛУЧШЕЙ ИНТЕРПОЛЯЦИИ, МАКСИМИЗАЦИЯ ОПРЕДЕЛИТЕЛЕЙ, СВЯЗАННЫХ С МОМЕНТАМИ, И ПРИЛОЖЕНИЯ К ТЕОРИИ ПЛАНИРОВАНИЯ ЭКСПЕРИМЕНТА

## § 1. Введение

Мы начинаем эту главу с постановки трех задач теории аппроксимации. Эти задачи имеют различное происхождение и их формулировки кажутся несвязанными, однако точные решения в некоторых случаях совпадают. Общая постановка, объясняющая это явление, приводится в § 2. В дальнейшем мы существенно обобщим эти частные случаи. В конце главы формулируются некоторые нерешенные задачи, вытекающие из общей постановки.

Задача 1.1. Найти значения $x_0, x_1, \ldots, x_n$, удовлетворяющие неравенствам $a \leqslant x_0 < x_1 < \ldots < x_n \leq b$, которые максимизируют

$$T(x_0, x_1, \ldots, x_n) = \prod_{i=0}^{n} w(x_i) \prod_{0 \leqslant i < j \leqslant n} (x_i - x_j)^2. \tag{1.1}$$

Здесь $w(x)$ — фиксированная весовая функция.

При специальном выборе $w(x)$ величина $T(x_0, x_1, \ldots, x_n)$ связана с дискриминантной функцией обычных многочленов. Мы можем также интерпретировать выражение (1.1) как меру электростатической энергии, ассоциированной с $n+1$ зарядами, распределенными на интервале $[a, b]$. Именно, рассмотрим $n+1$ единичных «масс», расположенных в точках $x_0, x_1, \ldots, x_n$ на отрезке $[-1, 1]$, и фиксированные массы $p$ и $q$ в точках $+1$ и $-1$, соответственно. Единичные массы создают силы отталкивания в соответствии с законом логарифмического потенциала, и энергия системы оценивается как $\log T^{-1/2}$, где

$$T = T(x_0, x_1, \ldots, x_n) = \prod_{i=0}^{n} (1 - x_i)^{2p} (1 + x_i)^{2q} \prod_{0 \leqslant i < j \leqslant n} (x_i - x_j)^2. \tag{1.2}$$

Положение точек $x_0, \ldots, x_n$, для которого достигается максимум (1.2), соответствует условию электростатического равновесия. Выражение (1.2) есть специальный случай (1.1), где

$$w(x) = (1-x)^{2p}(1+x)^{2q} \qquad (p, q > 0).$$

Стилтьес доказал, что (1.2) достигает своего максимума, когда множество $\{x_i\}_0^n$ есть множество нулей ногочлена Якоби $P_{n+1}^{(\alpha, \beta)}(x)$, где $\alpha = 2p-1$, $\beta = 2q-1$.

Задача 1.2. Пусть $x_0, x_1, \ldots, x_n$ аданы на $[-1, 1]$. Многочлены $n$-й степени $l_\nu(x)$, $\nu = 0, 1, \ldots, n$, определенные условиями $l_\nu(x_i) = \delta_{i\nu}$, называются фундаментальными интерполяционными многочленами Лагранжа, порожденными точками $x_0, x_1, \ldots, x_n$. Эти многочлены имеют явный вид

$$l_\nu(x) = l(x)\,[l'(x_\nu)(x - x^\nu)]^{-1}, \tag{1.3}$$

где $l(x) = \prod\limits_{i=0}^{n}(x - x_i)$.

Отметим, что многочлен $\sum\limits_{\gamma=0}^{n} y_\nu l_\nu(x)$ интерполирует функцию со значениями $y_0, y_1, \ldots, y_n$ в точках $x_0, x_1, \ldots, x_n$.

Для данной весовой функции $w(x) > 0$, заданной на $[a, b]$, мы нормируем многочлены $l_\nu(x)$ так, что $w(x_i)\,r_\nu^2(x_i) = \delta_{i\nu}$ для $i, v = 0, 1, \ldots, n$, положив

$$r_\nu(x) = w^{-1/2}(x_\nu)\,l_\nu(x). \tag{1.4}$$

Задача состоит в отыскании $x_0, x_1, \ldots, x_n$, для которых

$$\sup_{a \leqslant x \leqslant b} w(x)\,\{r_0^2(x) + \ldots + r_n^2(x)\} \tag{1.5}$$

достигает минимума.

В случае, когда $w(x) \equiv 1$ и $[a, b] = [-1, 1]$, Фейер доказал, что минимум (1.5) достигается, когда $x_0 = -1$, $x_n = 1$ и $x_1, \ldots, x_{n-1}$ суть нули уравнения

$$P_n'(x) = 0, \tag{1.6}$$

где $P_n(x)$ есть многочлен Лежандра $n$-й степени.

Задача 1.3. Эта задача также касается интерполяционных систем для различных весовых функций $w(x)$. Пусть $x_0, x_1, \ldots, x_n$ суть $n+1$ точек в $[a, b]$, и пусть $p_\nu(x)$, $\nu = 0, 1, \ldots, n$, суть $n+1$ многочленов произвольной степени, удовлетворяющих условию

$$w(x_i)\,p_\nu(x_i) = \delta_{i\nu}, \qquad i, \nu = 0, 1, \ldots, n.$$

Функция $w(x)\sum\limits_{\nu=0}^{n} y_\nu p_\nu(x)$, очевидно, принимает значения $y_0, y_1, \ldots$

$\ldots, y_n$ в точках $x_0, x_1, \ldots, x_n$ соответственно. По этой причине система $\{w(x) p_\nu(x)\}_0^n$ называется интерполяционной системой, ассоциированной с $\{x_i\}_0^n$. Она может отличаться от системы интерполяционных многочленов Лагранжа тем, что $p_\nu$ не являются обязательно многочленами $n$-ой степени.

Система $\{w(x) p_\nu(x)\}_0^n$ называется *устойчивой*, если

$$0 \le w(x) \sum_{\nu=0}^{n} y_\nu p_\nu(x) \le \max_\nu y_\nu, \qquad a \le x \le b,$$

выполняется всякий раз, когда $y_\nu \geqslant 0$, $\nu = 0, 1, \ldots, n$.

Устойчивая интерполяционная система называется *наиболее экономичной* при условии, что величина $\sum_{\nu=0}^{n} [\deg p_\nu(x)]$ минимальна.

Желательно определить точки, ассоциированные с наиболее экономичной устойчивой интерполяционной системой. Для некоторых классических весовых функций мы докажем, что решение этой задачи совпадает с решением задачи 1.2.

Стилтьес получил решение задачи 1.1 в качестве новой характеризации классических ортогональных многочленов (см. Сеге [1959], стр. 139). Специальный случай задачи 1.2, когда $w(x) = 1$ был сформулирован и решен Фейером [1932]. Понятие и анализ задачи 1.3 восходит к Егервари и Турану [1958] и [1959].

## § 2. Общая теорема эквивалентности

В этом параграфе мы вводим общую постановку, которая включает в себя формулировки задач 1.1—1.3. При этом мы установим принцип эквивалентности, который утверждает, что решения этих задач совпадают.

Пусть $f(x) = (f_0(x), \ldots, f_n(x))$ — векторнозначная функция, составленная из $n+1$ линейно независимых непрерывных вещественных функций $f_0, \ldots, f_n$, определенных на компактном пространстве $X$. Пусть $\mathcal{D}$ обозначает множество всех мер $\xi$, определенных на борелевском поле $\mathcal{B}$ (порожденном открытыми множествами $X$), таких, что $\int_X \xi(dx) = 1$, т. е. множество всех вероятностных мер. Для всякого $\xi \in \mathcal{D}$ пусть $\mathbf{M}(\xi)$ обозначает матрицу

$$\mathbf{M}(\xi) = \| m_{ij} \|_{i,j=0}^{n}, \tag{2.1}$$

где

$$m_{ij} = \int_X f_i(x) f_j(x) \xi(dx), \qquad i, j = 0, 1, \ldots, n. \tag{2.2}$$

Допустим на протяжении этой главы, что $\mathcal{B}$ включает все одноточечные множества. В случае, когда мера $\xi_0$ сосредоточена в единственной точке $x_0$, матрица $\mathbf{M}(\xi)$ сводится к матрице $\mathbf{M}(\xi_0)$ единичного ранга с элементами $m_{ij} = f_i(x_0) f_j(x_0)$. Мы также выделим случай, когда $\xi$ приписывает массы $\lambda_1, \ldots, \lambda_p$ $p$ различным точкам $x_1, \ldots, x_p$ соответственно. В этом случае $\mathbf{M}(\xi)$ имеет вид

$$\sum_{i=1}^{p} \lambda_i [\mathbf{f}(x_i)][\mathbf{f}(x_i)]', \tag{2.3}$$

и элементы этой матрицы суть

$$m_{ij} = \sum_{l=1}^{p} \lambda_l f_i(x_l) f_j(x_l).$$

Пояснение к обозначениям. Символ $\mathbf{xy}'$ обозначает на протяжении этой главы матрицу единичного ранга, полученную обычным матричным умножением вектор-столбца $\mathbf{x}$ на вектор-строку $\mathbf{y}$. Аналогично $\mathbf{y}'\mathbf{x}$ обозначает скаляр $(\mathbf{x}, \mathbf{y})$. Под $\mathbf{Ax}$, где $\mathbf{A}$ — матрица, а $\mathbf{x}$ — вектор, понимается матричное произведение, в котором $\mathbf{x}$ берется как вектор-столбец. Символы $|\mathbf{A}|$ и $\mathbf{A}'$ будут, как обычно, использованы для определителя и транспонирования матрицы $\mathbf{A}$ соответственно.

В случае, когда $f_i(x) = x^i$, $i = 0, \ldots, n$, и $X = [a, b]$, матрицы $\mathbf{M}(\xi)$ сводятся к классическим генкелевым матрицам $\| c_{i+j} \|_{i,j=0}^{n}$, где

$$c_k = \int_a^b x^k \xi(dx), \qquad k = 0, 1, \ldots, 2n.$$

Отметим несколько простых свойств матриц $\mathbf{M}(\xi)$ в виде следующей леммы.

Лемма 2.1. *Пусть* $\mathbf{M}(\xi)$ *определено соотношением* (2.1). *Тогда*:

(1) *для всякого* $\xi$ *матрица* $\mathbf{M}(\xi)$ *положительно полуопределена*;

(2) $|\mathbf{M}(\xi)| = 0$ *всякий раз, когда спектр* $\xi$ *сосредоточен менее чем в* $n+1$ *точках*;

(3) *семейство матриц* $\mathbf{M}(\xi)$ *при* $\xi$ *из* $\mathcal{D}$ *образует выпуклое компактное множество*;

(4) *при любом* $\xi$ *матрица* $\mathbf{M}(\xi)$ *может быть записана в форме* (2.3), *где* $p \leq (n+1)(n+2)/2 + 1$.

Доказательство.

(1) При любом выборе $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ имеем

$$(\boldsymbol{\alpha}, \mathbf{M}(\xi)\boldsymbol{\alpha}) = \sum_{i,j=0}^{n} m_{ij}(\xi) \alpha_i \alpha_j = \int_X \left| \sum_{i=0}^{n} \alpha_i f_i(x) \right|^2 \xi(dx) \geqslant 0,$$

откуда видно, что матрица $\mathbf{M}(\xi)$ положительно полуопределена.

(2) При сделанном предположении ранг $\mathbf{M}(\xi)$ не превосходит $n$, и, следовательно, $\mathbf{M}(\xi)$ — особенная матрица.

(3) Утверждение о выпуклости очевидно.

Пусть $z_1(x), \ldots, z_m(x)$, $(m = (n+1)(n+2)/2)$ обозначает множество функций $f_i(x) f_j(x)$ $(j = 0, 1, \ldots, i;\ i = 0, 1, \ldots, n)$, упорядоченных некоторым способом. Рассмотрим множество из $R^m$, образованное координатными функциями, т. е.

$$C_m = \{\mathbf{z}(x) = (z_1(x), \ldots, z_m(x)) \mid x \in X\}.$$

Пусть $\mathscr{C}(C_m)$ обозначает выпуклую оболочку $C_m$. Классическая теорема Каратеодори утверждает (ср. примечание на стр. 50), что всякая точка $\mathbf{c} = (c_1, \ldots, c_m)$ из $\mathscr{C}$ имеет представление в виде

$$c_i = \sum_{j=1}^{m+1} \alpha_j z_i(t_j), \qquad i = 1, 2, \ldots, m, \tag{2.4}$$

где $\alpha_j \geqslant 0$, $j = 1, \ldots, m+1$ и $\sum\limits_{j=1}^{m+1} \alpha_j = 1$.

Так как функции $z_1, \ldots, z_m$ непрерывны, а $X$ компактно, мы выводим из (2.4), что $\mathscr{C}(C_m)$ — компакт. Кроме того, из замкнутости $\mathscr{C}(C_m)$ следует, что $\mathscr{C}(C_m) = \mathscr{M}_m$ (ср. теорему 2.1 гл. V), где

$$\mathscr{M}_m = \left\{(c_1, \ldots, c_m) \mid c_i = \int z_i(x)\,\xi(dx), \qquad i = 1, \ldots, m,\ \xi \in \mathscr{D}\right\}.$$

Компактность семейства матриц $\mathbf{M}(\xi)$ теперь очевидна.

(4) Это утверждение есть формулировка теоремы Каратеодори для данного случая.

Для каждого $\xi$ такого, что $|\mathbf{M}(\xi)| \neq 0$, введем функцию

$$d(x, \xi) = (f(x), \mathbf{M}^{-1}(\xi) f(x)). \tag{2.5}$$

Мы теперь можем перейти к формулировке двух основных задач, в известной мере подготовленных исследованиями § 1. В самом деле, соображения этого параграфа были подходящими для исследования некоторых статистических задач (см. § 7). В ходе этих исследований результаты, относящиеся к теории аппроксимации и интерполяции, получались в качестве побочных. Наше основное внимание сосредоточено на аналитических аспектах, относящихся к геометрии пространств моментов.

Задача 2.1. Определить меры $\xi \in \mathscr{D}$, которые максимизируют $|\mathbf{M}(\xi)|$.

Задача 2.2. Определить $\xi \in \mathscr{D}$, для которых $\sup\limits_x d(x, \xi)$ достигает минимума.

В §§ 3 и 4 мы показываем, что задачи 1.1 и 1.2, поставленные в § 1, в действительности вытекают из более общих задач, сформулированных выше. Оставшаяся часть настоящего параграфа будет посвящена доказательству того, что решения задач 2.1 и 2.2 совпадают.

Теорема 2.1 (Теорема эквивалентности). *Условия*:

(1) $\xi^*$ *максимизирует* $|\mathbf{M}(\xi)|$;

(2) $\xi^*$ *минимизирует* $\sup\limits_x d(x, \xi)$;

(3) $\sup\limits_x d(x, \xi^*) = n + 1$

*определяют одно и то же множество.*

*Множество* $B$, *состоящее из всех* $\xi$, *удовлетворяющих* (1), (2) *или* (3), *выпукло и замкнуто и матрица* $\mathbf{M}(\xi)$ *одна и та же для всех* $\xi \in B$.

Замечание 2.1. Для того чтобы облегчить доказательство теоремы 2.1, удобно использовать терминологию и понятия теории игр. Для наших целей игра определяется как тройка $(\Xi, \Theta, K)$, где $\Xi$ и $\Theta$ суть компактные выпуклые множества и $K = K(\xi, \theta)$ — непрерывная функция, определенная на $\Xi \times \Theta$, вогнутая по $\xi$ при фиксированном $\theta \in \Theta$ и выпуклая по $\theta$ при фиксированном $\xi \in \Xi$. На пространство $\Xi$ мы будем ссылаться как на пространство стратегий для игрока I и на $\Theta$ как на пространство стратегий игрока II. Когда игроками I и II определены стратегии $\xi$ и $\theta$ соответственно, значение ядра $K(\xi, \theta)$ интерпретируется как выплата игроку I, т. е. игрок II «платит» игроку I сумму $K(\xi, \theta)$. Основная цель игрока I состоит в максимизации его вознаграждения, тогда как действия игрока II руководятся противоположными мотивами.

Основная теорема теории игр утверждает, что

$$\max_{\xi \in \Xi} \min_{\theta \in \Theta} K(\xi, \theta) = v = \min_{\theta \in \Theta} \max_{\xi \in \Xi} K(\xi, \theta). \tag{2.6}$$

Кроме того, существуют $\xi_0 \in \Xi$ и $\theta_0 \in \Theta$ такие, что $K(\xi_0, \theta) \geqslant v$ для всех $\theta \in \Theta$ и $K(\xi, \theta_0) \leqslant v$ для всех $\xi \in \Xi$. Величина $v$, входящая в (2.6), называется ценой игры, а $\xi_0$ и $\theta_0$ — суть оптимальные стратегии для игрока I и игрока II соответственно.

Доказательство этой теоремы читатель может найти в книге Карлина [1959]. В настоящем применении пространства $\Xi$ и $\Theta$ отождествляются с пространством вероятностных мер на компактном пространстве $X$.

Замечание 2.2. Сформулированная теорема эквивалентности была открыта Кифером и Вольфовицем [1960] в ходе их исследований по оптимальным планам для построения регрессионной модели по экспериментальным данным. В ряде изящных и глубоких работ Кифер [1959], [1960], [1961], [1962a, b] и Кифер и Вольфовиц [1959] заложили основание статистической теории эксперимента. Родственный геометрический подход был выдвинут Элвингом [1952].

Доказательство теоремы 2.1. Рассмотрим игру с ядром

$$K_\varepsilon(\xi, \theta) = \operatorname{tr} \mathbf{M}^{-1}(\theta)\, \mathbf{M}(\xi), \tag{2.7}$$

где $\xi$ и $\theta$ пробегают множество

$$\Xi = \{\xi \mid \xi \in D, \text{ все собственные числа } \mathbf{M}(\xi) \text{ не меньше } \varepsilon\}. \tag{2.8}$$

Ядро $K_\varepsilon(\xi, \theta)$ линейно и, таким образом, вогнуто по $\xi$. Покажем теперь, что $K_\varepsilon(\xi, \theta)$ выпукло по $\theta$. Доказательство выпуклости может быть основано на известном факте, что $\mathbf{M}^{-1}(\theta)$, как матричная функция, выпукла, т. е. матрица

$$\lambda\mathbf{M}^{-1}(\theta_1) + (1-\lambda)\mathbf{M}^{-1}(\theta_2) - [\lambda\mathbf{M}(\theta_1) + (1-\lambda)\mathbf{M}(\theta_2)]^{-1}$$

положительно определена для $0 < \lambda < 1$ при условии, что $\mathbf{M}(\theta_1) \neq \mathbf{M}(\theta_2)$. Умножение слева и справа на $\mathbf{M}^{1/2}(\xi)$ ничего не изменяет. Остается заметить, что $\operatorname{tr}$ есть линейный функционал.

Из замечания 2.1 следует, что каждая из игр $\{\Xi_\varepsilon, \Xi_\varepsilon, K_\varepsilon\}$ имеет цену $v_\varepsilon$ и соответствующие оптимальные стратегии. Ясно, что для всякого $\xi$ $\inf_\theta \operatorname{tr}\mathbf{M}^{-1}(\theta)\mathbf{M}(\xi) \leqslant n+1$, так как можно взять $\theta = \xi$. Следовательно,

$$\sup_\xi \inf_\theta \operatorname{tr}\mathbf{M}^{-1}(\theta)\mathbf{M}(\xi) \leqslant n+1, \tag{2.9}$$

где нижняя и верхняя грани берутся по множеству $\Xi_\varepsilon$.

Применяя неравенство между средним арифметическим и средним геометрическим[1]), получим

$$\operatorname{tr}\mathbf{M}^{-1}(\theta)\mathbf{M}(\xi) \geqslant (n+1)\frac{|\mathbf{M}(\xi)|^{1/(n+1)}}{|\mathbf{M}(\xi)|^{1/(n+1)}}; \tag{2.10}$$

равенство достигается в том и только в том случае, когда $\mathbf{M}(\xi)$ пропорциональна $\mathbf{M}(\theta)$. Так как $\prod_{i=0}^{n}\lambda_i = |\mathbf{M}(\xi)|$, где $\lambda_i$ — собственные числа $\mathbf{M}(\xi)$ и

$$\max_i \lambda_i = \sup_{\mathbf{v}} \frac{(\mathbf{v}, \mathbf{M}\mathbf{v})}{(\mathbf{v}, \mathbf{v})} = \sup_{\mathbf{v}} \frac{\int \sum_{i,j=0}^{n+1} v_i v_j f_i f_j d\xi}{(\mathbf{v}, \mathbf{v})} =$$

$$= \sup_{\mathbf{v}} \frac{\int \left(\sum_0^{n+1} v_i f_i\right)^2 d\xi}{(\mathbf{v}, \mathbf{v})} \leqslant \int \sum_0^{n+1} f_i^2(x)\,\xi(dx),$$

то мы получаем, что матрицы $\mathbf{M}(\xi_0)$ для $\xi_0$, на которых достигается $\sup_\xi |\mathbf{M}(\xi)|$, с необходимостью имеют собственные числа, превосходящие некоторое $\varepsilon_0$.

---

[1]) В виде $(\operatorname{tr}\mathbf{P})/(n+1) \geqslant |\mathbf{P}|^{1/(n+1)}$ для положительно полуопределенных матриц $\mathbf{P}$, где равенство достигается тогда и только тогда, когда $\mathbf{P}$ есть единичная матрица, умноженная на константу.

При $\varepsilon \leqslant \varepsilon_0$ из (2.10 следует, что

$$\sup_{\xi \in \Xi_\varepsilon} \inf_{\theta \in \Xi_\varepsilon} \operatorname{tr} \mathbf{M}^{-1}(\theta) \mathbf{M}(\xi) \geqslant n+1. \tag{2.11}$$

Сравнивая (2.9) и (2.11), видим, что цена игры $(\Xi_\varepsilon, \Xi_\varepsilon, K_\varepsilon)$ равна $v_\varepsilon = n+1$ (при $\varepsilon \leqslant \varepsilon_0$).

Пусть $\Xi_0(\varepsilon)$ и $\Theta_0(\varepsilon)$ — классы оптимальных стратегий для игроков I и II соответственно. Если $\theta_0 \in \Theta_0(\varepsilon)$, т. е.

$$\operatorname{tr} \mathbf{M}^{-1}(\theta_0) \mathbf{M}(\varepsilon) \leqslant n+1 \tag{2.12}$$

для всех $\xi \in \Xi_\varepsilon$, то по (2.10) $|\mathbf{M}(\theta_0)| = \sup_\xi |\mathbf{M}(\xi)|$. Следовательно, $\Theta_0(\varepsilon) \subset B$, где $B = \{\xi \mid |\mathbf{M}(\xi)| = \sup_\eta |\mathbf{M}(\eta)|\}$. Далее из уравнения (2.10) следует, что если $\xi_0 \in B$, то

$$\operatorname{tr} \mathbf{M}^{-1}(\theta) \mathbf{M}(\xi_0) \geqslant n+1 \text{ для всех } \theta \in \Xi_\varepsilon,$$

так что $\xi_0 \in \Xi_0(\varepsilon)$ или $B \subset \Xi_0(\varepsilon)$. Таким образом, мы показали, что $\Theta_0(\varepsilon) \subset B \subset \Xi_0(\varepsilon)$.

Пусть теперь $\xi_0 \in B \subset \Xi_0(\varepsilon)$ и $\theta \in \Theta_0(\varepsilon)$. Оптимальность $\xi_0$ и $\theta_0$ требует соотношения $\operatorname{tr} \mathbf{M}^{-1}(\theta_0) \mathbf{M}(\xi) = n+1$. Более того, так как $\xi_0$ и $\theta_0$ принадлежат $B$, то $|\mathbf{M}(\theta_0)| = |\mathbf{M}(\xi_0)|$, так что из (2.10) следует

$$n+1 = \operatorname{tr} \mathbf{M}^{-1}(\theta_0) \mathbf{M}(\xi) \geqslant (n+1) \frac{|\mathbf{M}(\xi_0)|^{1/(n+1)}}{|\mathbf{M}(\theta_0)|^{1/(n+1)}} = n+1. \tag{2.13}$$

Но равенство достигается в (2.13), только когда $\mathbf{M}(\xi_0)$ пропорциональна $\mathbf{M}(\theta_0)$, а отсюда легко вытекает, что $\mathbf{M}(\xi_0) = \mathbf{M}(\theta_0)$. Зафиксировав $\xi_0$, мы выводим, что матрицы $\mathbf{M}(\theta_0)$, $\theta_0 \in \Theta(\varepsilon)$, все совпадают. Кроме того, $\xi_0$ выбирается произвольно в $B$, так что $\Theta_0(\varepsilon) = B$. Следовательно, множество матриц $B = \Theta_0(\varepsilon)$ выпукло и замкнуто.

Мы можем теперь доказать, что условия (1), (2) и (3) определяют одно и то же множество. Так как

$$\sup_x d(x, \xi^*) = \sup_\xi \operatorname{tr} \mathbf{M}^{-1}(\xi^*) \mathbf{M}(\xi) \geqslant n+1,$$

то (2) и (3) определяют множество $\Theta_0(\varepsilon)$. Кроме того, (1) есть в точности множество $B$, так что желаемый результат установлен.

Представляет интерес приводимое далее следствие из теоремы 2.1.

Следствие 2.1. Пусть $f_0, f_1, \ldots, f_n$ — линейно независимые непрерывные функции на компактном пространстве $X$. Тогда существует вероятностная мера $\xi^*$ на $X$ с конечным спектром и множество функций

$$g_i = \sum_{j=0}^{n} a_{ij} f_j, \qquad i = 0, 1, \ldots, n,$$

где матрица $\mathbf{A} = \|a_{ij}\|_{i,j=0}^{n}$ положительно определена, таких, что

(1) функции $g_0, g_1, \ldots, g_n$ ортонормальны по отношению к $\xi^*$, т. е. $\int\limits_X g_i g_j d\xi^* = \delta_{ij}, i, j = 0, 1, \ldots, n$;

(2) $\max\limits_x \sum\limits_{i=0}^{n} g_i^2(x) = \int\limits_X \sum\limits_{i=0}^{n} g_i^2 d\xi^* = n + 1.$

Доказательство. Выберем $\xi^* \in B$. По лемме 2.1 мера $\xi^*$ может быть выбрана так, что ее спектр конечен. Если положить $\mathbf{A} = \mathbf{M}^{-1/2}(\xi^*)$, то утверждения (1) и (2) следуют непосредственно из теоремы 2.1.

## § 3. Максимизация некоторых определителей

В этом параграфе наша цель будет состоять в нахождении явного решения задач 2.1 и 2.2 в некоторых специальных случаях. Осуществляя эту цель, мы обнаружим, что задача 2.1 наиболее доступна для непосредственного анализа. Мы не всегда можем представить решение в законченной форме, но некоторые характеризации удаются.

Более точно, мы хотим найти меры $\xi_0$, которые максимизируют определитель матрицы $\mathbf{M}(\xi) = \|m_{ij}\|$, где $m_{ij} = \int\limits_X f_i(x) f_j(x) \xi(dx)$, $i, j = 0, 1, \ldots, n$, при специальных выборах $f_0, f_1, \ldots, f_n$.

Положим, что мера $\xi$ сосредоточивает массы $\lambda_0, \lambda_1, \ldots, \lambda_n$ $\left(\sum\limits_{j=0}^{n} \lambda_j = 1\right)$ в $n+1$ различных точках $x_0, x_1, \ldots, x_n$ соответственно. При таком выборе $\xi$ элементы $\mathbf{M}(\xi)$ принимают вид

$$m_{ij} = \sum_{l=0}^{n} \lambda_l f_i(x_l) f_j(x_l). \tag{3.1}$$

Записывая $a_{il} = \lambda_l f_i(x_l)$ и $b_{jl} = f_j(x_l)$ $(i, j, l = 0, 1, \ldots, n)$, получаем отсюда, что детерминант $\mathbf{M}(\xi)$ имеет специальный вид

$$\left\{\prod_{l=0}^{n} \lambda_l\right\} \{\det \| f_i(x_j) \|_{i,j=0}^{n}\}^2. \tag{3.2}$$

Если функции $f_0, \ldots, f_n$ определены как

$$f_i(x) = x^i w^{1/2}(x), i = 0, 1, \ldots, n, x \in [a, b], \tag{3.3}$$

то (3.2) принимает вид

$$\left(\prod_{l=0}^{n} \lambda_l\right) \prod_{l=0}^{n} w(x_l) \prod_{0 \leqslant i < j \leqslant n} (x_i - x_j)^2. \tag{3.4}$$

Это выражение есть по существу функция $T(x_0, \ldots, x_n)$, определенная в (1.1).

Мы видим из (3.2), что два множества переменных $x_0, x_1, \ldots$ $\ldots, x_n$ и $\lambda_0, \lambda_1, \ldots, \lambda_n$ разделены. Максимизация (3.2), следовательно, может быть выполнена в две стадии. Простое вычисление показывает, что максимум $\prod_{l=0}^{n} \lambda_l$ при условиях $\lambda_l \geqslant 0$, $\sum_{l=0}^{n} \lambda_l = 1$, достигается при $\lambda_l = 1/(n+1)$, $l = 0, 1, \ldots, n$. В общем случае максимизация $\det \| f_i(x_j) \|$ или даже более простая задача характеризации решения оказывается чрезвычайно трудной.

Замечание 3.1. Из проведенного выше обсуждения вытекает ряд интересных задач. Например, было бы важно установить условия на функции $f_0, f_1, \ldots, f_n$, которые гарантируют, что максимум $|\mathbf{M}(\xi)|$ достигается для мер, сосредоточенных в $n+1$ точках. В этом случае $|\mathbf{M}(\xi)|$ упрощается до (3.2).

Важный специальный случай касается задачи характеризации весовых функций $w(x)$, для которых максимизирующая мера $\xi$, связанная с функциями $f_i(x) = x^i w^{1/2}(x)$, $i = 0, 1, \ldots, n$, содержит $n+1$ скачок.

Частичное решение этой задачи устанавливается ниже в теореме 3.6. Было бы важно также охарактеризовать точки $x_0, x_1, \ldots$ $\ldots, x_n$, которые дают максимум в (3.2) или (3.4).

Обратимся теперь к максимизации определителя $|\mathbf{M}(\xi)|$ при некоторых специальных выборах $f_0, \ldots, f_n$. Так как доказательства теорем 3.1 — 3.6 сходны, то мы формулируем все эти теоремы, а доказываем только теорему 3.4.

Теорема 3.1. *Если* $\mu_\nu(\xi) = \int_{-1}^{1} x^\nu \xi(dx)$, $\nu = 0, 1, \ldots, 2n$, *где* $\xi$ — *вероятностная мера на* $[-1, 1]$, *то*

$$\det [\, \| \mu_{i+j} \|_{i,j=0}^{n} ] \leqslant 2^{n(n+1)} \left( \prod_{1}^{n} \nu^\nu \right)^4 n^{-n} \prod_{1}^{2n} \nu^{-\nu},$$

*где равенство достигается тогда и только тогда, когда $\xi$ состоит из $n+1$ равных масс, помещенных в нулях $x_0, x_1, \ldots, x_n$ уравнения $(1-x^2) P_n'(x) = 0$, где $P_n(x)$ есть многочлен Лежандра $n$-й степени.*

Теорема 3.2. *Если* $J_\nu(\xi) = \int_{-1}^{1} x^\nu (1-x)^{\alpha+1} (1+x)^{\beta+1} \xi(dx)$, $\nu = 0, 1, \ldots, 2n$ $(\alpha > -1, \beta > -1)$, *то*

$$\det [\| J_{i+j} \|_{i,j=0}^{n}] \leqslant \frac{\prod_{1}^{n} \nu^\nu \prod_{1}^{n+1} (\nu+\alpha)^{\nu+\alpha} (\nu+\beta)^{\nu+\beta}}{\prod_{1}^{n+1} (\nu+n+1+\alpha+\beta)^{\nu+n+1+\alpha+\beta}} \, 2^{(n+1)(n+\alpha+\beta+2)},$$

*где равенство достигается тогда и только тогда, когда $\xi$ состоит из $n+1$ равных масс, помещенных в нулях многочлена Якоби $P_{n+1}^{(\alpha,\beta)}(x)$.*

Для многочленов Лагерра имеем следующий результат.

Теорема 3.3 *Если $l_\nu(\xi) = \int_0^\infty xe^{-x}\xi(dx)$, $\nu = 0, 1, \ldots, 2n$, то*

$$\det[\| l_{i+j}\|_{i,j=0}^n] \leqslant e^{-n(n+1)}(2^2 3^3 \ldots n^n)^2,$$

*где равенство достигается тогда и только тогда, когда $\xi$ состоит из $n+1$ равных масс, помещенных в $x_0 = 0$ и в $n$ нулях многочлена Лагерра $L_n^{(1)}(x)$ с ассоциированным параметром 1.*

Теорема 3.4. *Если $L_\nu(\xi) = \int_0^\infty x^{\nu+\alpha+1}e^{-x}\xi(dx)$ $(\alpha > -1)$, то*

$$\det[\| L_{i+j}\|_{i,j=0}^n] \leqslant e^{-(n+1)(n+1+\alpha)}(n+1+\alpha)^{n+1+\alpha}\prod_1^n \nu^\nu(\nu+\alpha)^{\nu+\alpha},$$

*где равенство достигается тогда и только тогда, когда $\xi$ состоит из $n+1$ равных масс, сосредоточенных в нулях $L_{n+1}^{(\alpha)}(x)$.*

Наконец, для нормальных весовых функций, ассоциированных с многочленами Эрмита, справедлива

Теорема 3.5. *Если $H_\nu(\xi) = \int_{-\infty}^\infty x^\nu e^{-x^2}\xi(dx)$, $\nu = 0, 1, \ldots, 2n$, то*

$$\det[\| H_{i+j}\|_{i,j=0}^n] \leqslant (2e)^{-[n(n+1)]/2}\prod_1^n \nu^\nu,$$

*где равенство достигается тогда и только тогда, когда $\xi$ состоит из $n+1$ равных масс, помещенных в нулях многочлена Эрмита $H_{n+1}(x)$.*

Доказательство теоремы 3.4. Положим $D(\xi) = \det\| L_{i+j}(\xi)\|$ и отметим, что $0 < \sup_\xi D(\xi) = D_0 < \infty$. Применяя классическую теорему выбора Хелли, мы выберем последовательность $\{\xi_n\}$ такую, что $\lim_{n\to\infty} D(\xi_n) = D_0$ и $\xi_n$ слабо сходится к $\xi_0$, которая априори может не быть вероятностной мерой, т. е. $\xi_0$ может иметь полную меру меньшую единицы. Так как каждая из функций $x^{\nu+\alpha+1}e^{-x}$, $\nu = 0, 1, \ldots, 2n$, обращается в нуль на бесконечности, то имеем $D(\xi_0) = D_0$, так что максимум достигается для $\xi_0$. Мы утверждаем, что $\xi_0$ не может помещать массу в начало координат, а также что полная мера $\xi_0(\infty)$ не может быть меньше единицы. Если мы примем, напротив, что $\xi_0(\infty) < 1$, то мера $\xi_0(dx)/\xi_0(\infty)$ дает определитель с большим чем $D_0$ значением. Возможность того, что $\xi_0$ сосредоточивает некоторую массу в нуле, исключается по аналогичной причине.

Пусть теперь $L_\nu(\xi_0) = L_\nu^0$. Так как $\max D(\xi)$ достигается при $\xi = \xi_0$, то максимум $L_{2n}(\xi)$, взятый по всем вероятностным мерам, первые $2n$ моментов которых равны $L_\nu^0$, $\nu = 0, \ldots, 2n-1$, достигается при $\xi = \xi_0$. В этом случае точка $(L_0^0, \ldots, L_{2n}^0)$ есть граничная точка выпуклого множества

$$\mathcal{M} = \left\{(L_0, \ldots, L_{2n}) \mid L_\nu = \int_0^\infty x^{\nu+\alpha+1} e^{-x} \xi(dx),\ \nu = 0, 1, \ldots, 2n\right\},$$

где $\xi$ пробегает множество всех вероятностных мер на $[0, \infty)$. Поэтому существует опорная гиперплоскость к $\mathcal{M}$ в $(L_0^0, \ldots, L_{2n}^0)$, т. е. существуют вещественные константы $a_0, \ldots, a_{2n}$ и $d$, $\sum_0^{2n} a_i^2 > 0$ такие, что $\sum_0^{2n} a_i L_i \leqslant d$ для всех $(L_0, \ldots, L_{2n}) \in \mathcal{M}$ и $\sum_0^{2n} a_i L_i^0 = d$. Эквивалентно $\int_0^\infty \left(\sum_{i=0}^{2n} a_i x^{i+\alpha+1} e^{-x} - d\right) \xi(dx) \leq 0$ для всех $\xi$, причем равенство достигается при $\xi = \xi_0$. Отсюда следует, что $\sum_{i=0}^{2n} a_i x^{i+\alpha+1} e^{-x} \leq d$ для всех $x \in [0, \infty)$ и равенство достигается во всех точках роста $\xi_0$. Докажем теперь, что равенство выполняется самое большее в $n+1$ точках, больших чем нуль. Если равенство достигается в $n+2$ точках, то функция $g(x) = e^{-x} x^{\alpha+1} P_{2n}(x) - d$, где $P_{2n}(x) = \sum_0^{2n} a_i x^i$, имеет по крайней мере $2n+4$ нуля с учетом кратности. По теореме Ролля мы заключаем, что $g'(x) = e^{-x} x^\alpha [-x P_{2n} + x P'_{2n} + (\alpha+1) P_{2n}]$ имеет по крайней мере $2n+3$ нуля на $(0, \infty)$.

Тогда должно быть

$$-x P_{2n}(x) + x P'_{2n}(x) + (\alpha+1) P_{2n}(x) \equiv 0$$

и, следовател но, $P_{2n}(x) \equiv 0$, что противоречит условию $\sum_{i=0}^{2n} a_i^2 > 0$. Так как $D(\xi_0) > 0$, то ясно, что $\xi_0$ имеет $n+1$ точку роста. Проделанный анализ устанавливает, что мера $\xi_0$ сосредоточена в точках $x_0, x_1, \ldots, x_n$, $x_i > 0$ $(i = 0, \ldots, n)$ с соответствующими весами $\lambda_0, \lambda_1, \ldots, \lambda_n$, $\sum_{i=0}^{n} \lambda_i = 1$.

Мы можем, следовательно, записать

$$L_\nu^{(0)} = \sum_{i=0}^{n} x_i^{\nu+\alpha+1} e^{-x_i} \lambda_i. \qquad (3.5)$$

Применение основной композиционной формулы (3.12) гл. I к (3.5) дает выражение

$$D(\xi_0) = \left(\prod_0^n \lambda_i\right) e^{-\Sigma x_i} \prod_0^n x_i^{\alpha+1} \prod_{0 \leqslant i < j \leqslant n} (x_i - x_j)^2. \qquad (3.6)$$

Теперь можно действовать, как при доказательстве теорем 6.7.1 — 6.7.3 у Сеге [1959] (стр. 140). В данном случае мы можем продифференцировать $\log D(\xi_0) \left(\prod_0^n \lambda_i\right)^{-1}$ и проверить, что $f(x) = \prod_0^n (x - x_i)$ удовлетворяет дифференциальному уравнению, которое определяет многочлен Лагерра $L_{n+1}^{(\alpha)}(x)$ с точностью до постоянного множителя. Кроме того, напомним, что максимум $\Pi\lambda_i$ при ограничениях $\lambda_i \geqslant 0, \Sigma\lambda_i = 1$, достигается при $\lambda_i = (n+1)^{-1}$. Суммируя сказанное, получаем, что максимум $D(\xi)$ равен

$$D(\xi_0) = (n+1)^{-n-1} e^{-(n+1)(n+1+\alpha)} \prod_0^n x_i^{\alpha+1} \prod_{0 \leqslant i < j \leqslant n} (x_i - x_j)^2,$$

где $x_0, x_1, \ldots, x_n$ — нули $L_{n+1}^{(\alpha)}(x)$, и это значение не достигается ни при каком другом выборе $x_0, \ldots, x_n$.

Численное значение $D(\xi_0)$, данное в формулировке теоремы 3.4, можно установить, сравнивая его со значением дискриминантной функции $L_{n+1}^{(\alpha)}(x)$, данной у Сеге [1959] (стр. 141).

Некоторые из предшествующих рассуждений, относящиеся к числу точек спектра максимизирующей меры $\xi_0$ для задачи 2.1, непосредственно приложимы для случая $f_i(x) = \sqrt{w(x)}\, x^i$ ($i = 0, 1, \ldots, n, a \leqslant x \leqslant b$), где $w(x)$ — весовая функция общего вида.

Полагаем, что $w(x)$ непрерывна на интервале $(a, b)$, который может быть конечным или бесконечным, и что если $a = -\infty$ ($b = +\infty$), то предел $w(x)\, x^{2n}$ при $x \to a$ ($x \to b$) конечен. Допустим, что $\mu_\nu = \int_a^b w(x)\, x^\nu \xi(dx)$, $\nu = 0, 1, \ldots, 2n$, и обозначим, как прежде, определитель $\| \mu_{i+j} \|_{i,j=0}^n$ через $D(\xi)$. Будем предполагать, что $w(x) > 0$ по крайней мере для $n+1$ точки, иначе $D(\xi) \equiv 0$.

Если $D(\xi_0) = \max_\eta D(\eta)$, то можно вывести (используя, как в доказательстве теоремы 3.4, опорную гиперплоскость), что верхняя граница для числа точек в спектре $\xi_0$ есть максимальное число точек, для которых в соотношении $w(x)\sum_{i=0}^{2n} a_i x^i \leq d$, $x \in [a, b]$, $\sum_{i=0}^{2n} a_i^2 > 0$, достигается равенство.

Используя этот факт, мы докажем следующую теорему.

Теорема 3.6. *Максимум $D(\eta)$, взятый по множеству всех вероятностных мер, достигается на мере $\xi_0$, которая сосредоточена в точности в $n+1$ точке тогда и только тогда, когда выполняется какое-либо из следующих условий:*

*(1) система $\{1, w(x), xw(x), \ldots, x^{2n}w(x)\}$ есть $T$-система на $[a, b]$;*

*(2) $w(x) = 1/P(x)$, где $P(x)$ есть многочлен, положительный на $[a, b]$, и $P^{(2n+1)}(x)$ не имеет нулей в открытом интервале $(a, b)$, например, если $P(x)$ имеет только вещественные нули, которые все меньше $a$ или все больше $b$;*

*(3) $w(x)$ можно равномерно аппроксимировать весовыми функциями типа указанных в п. (2);*

*(4) $w(x) = 1/P(x)$, где $P(x)$ — многочлен степени, не большей $2n$, положительный на $[a, b]$.*

Замечание 3.2. Пункт (3) применим, если $a = 0$, $b = \infty$, а $w(x)$ — преобразование Лапласа от частотной функции Полиа на $[0, \infty)$ (см. Шенберг [1951]).

Доказательство. Пусть (1) выполняется. В замечаниях перед формулировкой теоремы показано, что существуют вещественные константы $a_i$, $i = 0, 1, \ldots, 2n$, $(\Sigma a_i^2 > 0)$ такие, что $w(x)\sum_{0}^{2n} a_i x^i \leqslant d$ для некоторого вещественного $d$, и равенство достигается на спектре $\xi_0$. Если существуют по крайней мере $n+2$ точки $x_1 < x_2 < \ldots < x_{n+2}$ в спектре $\xi_0$, то могут встретиться три случая: $x_i \in (a, b)$ для всех $i$; $n+1$ из $x_i$ лежит в $(a, b)$ и одно расположено в граничной точке $a$ или $b$; $x_2, \ldots, x_{n+1}$ принадлежат $(a, b)$, $x_1 = a$, $x_{n+2} = b$. Для каждой из трех возможностей «многочлен» $w(x)\sum_{0}^{2n} a_i x^i - d$ имеет по крайней мере $2n+2$ нуля, где мы считаем неузловые нули дважды, а узловые нули по одному разу. Это невозможно по теореме 4.2. гл. I.

Если выполняется условие (2), то, как и выше, функция $g(x) = \sum_{0}^{2n} a_i x^i - dP(x)$ имеет по крайней мере $2n+2$ нуля в интервале $[a, b]$, где нули в точках на концах $a$ и $b$ считаются один раз. Если $d = 0$, то это невозможно, а если $d \neq 0$, то функция $g^{(2n+1)}(x) = -dP^{(2n+1)}(x)$ имеет нуль в промежутке $(a, b)$, что противоречит предположению.

Третья часть доказывается переходом к пределу. Условие (4) трактуется несколько иным способом. Пусть $M = \{(\mu_0(\xi), \ldots$ $\ldots, \mu_{2n}(\xi)\} \left( \mu_\nu = \int\limits_a^b w(x)\, x^\nu \xi(dx) \right)$ обозначает пространство моментов, полученное варьированием $\xi$ над классом всех вероятностных мер на $[a, b]$, и пусть $\mathcal{M}$ обозначает замкнутый выпуклый конус, порожденный $M$. Так как $P(x)$ — многочлен степени не выше $2n$, мы можем записать $P(x) = \sum\limits_{i=0}^{2n} b_i x^i$. Отсюда следует, что $M$ есть сечение конуса $M$, состоящее из точек $(c_0, \ldots, c_{2n}) \in \mathcal{M}$, удовлетворяющих условию нормировки $\sum\limits_{i=0}^{2n} b_i c_1 = 1$. Теперь, если $\xi_0$ максимизирует определитель $D(\xi)$, то $(\mu(\xi_0), \ldots, \mu_{2n}(\xi_0))$ есть граничная точка $M$. Так как $M$ — сечение $\mathcal{M}$, то точка $(\mu_0(\xi_0), \ldots, \mu_{2n}(\xi_0))$ должна быть также граничной точкой конуса $\mathcal{M}$. В этом случае существуют вещественные константы $a_0, a_1, \ldots, a_{2n} \left( \sum\limits_{i=0}^{2n} a_i^2 > 0 \right)$ такие, что

$$\sum_{i=0}^{2n} a_i \frac{x^i}{w(x)} \leq 0, \qquad a \leq x \leq b, \tag{3.7}$$

и равенство выполняется для $x$, принадлежаших спектру $\xi_0$. Но из $|\mathbf{M}(\xi_0)| > 0$ следует, что спектр $\xi_0$ содержит по крайней мере $n+1$ точку. Однако $w(x) > 0$ для $x \in [a, b]$, так что равенство может достигаться в (3.7) самое большее для $n+1$ точки, куда включаются две точки на концах $a$ и $b$. Следовательно, спектр $\xi_0$ содержит в точности $n+1$ точку.

Теоремы 3.1—3.5 восходят к Шенбергу [1959]. Он использовал более громоздкое вариационное доказательство. Наши методы являются геометрическими и более простыми, кроме того, они приводят к общей формулировке теоремы 3.6.

З а м е ч а н и е 3.3. Следует отметить, что условия в теореме 3.6 не исчерпываются классическими весовыми функциями из теорем 3.1—3.5.

## § 4. Обобщение задачи Фейера

В § 2 мы доказали, что класс вероятностных мер $\xi$, которые обеспечивают максимум для $|\mathbf{M}(\xi)|$, совпадает с классом вероятностных мер, для которых $\max\limits_x d(x, \xi)$ достигает минимума. В каждом из особых случаев, рассмотренных в § 3, меры, для которых максимум $|\mathbf{M}(\xi)|$ достигается, приписывают равные массы $n+1$ точке. Здесь $n+1$ есть размер матрицы $\mathbf{M}(\xi)$. Мы докажем (лемма 4.1), что при $\xi$, приписывающей равные массы $n+1$ точкам $x_0, x_1, \ldots, x_n$ и $f_i(x) = x^i w^{1/2}(x)$, $i = 0, 1, \ldots, n$, выражение $d(x, \xi)$ сводится к

$$d(x, \xi) = \frac{1}{n+1} w(x) \sum_{i=0}^{n} r_i^2(x)$$

(см. (1.4)). Таким образом, $\sup_x d(x, \xi)$ есть в точности величина (1.5), которую требуется минимизировать в задаче 1.2. С помощью теоремы эквивалентности (теорема 2.1) мы получаем, что теоремы из § 3 дают решение задачи 1.2 для соответствующих весовых функций $w(x)$.

Для удобства мы напомним обозначения. Пусть $w(x)$ — неотрицательная функция такая, что функция $w(x)x^{2n}$ интегрируема на $[a, b]$. Пусть $\Phi_{n+1}$ — класс последовательностей $(x_0, x_1, \ldots, x_n)$, удовлетворяющих неравенствам $a \leq x_0 < x_1 < \ldots < x_n \leq b$ и ограничениям $w^{1/2}(x_i) r_\nu(x_i) = \delta_{i\nu}$, $i, \nu = 0, 1, \ldots, n$, где $r_\nu(x)$ — многочлен степени $n$, обращающийся в нуль в точках $x_0, \ldots, x_{\nu-1}, x_{\nu+1}, \ldots, x_n$. Многочлен $r_\nu(x)$ есть с точностью до постоянного множителя $\nu$-й фундаментальный интерполяционный многочлен Лагранжа. Более точно

$$r_\nu(x) = w^{-1/2}(x_\nu)\, l_\nu(x), \tag{4.1}$$

где

$$l_\nu(x) = l(x)\,[l'(x_\nu)(x - x_\nu)]^{-1},$$

$$l(x) = \sum_{j=0}^{n} (x - x_j).$$

Отметим, что точки $x_\nu$ должны быть определены так, что $w(x_\nu) > 0$ и что степень $r_\nu$ всегда на единицу меньше числа заданных точек.

Лемма 4.1. *Положим, что* $\mathfrak{f}(x) = (f_0(x), \ldots, f_n(x))$, *где* $f_i(x) = \sqrt{w(x)}\, x^i$, $i = 0, \ldots, n$, *и что* $w(x) \geqslant 0$ — *непрерывная на* $[a, b]$ *функция. Пусть* $r_\nu(x)$, $\nu = 0, \ldots, n$ *определена соотношением* (4.1), *тогда*

$$w(x) \sum_{i=0}^{n} r_i^2(x) = \frac{1}{n+1} (\mathfrak{f}(x),\ \mathbf{M}^{-1}(\xi_0)\, f(x)), \tag{4.2}$$

*где* $\xi_0$ — *мера, приписывающая массу* $1/(n+1)$ *каждой из точек* $x_0, x_1, \ldots, x_n$.

Доказательство. Поскольку $w^{1/2}(x_i)\, r_\nu(x_i) = \delta_{i\nu}$, то получаем

$$w^{1/2}(x)\, \mathbf{r}(x) = \mathbf{A}\mathfrak{f}(x),$$

где $\mathbf{A}$ есть матрица размера $(n+1)\times(n+1)$ такая, что $\mathbf{A}^{-1} = \|b_{ij}\|$, $b_{ij} = w^{1/2}(x_j)\, x_j^i$, $i, j = 0, 1, \ldots, n$. Мы использовали тот факт, что многочлены степени $n$, совпадающие в $n+1$ точке, тождественно равны. Итак,

$$w(x)\sum_{i=0}^{n} r_i^2(x) = w(x)(r(x), r(x)) = (\mathbf{Af}(x), \mathbf{Af}(x)).$$

Легко проверить, что $[(n+1)\,\mathbf{A}'\mathbf{A}]^{-1} = \mathbf{M}(\xi_0)$, где $\xi_0$ — мера, приписывающая массу $1/(n+1)$ каждой из точек $x_0, x_1, \ldots, x_n$. Мы поэтому заключаем, что

$$w(x)\sum_{i=0}^{n} r_i^2(x) = \frac{1}{n+1}(\mathbf{f}(x), \mathbf{M}^{-1}(\xi_0)\,\mathbf{f}(x)).$$

Соединяя результаты леммы 4.1 и теоремы 2.1, мы получаем следующие теоремы.

Теорема 4.1. *Если* $w(x) = (1-x)^{\alpha+1}(1+x)^{\beta+1}$, $(\beta > -1$, $\alpha > -1)$, $x \in [-1, 1]$, *то*

$$\inf_{\Phi_{n+1}} \sup_{-1\leqslant x\leqslant 1} (1-x)^{\alpha+1}(1+x)^{\beta+1}\{r_0^2(x) + \ldots + r_n^2(x)\} = 1,$$

*и нижняя грань достигается в единственном случае, когда* $x_0, x_1, \ldots$ $\ldots, x_n$ *суть нули многочлена Якоби* $P_{n+1}^{(\alpha,\beta)}(x)$.

Теорема 4.2. *Если* $w(x) = e^{-x}$, $x \in [0, \infty)$, *то*

$$\inf_{\Phi_{n+1}} \sup_{0\leqslant x<\infty} e^{-x}\{r_0^2(x) + \ldots + r_n^2(x)\} = 1,$$

*и нижняя грань достигается тогда и только тогда, когда* $x_0 = 0$, *а* $x_1, \ldots, x_n$ *суть нули многочлена Лагерра* $L_n^{(1)}(x)$ *с параметром* 1.

Теорема 4.3. *Если* $w(x) = x^{\alpha+1}e^{-x}$ $(\alpha > -1)$, $x \in [0, \infty)$, *то*

$$\inf_{\Phi_{n+1}} \sup_{0\leqslant x<\infty} x^{\alpha+1}e^{-x}\{r_0^2(x) + \ldots + r_n^2(x)\} = 1,$$

*и нижняя грань достигается в единственном случае, когда* $x_0, x_1, \ldots$ $\ldots, x_n$ *суть нули* $L_{n+1}^{(\alpha)}(x)$ *многочлена Лагерра с параметром* $\alpha$.

Наконец, для нормальной весовой функции мы имеем следующий результат.

Теорема 4.4. *Если* $w(x) = e^{-x^2}$, $x \in (-\infty, \infty)$, *то*

$$\inf_{\Phi_{n+1}} \sup_{-\infty<x<\infty} e^{-x^2}\{r_0^2(x) + \ldots + r_n^2(x)\} = 1,$$

*и нижняя грань достигается в единственном случае, когда* $x_0, x_1, \ldots$ $\ldots, x_n$ *суть нули многочлена Эрмита* $H_{n+1}(x)$.

Мы обсудим только доказательство теоремы 4.2, так как остальные доказательства получаются точно так же.

Доказательство теоремы 4.2. В данном случае $[a, b] = [0, \infty]$ и

$$f_i(x) = \begin{cases} x^i e^{-x/2}, & 0 \leq x < \infty \\ 0, & x = \infty \end{cases} \qquad (i = 0, 1, \ldots, n).$$

По лемме 4.1

$$\sup_{0 \leqslant x < \infty} e^{-x} \sum_{i=0}^{n} r_i^2(x) = \sup_{0 \leqslant x < \infty} (n+1)^{-1} (\mathbf{f}(x), \mathbf{M}^{-1}(\xi) \mathbf{f}(x)), \qquad (4.3)$$

где $\xi$ определено на $[0, \infty]$ и приписывает равные массы $n+1$ точкам $x_0, x_1, \ldots, x_n$. Кроме того, теорема 2.1 утверждает, что нижняя грань правой части (4.3) равна единице, и это значение достигается тогда и только тогда, когда $\xi$ максимизирует $|\mathbf{M}(\xi)|$. Вместе с тем теорема 3.3 утверждает, что $\max_{\xi} |\mathbf{M}(\xi)|$ достигается в единственном случае на мере $\xi_0$, которая приписывает массу $(n+1)^{-1}$ каждой из $n+1$ точек $x_0, x_1, \ldots, x_n$, где $x_0 = 0$ и $x_1, x_2, \ldots, x_n$ суть нули $L_n^{(1)}(x)$.

Следовательно,

$$\inf_{\Phi_{n+1}} \sup_{0 \leqslant x < \infty} e^{-x} \sum_{i=0}^{n} r_i^2(x) = 1,$$

и нижняя грань достигается тогда и только тогда, когда $x_0 = 0$ и $x_1, x_2, \ldots, x_n$ таковы, как указано выше. Доказательство закончено.

Замечание 4.1. В случае $f_i(x) = \sqrt{w(x)}\, x^i$, $i = 0, 1, \ldots, n$, теорема 3.6 обеспечивает достаточные условия для того, чтобы максимум $|\mathbf{M}(\xi)|$ достигался на мере, сосредоточивающей равные массы в $n+1$ различных точках $x_0, x_1, \ldots, x_n$. При выполнении любого из условий теоремы 3.6 лемма 4.1 в соединении с (3.4) утверждает, что

$$\sup_{[a,b]} w(x) \left[ \sum_{i=0}^{n} r_i^2(x) \right]$$

достигает минимума при условии, что $r_i(x)$ — фундаментальные интерполяционные многочлены, порожденные точками $x_0, x_1, \ldots, x_n$, которые максимизируют выражение

$$\prod_{i=0}^{n} w(x_i) \prod_{0 \leqslant i < j \leqslant n} (x_i - x_j)^2,$$

## § 5. Устойчивые и экономичные интерполяционные системы

Метод исследования наиболее экономичных устойчивых систем комбинирует идеи, заключенные в работе Егервари и Турана [1958], и принцип эквивалентности § 2.

Обратимся к задаче 1.3 § 1. Результаты теорем 4.1 — 4.4 будут применены для характеризации наиболее экономичных устойчивых интерполяционных систем.

Пусть $w(x)$ — неотрицательная интегрируемая функция, заданная на $[a, b]$, такая, что $w(x) > 0$ на $(a, b)$. Пусть для всякого множества $n+1$ точек $x_0, x_1, \ldots, x_n$ в интервале $[a, b]$ $p_\nu(x)$, $\nu=0, 1, \ldots, n$, обозначает $n$ многочленов произвольной степени таких, что

$$w(x_i)\, p_\nu(x_i) = \delta_{i\nu}, \qquad i, \nu = 0, 1, \ldots, n. \tag{5.1}$$

Подчеркнем, что мы позволяем $p_\nu$ иметь произвольную степень, тогда как в (4.1) многочлены $r_\nu$ имеют минимальную степень. Система $\{w(x)\, p_\nu(x)\}_0^n$ называется устойчивой интерполяционной системой, если всякий раз, когда $y_\nu \geqslant 0$, $\nu = 0, 1, \ldots, n$, имеем

$$0 \leq \sum_{\nu=0}^{n} y_\nu w(x)\, p_\nu(x) \leq \max_\nu y_\nu. \tag{5.2}$$

Отметим, что функция $\sum\limits_{\nu=0}^{n} y_\nu w(x)\, p_\nu(x)$ принимает значения $y_\nu$ в точках $x_\nu$. Следует отметить, как и в § 4, что $x_i$ не могут быть выбраны в точках $x$, где $w(x)$ обращается в нуль. Устойчивая интерполяционная система называется наиболее экономичной при условии, что величина $\sum\limits_{\nu=0}^{n} \deg p_\nu(x)$ минимальна.

Из определения устойчивой системы вытекает

$$0 \leq w(x)\, p_\nu(x) \leq \sum_{\nu=0}^{n} w(x)\, p_\nu(x) \leq 1. \tag{5.3}$$

И наоборот, неравенства (5.3) гарантируют условия устойчивости (5.2).

В случае, когда $w(x)$ обращается в нуль в обеих конечных точках $a$ и $b$, многочлен $p_\nu(x)$ имеет нули четной кратности в $x_0, x_1, \ldots$ $\ldots, x_{\nu-1}, x_{\nu+1}, \ldots, x_n$, так что $\deg p_\nu(x) \geqslant 2n$ и $\min \sum\limits_{\nu=0}^{n} \deg p_\nu(x) \geqslant$ $\geqslant 2n(n+1)$.

Далее, для весовых функций $(1-x)^{\alpha+1}(1+x)^{\beta+1}$, $x^{\alpha+1}e^{-x}$ и $e^{-x^2}$ мы выводим на основании теорем 4.1, 4.3 и 4.4, что если $p_\nu(x) = r_\nu^2(x)$, $\nu = 0, 1, \ldots, n$, и $r_\nu(x)$ определено в (4.1), то

$$\sup_{a \leqslant x \leqslant b} w(x) \sum_{\nu=0}^{n} p_\nu(x) = \sup_{a \leqslant x \leqslant b} w(x) \sum_{\nu=0}^{n} r_\nu^2(x) \geqslant 1. \tag{5.4}$$

Кроме того, равенство в (5.4) достигается в том и только в том случае, когда $x_0, x_1, \ldots, x_n$ суть нули $P_{n+1}^{(\alpha,\beta)}(x)$, $L_{n+1}^{(\alpha)}(x)$ и $H_{n+1}(x)$ соответственно.

Так как равенство в (5.4) достигается в том случае, когда $x_0, x_1, \ldots, x_n$ суть указанные нули, то результирующие системы $\{w(x)\, r_\nu^2(x)\}$ устойчивы. Более того, $r_\nu^2(x)$ имеет степень $2n$, так что $\sum_{\nu=0}^{n} \deg r_\nu^2(x) = 2n(n+1)$, что совпадает с установленной ранее нижней границей. Следовательно, система $\{w(x)\, r_\nu^2(x)\}$ является наиболее экономичной. Наоборот, если $\min \sum_{\nu=0}^{n} \deg p_\nu(x) = 2n(n+1)$, то многочлены $p_\nu(x)$, $\nu = 0, 1, \ldots, n$, необходимо имеют вид $p_\nu(x) = r_\nu^2(x)$, если потребовать, чтобы система $\{w(x)\, p_\nu(x)\}$ была наиболее экономичной и устойчивой. Но равенство достигается в (5.4) только при условиях, упомянутых выше, т. е. для единственного множества точек $x_0, x_1, \ldots, x_n$.

Проведенное рассуждение есть доказательство следующей теоремы.

Теорема 5.1. *Если*

$$r_\nu(x) = w^{-1/2}(x)\, \frac{l(x)}{l'(x_\nu)(x - x_\nu)},$$

*где* $l(x) = \sum_{0}^{n} (x - x_\nu)$, *то в следующих трех случаях*:

(1) $w(x) = (1-x)^{\alpha+1}(1+x)^{\beta+1}$, $x \in [-1, 1]$, $\alpha > -1$, $\beta > -1$,
(2) $w(x) = x^{\alpha+1} e^{-x}$, $x \in [0, \infty)$, $\alpha > -1$,
(3) $w(x) = e^{-x^2}$, $x \in (-\infty, \infty)$

*система* $\{w(x)\, p_\nu(x)\}$ *является наиболее экономичной устойчивой интерполяционной системой тогда и только тогда, когда* $p_\nu(x) = r_\nu^2(x)$ *и* $x_0, x_1, \ldots, x_n$ *суть соответственно нули*:

(1) $P_{n+1}^{(\alpha,\beta)}(x) = 0$,
(2) $L_{n+1}^{(\alpha)}(x) = 0$,
(3) $H_{n+1}(x) = 0$.

Случаи, когда $w(x) = e^{-x}$, $x \in [0, \infty)$ или $w(x) = 1$, $x \in [-1, 1]$ требуют несколько иного анализа, чем данный выше.

Если $w(x) = e^{-x}$, $x \in [0, \infty)$, то положим, что $x_0, x_1, \ldots, x_n$ — $n+1$ точек из $[0, \infty)$, и рассмотрим систему $n+1$ функций $e^{-x} p_\nu(x)$, $\nu = 0, 1, \ldots, n$, где $e^{-x_i} p_\nu(x_i) = \delta_{i\nu}$, а $p_\nu$ — многочлены произвольной степени.

Так как $w(0) \neq 0$, то точка $x_0$ может быть выбрана равной нулю. Принимая во внимание эту возможность, мы заключаем, как и ранее, что если система $\{w(x)\,p_\nu(x)\}_0^n$ устойчива, то $\sum_{\nu=0}^{n}\deg p_\nu(x)\geqslant \geqslant 2n+n[2(n-1)+1]=n(2n+1)$. Определим

$$\tilde{p}_\nu(x)=\begin{cases} L_n^2(x), & \nu=0,\\ \dfrac{xe^{x_\nu}}{x_\nu}\left[\dfrac{L_n(x)}{(x-x_\nu)L_n'(x_\nu)}\right]^2, & \nu=1,2,\ldots,n,\end{cases}\qquad (5.5)$$

где $x_1, x_2, \ldots, x_n$ суть нули $L_n(x)=L_n^{(1)}(x)$ многочлена Лагерра. Очевидно, $e^{-x_i}\tilde{p}_\nu(x_i)=\delta_{i\nu}$, $i,\nu=0, 1, \ldots, n\,(x_0=0)$ и $\sum_{\nu=0}^{n}\deg\tilde{p}_\nu(x)= = n(2n+1)$. Чтобы проверить, что $\{e^{-x}p_\nu(x)\}_0^n$ — устойчивая система (и, следовательно, наиболее экономичная устойчивая система), достаточно установить неравенство

$$1-xe^{-x}\sum_{1}^{n}r_\nu^2(x)\geqslant e^{-x}L_n^2(x),\qquad x\in[0,\infty),\qquad (5.6)$$

где

$$r_\nu^2(x)=\frac{e^{x_\nu}}{x_\nu}\left[\frac{L_n(x)}{L_n'(x_\nu)(x-x_\nu)}\right]^2$$

и $x_1, x_2, \ldots, x_n$ — нули $L_n(x)$. По теореме 4.3 при $\alpha=0$ и $n$, замененном на $n-1$, левая часть (5.6) всегда неотрицательна. Кроме того, левая часть (5.6) равна нулю в каждой из точек $x_\nu$, $\nu=1,\ldots$ $\ldots, n$, так что

$$g(x)=1-xe^{-x}\sum_{1}^{n}r_\nu^2(x)-e^{-x}L_n^2(x)$$

имеет двойной нуль в каждой из точек $x_\nu$, $\nu=1,\ldots,n$. Далее $L_n(0)=1$, так что нуль достигается также в точке $x=0$ и общее число нулей не менее $2n+1$. Так как $g(x)$ стремится к единице при $x\to\infty$, предположение $g(y)<0$ при некотором $y$ влечет по крайней мере $2n+2$ нулей $g(x)$. Это, однако, невозможно, так как тогда повторное применение теоремы Ролля показывает, что производная $(d^{2n+1}/dx^{2n+1})\,e^x g(x)=e^x$ обращается в нуль на $(0,\infty)$. Это, очевидно, невозможно. Поэтому неравенства $g(x)\geqslant 0$ и (5.6) справедливы.

Мы теперь применим теорему 4.3, чтобы показать, что $\{w(x)\times \times\tilde{p}_\nu(x)\}_0^n$ есть единственная наиболее экономичная устойчивая система для весовой функции $w(x)=e^{-x}$. В самом деле, если система $\{w(x)\,p_\nu(x)\}_0^n$ наиболее экономична и устойчива, то $p_\nu(x)$, $\nu=1, 2, \ldots$ $\ldots, n$, может быть выражена в виде $p_\nu(x)=xr_\nu^2(x)$, $\nu=1, 2, \ldots, n$, где $r_\nu(x)$ имеет степень $n-1$, $e^{-x_i}x_ir_\nu^2(x_i)=\delta_{i\nu}$, $i,\nu=1,\ldots,n$, и

$0 < x_1 < x_2 < \ldots < x_n$. Если система $\{e^{-x}p_\nu(x)\}_0^n$ устойчива, то мы знаем, что $e^{-x}\sum\limits_{i=0}^{n} p_i(x) \leq 1$ или

$$xe^{-x}\sum_{\nu=1}^{n} r_\nu^2(x) \leq 1 \tag{5.7}$$

(отметим, что член $p_0(x)e^{-x}$ опущен). Однако теорема 4.3 утверждает, что (5.7) выполняется только в случае, когда $x_1, x_2, \ldots x_n$ — нули $L_n^{(1)}(x) = L_n(x)$.

Таким образом мы доказали следующую теорему.

Теорема 5.2. *Если $w(x) = e^{-x}$, то единственная система $\{e^{-x}p_\nu(x)\}_0^n$, которая является устойчивой и наиболее экономичной, задается посредством соотношений $\tilde{p}_0(x) = L_n^2(x)$,*

$$\tilde{p}_\nu(x) = \frac{xe^{x_\nu}}{x_\nu}\left[\frac{L_n(x)}{(x-x_\nu)L_n'(x_\nu)}\right]^2, \qquad \nu = 1, \ldots, n,$$

*где $L_n(x)$ — многочлен Лагерра $L_n(x) = L_n^{(0)}(x)$, а $x_1, \ldots, x_n$ суть нули $L_n(x)$.*

Оставшийся случай, когда $w(x) \equiv 1$, $x \in [-1, 1]$, можно исследовать подобным образом. Так как $w(-1)$ и $w(1)$ не равны нулю, то возможности $x_0 = -1$ и (или) $x_n = 1$ не исключены. Нижняя граница для $\sum\limits_{\nu=0}^{n} \deg p_\nu(x)$ равна, как легко заметить, $2n^2$. Пусть $\tilde{p}_\nu(x)$, $\nu = 0, 1, \ldots, n$, определены как

$$\tilde{p}_\nu(x) = \begin{cases} \dfrac{1-x}{2}P_{n-1}^2(x), & \nu = 0, \\[2ex] \dfrac{1-x^2}{1-x_\nu^2}\left[\dfrac{P_{n-1}(x)}{(x-x_\nu)P_{n-1}'(x_\nu)}\right]^2, & \nu = 1, 2, \ldots, n-1, \\[2ex] \dfrac{1+x}{2}P_{n-1}^2(x), & \nu = n, \end{cases} \tag{5.8}$$

где $P_{n-1}(x) = P_{n-1}^{(0,0)}(x)$ есть $(n-1)$-й многочлен Лежандра и $x_1, x_2, \ldots, x_{n-1}$ — нули $P_{n-1}(x)$. Установим, что система (5.8) есть единственная наиболее экономичная устойчивая интерполяционная система для весовой функции $w(x) \equiv 1$.

Ясно, что $\sum\limits_{\nu=0}^{n} \deg \tilde{p}_\nu(x) = 2n^2$. Чтобы показать устойчивость данной системы, достаточно проверить тождество

$$1 - \sum_{\nu=1}^{n-1} \frac{1-x^2}{1-x_\nu^2}\left[\frac{P_{n-1}(x)}{(x-x_\nu)P_{n-1}'(x_\nu)}\right]^2 = P_{n-1}^2(x). \tag{5.9}$$

По теореме 4.1 при $\alpha = \beta = 0$ и $n$, замененном на $n-2$, левая часть (5.9) неотрицательна и имеет двойной нуль в каждом из нулей $P_{n-1}(x)$. Так как $P_{n-1}(1) = 1$ и степени обеих частей одинаковы, то (5.9) выполняется. Рассуждение, доказывающее, что система (5.8) является единственной наиболее экономичной устойчивой системой, проводится, так же как в случае $w(x) = e^{-x}$, в теореме 4.2.

Мы получаем следующий результат.

Теорема 5.3. *Если* $w(x) \equiv 1$, $x \in [-1, 1]$, *то единственная наиболее экономичная устойчивая система* $\{p_\nu(x)\}_0^n$ *задается посредством* (5.8).

Замечание 5.1. В теореме 5.3 обнаруживается более сильное свойство устойчивости для многочленов Лежандра, а именно, для всех $y_\nu \geqslant 0$, $\nu = 0, 1, \ldots, n$, выполняется

$$\min_\nu \{y_\nu\} \leq \sum_{\nu=0}^{n} y_\nu p_\nu(x) \leq \max_\nu \{y_\nu\}. \tag{5.10}$$

Соотношение (5.10) есть непосредственное следствие тождества (5.9). Это свойство отличается от обычного понятия устойчивости тем, что левая часть (5.2) заменена на $\min\limits_\nu y_\nu$. В случае других весовых функций левое неравенство в (5.10), вообще говоря, не выполняется.

Замечание 5.2. Предположим, что для данной весовой функции $w(x)$ интерполяционная система $\{w(x)p_\nu(x)\}_0^n$, порожденная точками $x_0, x_1, \ldots, x_n$, наиболее экономична и устойчива. Если $a < x_0$ и $x_n < b$, а $p_\nu$ — многочлены степени $2n$, то $p_\nu$ принимают вид $p_\nu(x) = r_\nu^2(x)$, $\nu = 0, 1, \ldots, n$, где $r_\nu(x)$ — многочлены степени $n$. Тогда $w(x)\sum\limits_{\nu=0}^{n} r_\nu(x) \leq 1$, так что по лемме 4.1 функция $d(x, \xi^*)$, определенная в (2.5), удовлетворяет равенству $\sup\limits_x d(x, \xi^*) = n+1$, где $\xi^*$ есть мера, приписывающая массу $1/(n+1)$ каждой из точек $x_0, x_1, \ldots, x_n$. Из теоремы 2.1 следует, что $\sup\limits_\xi |\mathbf{M}(\xi)|$ для случая $f_i(x) = \sqrt{w(x)}\,x^i$, $i = 0, 1, \ldots, n$, достигается для меры $\xi^*$, помещающей равные массы в точках $x_0, x_1, \ldots, x_n$.

## § 6. Тригонометрические системы

Этот параграф посвящен получению некоторых аналогов предыдущих теорем для тригонометрических функций.

Рассмотрим сначала систему функций $f_k(\theta) = \cos k\theta$, $k = 0, \ldots, n$, при $0 \leq \theta \leq \pi$. Отметим, что в качестве области определения взят интервал $[0, \pi]$, а не больший интервал $[-\pi, \pi]$. В этой ситуации мы можем свести рассмотрение к случаю обычных степенных функций $t^l$ на $[-1, 1]$. Полное множество тригонометрических функций, включающее и $\sin k\theta$ и $\cos k\theta$, будет изучено ниже.

Будут полезны две следующие элементарные леммы.

Лемма 6.1. *Пусть* $P_i(x) = \sum_{k=0}^{n} a_{ik} x^k$, $i = 0, 1, \ldots, n$, — *произвольная система многочленов, и пусть* $\mathbf{B}$ — *матрица с элементами* $b_{ij} = \int_{-1}^{1} P_i(x) P_j(x) \xi(dx)$ $(i, j = 0, 1, \ldots, n)$. *Тогда*

$$|\mathbf{B}| = |\mathbf{A}|^2 \det[\| \mu_{i+j} \|_{i,j=0}^{n}], \tag{6.1}$$

*где*

$$\mu_k = \int_{-1}^{1} x^k \xi(dx), \quad k = 0, 1, \ldots, 2n,$$

$$\mathbf{A} = \| a_{ij} \|_{i,j=0}^{n}.$$

Лемма 6.2. *Если* $m_{ij} = \int_{0}^{\pi} \cos i\theta \cos j\theta \zeta(d\theta)$, *то*

$$|\mathbf{M}(\xi)| = \det[\| m_{ij} \|_{i,j=0}^{n}] = |\mathbf{A}_0|^2 \det[\| \mu_{i+j} \|_{i,j=0}^{n}], \tag{6.2}$$

*где* $\eta(E) = \xi(\cos^{-1} E)$ *для всех* $E \subset [-1, 1]$ *и* $\mathbf{A}_0$ — *треугольная матрица, i-й столбец которой есть вектор коэффициентов i-го многочлена Чебышева первого рода.*

По теореме 3.1 $\det[\| \mu_{i+j} \|]$ максимизируется единственным образом мерой $\eta$, которая помещает массы $(n+1)^{-1}$ в нули $x_0, x_1, \ldots, x_n$ многочлена $(1-x^2) P'_n(x)$, где $P_n(x)$ — $n$-й многочлен Лежандра. Следовательно, $|\mathbf{M}(\xi)|$ максимизируется единственным образом мерой $\xi$, которая помещает массу $(n+1)^{-1}$ в точки $\theta_k$, $k = 0, 1, \ldots, n$, где $x_k = \cos\theta_k$. Используя тот факт, что $\mathbf{A}_0$ — треугольная матрица, а коэффициент при $x^k$ $(k \neq 0)$ в $T_k(x)$ равен $2^{k-1}$, получим следующий результат.

Теорема 6.1. *Если* $f_k(\theta) = \cos k\theta$, $k = 0, 1, \ldots, n$, *при* $0 \leq \theta \leq \pi$, *то*

$$|\mathbf{M}(\xi)| \leq 4^{n^2} \left( \prod_{1}^{n} \nu^{\nu} \right)^4 n^{-n} \prod_{1}^{2n} \nu^{-\nu},$$

*и равенство достигается тогда и только тогда, когда* $\xi$ *сводится к мере, приписывающей равные массы* $n+1$ *точкам* $\theta_k$, $k = 0, \ldots, n$, *где* $\cos\theta_k = x_k$ *и* $x_0, \ldots, x_n$ *суть нули* $(1-x^2) P'_n(x)$ ($P_n(x)$ — *n-й многочлен Лежандра*).

Следующая теорема является аналогом результата Фейера (сравните теоремы 4.1—4.4) в случае тригонометрических многочленов.

Доказательство использует рассуждения, проведенные для теоремы 4.2 и поэтому опущено.

Теорема 6.2. *Пусть* $\theta_0, \theta_1, \ldots, \theta_n$ — *различные точки, принадлежащие отрезку* $[0, \pi]$. *Определим* $\lambda_k(\theta)$, $k = 0, 1, \ldots, n$, *как многочлен по* $\cos k\theta$, $k = 0, 1, \ldots, n$, *удовлетворяющий условию* $\lambda_k(\theta_\nu) = \delta_{k\nu}$ *при* $k, \nu = 0, 1, \ldots, n$.

*Тогда*

$$\sup_{0 \leqslant \theta \leqslant \pi} \sum_{k=0}^{n} [\lambda_k(\theta)]^2$$

*достигается единственным образом в фундаментальных интерполяционных точках* $\theta_k = \cos x_k$, $k = 0, 1, \ldots, n$, *где* $x_0, x_1, \ldots$ $\ldots, x_n$ — *нули* $(1 - x^2) P_n'(x)$.

Для исследования $\mathbf{M}(\xi)$ и $d(x, \xi)$, построенных по функциям 1, $\cos\theta$, $\sin\theta, \ldots, \cos n\theta$, $\sin n\theta$, определенным на $0 \leq \theta \leq 2\pi$, удобно привлечь некоторые понятия и результаты теории меры, обладающие определенными свойствами инвариантности по отношению к группе преобразований.

Пусть $f_i$, $i = 0, 1, \ldots, n$, — линейно независимые функции на произвольном компактном пространстве $X$. Пусть $\mathcal{G}$ — компактная группа преобразований, действующих на $X$ с мерой Хаара $\mu$, $\mu(\mathcal{G}) = 1$. Предположим, что для всякого $g$, принадлежащего $\mathcal{G}$, соответствующее преобразование $\bar{g}$ на пространстве вероятностных мер $\xi$, определенное посредством $\bar{g}\xi = \xi g$, т. е. $\bar{g}\xi(A) = \xi g(A)$ для всех борелевских множеств $A$, удовлетворяет соотношению

$$d(gx, \xi) = d(x, \bar{g}\xi). \tag{6.3}$$

Образуем $\xi^* = \int\limits_{\mathcal{G}} (\bar{g}\xi)\, \mu(dg)$.

Легко показать, что $\xi^*$ — инвариантная мера, т. е. $\xi^*(gA) = \xi^*(A)$ для всех $g$ и всех борелевских множеств $A$.

Теперь, используя тот факт, что матрица $\lambda\mathbf{P}^{-1} + (1 - \lambda)\mathbf{Q}^{-1} - [\lambda\mathbf{P} + (1 - \lambda)\mathbf{Q}]^{-1}$ положительно определена при $0 < \lambda < 1$ всякий раз, когда $\mathbf{P}$ и $\mathbf{Q}$ положительно определены и $\mathbf{P} \neq \mathbf{Q}$, мы получаем, что

$$\sup_x d(x, \xi) \geqslant \sup_x \int d(gx, \xi)\, \mu(dg) = \sup_x \int d(x, \bar{g}\xi)\, \mu(dg) =$$

$$= \sup_x \left(\mathbf{f}(x), \left[\int \mathbf{M}^{-1}(\bar{g}\xi)\, \mu(dg)\right] \mathbf{f}(x)\right) \geqslant$$

$$\geqslant \sup_x \left(\mathbf{f}(x), \left[\int \mathbf{M}(\bar{g}\xi)\, \mu(dg)\right]^{-1} \mathbf{f}(x)\right) = \sup_x d(x, \xi^*). \tag{6.4}$$

Эти соотношения показывают, что $\xi^*$ минимизирует $\sup\limits_x d(x, \xi)$ всякий раз, когда $\xi$ минимизирует это выражение. Теперь, применяя теорему эквивалентности (теорема 2.1), получаем, что $|\mathbf{M}(\xi^*)| = \max_\xi |\mathbf{M}(\xi)|$. Мы получили доказательство следующей теоремы.

Теорема 6.3. *Предположим, что $\mathcal{G}$ — компактная группа преобразований на $X$ такая, что для всякого $g \in \mathcal{G}$ и всякой вероятностной меры $\xi$ мы имеем $d(gx, \xi) = d(x, \bar{g}\xi)$, где $\bar{g}\xi(A) = \xi(gA)$ для всех $A$ в борелевском поле $\mathcal{B}$. Тогда существует инвариантная мера $\xi^*$, т. е. $\xi^*(gA) = \xi^*(A)$ для всех $A \in \mathcal{B}$, со свойствами*

$$(1) \qquad |\mathbf{M}(\xi^*)| = \sup_\xi |\mathbf{M}(\xi)|,$$

$$(2) \qquad \sup_x d(x, \xi^*) = \inf_\xi \sup_x d(x, \xi).$$

Эта теорема будет теперь применена для получения некоторых периодических аналогов теорем §§ 3 — 5. Положим

$$\mathbf{f}(\theta) = (1, \ \cos\theta, \ \sin\theta, \dots, \cos n\theta, \ \sin n\theta).$$

Пусть $\mathcal{G}$ — группа вращений на круге. Если $g\theta = \theta + \varphi$, то

$$\mathbf{f}(g\theta) = \mathbf{U}_g \mathbf{f}(\theta), \tag{6.5}$$

где $\mathbf{U}_g$ — матрица размера $(2n+1) \times (2n+1)$ с единицей в верхнем левом углу и далее блоками

$$\begin{Vmatrix} \cos k\varphi & -\sin k\varphi \\ \sin k\varphi & \cos k\varphi \end{Vmatrix}, \qquad k = 1, 2, \dots, n,$$

вдоль диагонали. Матрица $\mathbf{U}_g$ удовлетворяет уравнению $\mathbf{U}_g^{-1} = \mathbf{U}_{g^{-1}}$ ($\mathbf{U}_g$ является также ортогональной, но это свойство нам здесь не потребуется). Теперь

$$\mathbf{M}(\bar{g}\xi) = \int [\mathbf{f}(\theta)][\mathbf{f}(\theta)]' \, d\xi(g\theta) = \int [\mathbf{f}(g^{-1}\theta)][\mathbf{f}(g^{-1}\theta)]' \, d\xi(\theta) =$$

$$= \int \mathbf{U}_{g^{-1}} [\mathbf{f}(\theta)][\mathbf{f}(\theta)]' \, \mathbf{U}'_{g^{-1}} \, d\xi(\theta) = \mathbf{U}_{g^{-1}} \mathbf{M}(\xi) \, \mathbf{U}'_{g^{-1}}. \tag{6.6}$$

Следовательно,

$$d(\theta, \bar{g}\xi) = (\mathbf{f}(\theta), \mathbf{M}^{-1}(\bar{g}\xi) \, \mathbf{f}(\theta)) =$$

$$= (\mathbf{f}(\theta), [\mathbf{U}'_{g^{-1}}]^{-1} \mathbf{M}^{-1}(\xi) [\mathbf{U}_{g^{-1}}]^{-1} \mathbf{f}(\theta)) =$$

$$= (\mathbf{U}_{g^{-1}}^{-1} \mathbf{f}(\theta), \mathbf{M}^{-1}(\xi) \, \mathbf{U}_{g^{-1}}^{-1} \mathbf{f}(\theta)) = (\mathbf{U}_g \mathbf{f}(\theta), \mathbf{M}^{-1}(\xi) \, \mathbf{U}_g \mathbf{f}(\theta)) =$$

$$= (\mathbf{f}(g\theta), \mathbf{M}^{-1}(\xi) \, \mathbf{f}(g\theta)) = d(g\theta, \xi). \tag{6.7}$$

Из теоремы 6.3 мы выводим существование $\mathcal{G}$-инвариантной меры $\xi_0$, которая максимизирует $|\mathbf{M}(\xi)|$. Однако единственная инвариантная мера $\xi_0$ есть равномерная на $[0, 2\pi]$ мера. Следовательно, $|\mathbf{M}(\xi_0)| = \max_\xi |\mathbf{M}(\xi)|$, где $\xi_0$ — равномерная на $[0, 2\pi]$ мера.

В этом месте удобно нормировать тригонометрические функции подходящими скалярными множителями, а именно

$$\begin{aligned} \varphi_0(\theta) &= \sqrt{\frac{1}{2n+1}}, \\ \varphi_{2p-1}(\theta) &= \sqrt{\frac{2}{2n+1}}\cos p\theta, \\ \varphi_{2p}(\theta) &= \sqrt{\frac{2}{2n+1}}\sin p\theta, \qquad p = 1, 2, \ldots, n. \end{aligned} \tag{6.8}$$

Лемма 6.3. *Функции* $\varphi_0, \varphi_1, \ldots, \varphi_{2n}$, *определенные в* (6.8), *ортонормальны на* $[0, 2\pi]$ *по отношению к мере* $\eta_0$, *которая размещает равные массы* $(2n+1)^{-1}$ *в точках* $\theta_k = 2\pi k/(2n+1)$, $k = 0, 1, \ldots, 2n$.

Доказательство. Доказательство осуществляется прямым вычислением, использующим экспоненциальные выражения для $\cos p\theta$ и $\sin p\theta$.

С помощью леммы 6.3 легко видеть, что если $\eta_0$ размещает массы $(2n+1)^{-1}$ в точках $\theta_k = 2\pi k/(2n+1)$, $k = 0, 1, \ldots, 2n$, а $\xi_0$ — равномерная мера на $[0, 2\pi]$, то $\mathbf{M}(\xi_0) = \mathbf{M}(\eta_0)$, так что $\eta_0$ максимизирует $|\mathbf{M}(\xi)|$.

Мера $\xi$, максимизирующая $|\mathbf{M}(\xi)|$, очевидно, определена не однозначно. Заметим, что если $|\mathbf{M}(\xi)| > 0$, то $\xi$ необходимо содержит не менее $2n+1$ точек роста. Если мы теперь ограничим внимание мерами, помещающими равные массы в минимальном числе точек, то максимизирующая мера определяется единственным образом посредством $\eta_0$ с точностью до вращения. Чтобы показать это, нам потребуется следующая вспомогательная лемма.

Лемма 6.4. *Максимум выражения*

$$D(\theta_0, \theta_1, \ldots, \theta_{2n}) = \prod_{\nu>\mu} \sin\left(\frac{\theta_\nu - \theta_\mu}{2}\right)$$

*при* $0 \leq \theta_0 < \theta_1 < \ldots < \theta_{2n} < 2\pi$ *достигается только при*

$$\theta_k = \frac{2\pi k}{2n+1} + a, \qquad k = 0, 1, \ldots, 2n; \qquad 0 \leq a < \frac{2\pi}{2n+1}.$$

Доказательство. Так как максимум достигается, только когда $\sin[(\theta_\nu - \theta_\mu)/2] > 0$, мы можем рассмотреть

$$\log D(\theta_0, \ldots, \theta_{2n}) = \sum_{\nu>\mu} \log\sin\left(\frac{\theta_\nu - \theta_\mu}{2}\right) =$$

$$= \sum_{k=1}^{n} \left[\sum_{A(k)} \log\sin\left(\frac{\theta_\nu - \theta_\mu}{2}\right) + \sum_{B(k)} \log\sin\left(\frac{\theta_\nu - \theta_\mu}{2}\right)\right],$$

где

$$A(k) = \{(\nu, \mu) \mid \nu > \mu, \ \nu - \mu = k\},$$

$$B(k) = \{(\nu, \mu) \mid \nu > \mu, \ \nu - \mu = 2n + 1 - k\}.$$

Так как $\sin\theta = \sin(\pi - \theta)$, то предыдущее уравнение может быть записано в виде

$$\log D(\theta_0, \ldots, \theta_{2n}) =$$

$$= \sum_{k=1}^{n} \left[ \sum_{A(k)} \log\sin\left(\frac{\theta_\nu - \theta_\mu}{2}\right) + \sum_{B(k)} \log\sin\left(\frac{2\pi - \theta_\nu + \theta_\mu}{2}\right) \right].$$

Функция $\log\sin\theta$, как легко видеть, строго вогнута при $\theta \in (0, \pi)$. Поэтому из того, что число членов в $A(k) \bigcup B(k)$ всегда равно $2n + 1$, следует, что

$$\log D(\theta_0, \ldots, \theta_{2n}) \leq$$

$$\leq \sum_{k=1}^{n} (2n+1) \log\sin\frac{1}{2n+1} \left[ \sum_{A(k)} \left(\frac{\theta_\nu - \theta_\mu}{2}\right) + \sum_{B(k)} \left(\frac{2\pi - \theta_\nu + \theta_\mu}{2}\right) \right] =$$

$$= \sum_{k=1}^{n} (2n+1) \log\sin\frac{k\pi}{2n+1},$$

и равенство достигается при $\theta_k = (2\pi k/2n + 1) + a$. Кроме того, при $k = 1$ замечаем, что равенство достигается в этом выражении, только если $\theta_1 - \theta_0 = \theta_2 - \theta_1 = \ldots = \theta_{2n} - \theta_{2n-1} = 2\pi - (\theta_{2n} - \theta_0)$ или, эквивалентно, если $\theta_\nu - \theta_{\nu-1} = 2\pi(2n+1)$, $\nu = 1, \ldots, 2n$, что завершает доказательство.

С помощью леммы 6.4 мы теперь докажем, что справедлива

Теорема 6.4. *Если* $f_k(\theta)$, $k = 0, 1, \ldots, 2n$. *суть функции* $1$, $\cos\theta$, $\sin\theta, \ldots, \cos n\theta$, $\sin n\theta$ *при* $\theta \in [0, 2\pi]$, *то*

$$|\mathbf{M}(\xi)| \leq \left(\frac{1}{2}\right)^{2n}. \tag{6.9}$$

*Равенство в* (6.9) *достигается для равномерной меры* $\xi_0$. *Для мер* $\xi$, *сосредоточивающих массы* $(2n+1)^{-1}$ *в* $2n + 1$ *различных точках* $\theta_0, \theta_1, \ldots, \theta_{2n}$, *равенство в* (6.9) *достигается тогда и только тогда, когда*

$$\theta_k = \frac{2\pi k}{2n+1} + a, \qquad k = 0, 1, \ldots, 2n,$$

*где* $0 \leq a < \dfrac{2\pi}{2n+1}$.

Доказательство. Неравенство (6.9) вытекает из тривиального подсчета $|\mathbf{M}(\xi_0)| = |\mathbf{M}(\eta_0)| = 2^{-2n}$, где $\eta_0$ размещает массы $(2n+1)^{-1}$ в точках $(2\pi k/2n+1)$, $k = 0, 1, \ldots, 2n$, а $\xi_0$ — равномерная мера на $(0, 2\pi)$.

Для мер $\xi$, размещающих равные массы в $2n+1$ различных точках $\theta_0, \theta_1, \ldots, \theta_{2n}$, обратимся к (3.2) и запишем $|\mathbf{M}(\xi)|$ в виде

$$|\mathbf{M}(\xi)| = \left(\frac{1}{2n+1}\right)^{2n+1} [\det \| f_i(\theta_j) \|]^2.$$

Используя известное равенство

$$\det [\| f_i(\theta_j) \|] = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ \cos\theta_0 & \cos\theta_1 & \ldots & \cos\theta_{2n} \\ \sin\theta_0 & \sin\theta_1 & \ldots & \sin\theta_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ \sin n\theta_0 & \sin n\theta_1 & \ldots & \sin n\theta_{2n} \end{vmatrix} = 2^{2n} \prod_{\nu>\mu} \sin\left(\frac{\theta_\nu - \theta_\mu}{2}\right),$$

мы получаем, что

$$|\mathbf{M}(\xi)| = \frac{(2^{2n})^2}{(2n+1)^{2n+1}} D^2(\theta_0, \ldots, \theta_{2n}).$$

Доказательство завершается ссылкой на лемму 6.4.

Обратимся теперь к задаче Фейера для тригонометрических систем $1, \cos\theta, \sin\theta, \ldots, \cos n\theta, \sin n\theta$ $(0 \leqslant \theta \leqslant 2\pi)$. Пусть $\{\lambda_k(\theta)\}_0^{2n}$ обозначает систему тригонометрических многочленов степени $n$, обладающих свойством $\lambda_k(\theta_\nu) = \delta_{\nu k}$ при $k, \nu = 0, 1, \ldots, 2n$, где $\theta_0, \theta_1, \ldots, \theta_{2n}$ — различные точки в интервале $[0, 2\pi)$. Такая система называется *тригонометрической интерполяционной системой*.

В случае, когда $\theta_0, \theta_1, \ldots, \theta_{2n}$ определены как $\theta_k = (2\pi k/2n+1)$, $k = 0, 1, \ldots, 2n$, имеем

$$\lambda_k(\theta) = \sum_{\nu=0}^{2n} \varphi_\nu(\theta_k)\, \varphi_\nu(\theta),$$

где $\varphi_0, \varphi_1, \ldots, \varphi_{2n}$ определены в (6.8). Так как матрица коэффици ентов $\| \varphi_\nu(\theta_k) \|$ ортогональна, то мы получаем, что

$$\sum_{i=0}^{2n} [\lambda_i(\theta)]^2 = \sum_{i=0}^{2n} [\varphi_i(\theta)]^2 \equiv 1,$$

Рассуждение, аналогичное проведенному в теореме 4.2, доказывает, что если множество $2n+1$ точек в $[0, 2\pi)$ и ассоциированные с ним интерполяционные функции $\{\lambda_k(\theta)\}_{k=0}^{2n}$ удовлетворяют условию

$$\sup_{0 \leqslant \theta < 2\pi} \sum_{i=0}^{2n} [\lambda_i(\theta)]^2 = 1,$$

то мера, приписывающая равные массы каждой из точек $\theta_0, \theta_1, \ldots$

$\ldots, \theta_{2n}$, максимизирует $|\mathbf{M}(\xi)|$. Применяя теорему 6.4, мы заключаем, что решение задачи Фейера единственно с точностью до вращения интерполяционных точек. Этим завершается доказательство следующей теоремы.

Теорема 6.5. *Пусть* $\theta_0, \theta_1, \ldots, \theta_{2n}$ — *различные точки, принадлежащие* $[0, 2\pi)$, *а* $\lambda_k(\theta)$, $k = 0, 1, \ldots, 2n$, — *многочлен от* $1$, $\cos\theta$, $\sin\theta, \ldots, \cos n\theta$, $\sin n\theta$, *удовлетворяющий условиям* $\lambda_k(\theta_\nu) = \delta_{k\nu}$ *при* $k, \nu = 0, 1, \ldots, 2n$.

*Тогда минимум по* $\theta_0, \theta_1, \ldots, \theta_{2n}$ *выражения*

$$\sup_{0 \leqslant \theta < 2\pi} \sum_{k=0}^{2n} [\lambda_k(\theta)]^2$$

*достигается тогда и только тогда, когда*

$$\theta_k = \frac{2\pi k}{2n+1} + a, \qquad k = 0, 1, \ldots, 2n,$$

*где* $0 \leqslant a < \dfrac{2\pi}{2n+1}$ *и значение минимума равно единице.*

Используя теорему 6.5, нетрудно решить задачу отыскания устойчивой наиболее экономичной системы для тригонометрических многочленов. Если $\theta_0, \theta_1, \ldots, \theta_{2n}$ — $2n+1$ различных точек в $[0, 2\pi)$ и $\alpha_k(\theta)$, $k = 0, \ldots, 2n$ — тригонометрические многочлены произвольной степени, удовлетворяющие $\alpha_k(\theta_\nu) = \delta_{\nu k}$, система $\{\alpha_k(\theta)\}_0^{2n}$ называется *устойчивой*, если для всех $y_\nu \geqslant 0$, $\nu = 0, \ldots, 2n$ имеем

$$\min\{y_\nu\} \leqslant \sum_{k=0}^{2n} y_k \alpha_k(\theta) \leqslant \max\{y_\nu\}, \quad \theta \in [0, 2\pi].$$

Устойчивая система $\{\alpha_k(\theta)\}$ называется *наиболее экономичной*, если сумма степеней $\alpha_k(\theta)$ минимальна. Этот минимум ограничен снизу величиной $2n(2n+1)$.

Теорема 6.6. *Для системы* $1$, $\cos\theta$, $\sin\theta, \ldots, \cos n\theta$, $\sin n\theta$ *единственная устойчивая и наиболее экономичная система* $\{\alpha_k(\theta)\}_0^{2n}$ *задается равенствами* $\alpha_k(\theta) = \lambda_k^2(\theta)$, $k = 0, 1, \ldots, 2n$, *где* $\{\lambda_k(\theta)\}_0^{2n}$ — *тригонометрическая интерполяционная система, ассоциированная с точками* $\theta_k = (2\pi k/2n+1) + a$, $k = 0, 1, \ldots, 2n$, *где* $0 \leqslant a < (2\pi/2n+1)$.

Доказательство. Доказательство осуществляется с помощью теоремы 6.5. Рассуждения соответствуют основным этапам доказательства теоремы 5.1 и поэтому опускаются.

## § 7. Планы эксперимента

Исследования §§ 1—6 основывались на теории планирования эксперимента. Величины $d(x, \xi)$ и $\mathbf{M}(\xi)$, определенные в (2.5) и (2.1), имеют глубокий статистический смысл и играют важную роль в этой теории. Наша цель в настоящем параграфе состоит в том, чтобы изложить статистическую интерпретацию $d(x, \xi)$ и $\mathbf{M}(\xi)$ и проиллюстрировать, как геометрические методы, используемые на протяжении книги, помогают в получении оптимальных планов. В большинстве исследований настоящего параграфа мы следуем работам Кифера [1959], [1960], [1961] и Кифера и Вольфовица [1960].

Изучение оптимальных планов экспериментов, включающих статистические данные, приводит к следующей структуре. Пусть $f_0, f_1, \ldots, f_n$ обозначают непрерывные функции, определенные на компактном нространстве $X$. Точки $X$ представляют собой возможные уровни осуществимых экспериментов. Для каждого уровня $x \in X$ может быть осуществлен некоторый эксперимент, результатом которого является случайное наблюдение $y(x)$. Допустим, что наблюдение $y(x)$ имеет вид

$$y(x) = \sum_{j=0}^{n} \theta_j f_j(x) + \eta(x), \tag{7.1}$$

где $\eta(x)$ — случайная величина, такая, что

$$E(\eta(x)) = 0,$$

$$E\eta(x)\eta(x') = \begin{cases} 1, & x = x', \\ 0, & x \neq x', \end{cases}$$

а $E$ обозначает математическое ожидание соответствующей случайной величины. Функции $f_0, f_1, \ldots, f_n$, называемые *регрессионными* функциями, известны экспериментатору, в то время как параметры $\theta_0, \theta_1, \ldots, \theta_n$ неизвестны. Экспериментатор должен оценить значения параметров $\theta_0, \theta_1, \ldots, \theta_n$ или некоторую функцию от этих параметров на основе $N$ наблюдений (7.1), допуская возможность того, что разные наблюдения соответствуют разным уровням.

*План эксперимента* определяется как вероятностная мера, сосредоточивающая массы $p_1, \ldots, p_r$ в точках $x_1, \ldots, x_r$, где значения $p_i N = n_i$, $i = 1, \ldots, r$ — целые числа. Соответствующий эксперимент включает $n_i$ некоррелированных наблюдений случайных величин $y(x_i)$, $i = 1, \ldots, r$. После того как план задан и наблюдения проделаны, для оценки параметров $\theta_0, \theta_1, \ldots, \theta_n$ используется стандартная процедура.

Задача, с которой сталкивается экспериментатор, состоит в выборе плана, обладающего определенными оптимальными свойствами.

Статистические соображения (см. Кифер [1959]) придают интерес тем $\xi$, которые делают матрицу $\mathbf{M}(\xi)$ $\left(m_{ij} = \int_X f_i f_j d\xi\right)$ или определенную функцию от $\mathbf{M}(\xi)$ в некотором смысле большой. Мотивы для рассмотрения матрицы $\mathbf{M}(\xi)$ таковы. Если вектор неизвестных параметров $\theta = (\theta_0, \theta_1, \ldots, \theta_n)$ оценивается по методу наименьших квадратов, дающему наилучшую линейную несмещенную оценку, скажем $\hat{\theta}$, то ковариационная матрица $\hat{\theta}$ дается выражением

$$E(\hat{\theta}-\theta)(\hat{\theta}-\theta)' = \frac{1}{N}\mathbf{M}^{-1}(\xi), \tag{7.2}$$

где $\xi$ приписывает массу $p_i = n_i/N$ точке $x_i$, $i = 1, \ldots, r$.

Если матрица $[\mathbf{M}(\xi)]^{-1}$ «мала» в некотором смысле или $\mathbf{M}(\xi)$ «велико», то, грубо говоря, $\hat{\theta}$ близко к $\theta$, что выражается в «малости» (7.2). Большинство критериев оптимальности плана эксперимента основано на максимизации некоторого функционала от матрицы $\mathbf{M}(\xi)$, которая обычно называется *информационной* матрицей плана.

Информационная матрица, соответствующая плану $\xi$, который сосредоточивает всю массу в точке $x_0$, есть матрица единичного ранга $[\mathbf{f}(x_0)][\mathbf{f}(x_0)]'$, а если $\xi$ сосредоточивает массы $p_1, \ldots, p_r$ в точках $x_1, \ldots, x_r$, то информационная матрица есть

$$\sum_{i=1}^{r} p_i [\mathbf{f}(x_i)][\mathbf{f}(x_i)]'.$$

В дальнейшем мы будем предполагать, что $\xi$ — произвольная вероятностная мера на борелевских множествах $X$. Оправдание для допущения этой большей общности заключается в том, что информационная матрица произвольного $\xi$ может быть получена с помощью меры или плана, сосредоточенного в конечном числе точек.

Некоторые эквивалентные оптимальные процедуры. Простой критерий оптимальности плана состоит в том, чтобы выбирать план $\xi$, максимизирующий определитель информационной матрицы $\mathbf{M}(\xi)$. Эта процедура эквивалентна минимизации определителя ковариационной матрицы наилучших линейных несмещенных оценок для $\theta_0, \theta_1, \ldots, \theta_n$.

Другим возможным критерием оптимальности является минимизация максимума по $x$ дисперсии $N^{-1}d(x, \xi)$ наилучшей линейной оценки регрессионной функции $\sum_{i=0}^{n} \theta_i f_i(x)$. Функция $d(x, \xi)$ может быть записана в виде выражения

$$d(x, \xi) = (\mathbf{f}(x), \mathbf{M}^{-1}(\xi)\,\mathbf{f}(x)),$$

которое определено при условии, что матрица $\mathbf{M}(\xi)$ неособенная (величины $|\mathbf{M}(\xi)|$ и $d(x, \xi)$ те же, что и в § 2).

Первоначальным мотивом теоремы 2.1 было доказательство того, что два упомянутых критерия оптимальности эквивалентны.

Для удобства мы вновь сформулируем здесь теорему 2.1 в следующем виде.

Теорема 7.1. *Условия*

(1) $\xi^*$ *максимизирует* $|\mathbf{M}(\xi)|$,

(2) $\xi^*$ *минимизирует* $\max\limits_x d(x, \xi)$,

(3) $\max\limits_x d(x, \xi^*) = n + 1$

*эквивалентны. Множество* $B$ *всех* $\xi$, *удовлетворяющих этим условиям, выпукло и замкнуто и* $\mathbf{M}(\xi)$ *одно и то же для всех* $\xi \in B$.

Пример 7.1. Пусть $X$ обозначает $m$-симплекс $S_m$ точек $\mathbf{x} = (x_1, \ldots, x_{m+1})$, где $x_i \geqslant 0$, $i = 1, \ldots, m+1$ и $\sum\limits_{i=1}^{m+1} x_i = 1$, и пусть $\{f_i\}$ включает $(m+1)(m+2)/2$ функций: $1, x_r$ $(1 \leqslant r \leqslant m)$ и $x_r x_s$ $(1 \leqslant r \leqslant s \leqslant m)$. Пусть $S_{m,2}$ обозначает точки $S_m$ с двумя компонентами, равными $1/2$, и оставшимися компонентами, равными нулю.

Покажем, что план $\xi_{m,2}$, который приписывает массу $2/(m+1)\times(m+2)$ каждой из точек $S_{m,2}$ и вершинам $S_m$, является оптимальным в том смысле, что $\xi_{m,2}$ максимизирует $|\mathbf{M}(\xi)|$ или, эквивалентно, $\xi_{m,2}$ минимизирует $\sup\limits_x d(\mathbf{x}, \xi)$.

Пусть $\{g_i\}$ — система квадратичных многочленов

$$[2(m+1)(m+2)]^{1/2} x_i (x_i - 1/2), \qquad 1 \leqslant i \leqslant m+1,$$

$$[8(m+1)(m+2)]^{1/2} x_i x_j, \qquad 1 \leqslant i < j \leqslant m+1.$$

Непосредственный подсчет показывает, что написанная система ортонормальна по отношению к мере $\xi_{m,2}$. В этом случае $d(x, \xi_{m,2})$ может быть выражено как сумма квадратов от $\{g_i\}$. Простые преобразования приводят к соотношению

$$\frac{2}{(m+1)(m+2)} d(\mathbf{x}, \xi_{m,2}) = 1 - \sum_{i \neq j} x_i x_j \{2(x_i - x_j)^2 + (1 - 4x_i x_j)\}.$$

Так как величина в скобках неотрицательна, имеем

$$d(\mathbf{x}, \xi_{m,2}) \leqslant \frac{(m+1)(m+2)}{2}.$$

Сравнивая с теоремой 7.1 (3), мы заключаем, что $\xi_{m,2}$ оптимально.

Допустимые планы. Мера или план $\xi$ называются *допустимыми*, если не существует плана $\xi^*$ такого, что $\mathbf{M}(\xi^*) \geqslant \mathbf{M}(\xi)$ и $\mathbf{M}(\xi^*) \neq \mathbf{M}(\xi)$. В этом случае мы выбираем матрицу $\mathbf{M}(\xi)$ «большой» в смысле упорядочения, задаваемого положительной определенностью.

Если пространство $X$ есть интервал $[a, b]$ и $f_i(x) = x^i$, $i = 0, \ldots$ $\ldots, n$, то матрица $\mathbf{M}(\xi)$ образована элементами

$$m_{ij}(\xi) = \mu_{i+j}(\xi) = \int_a^b x^{i+j}\xi(dx),$$

и мы имеем следующий результат.

Теорема 7.2. $\mathbf{M}(\xi^*) \geqslant \mathbf{M}(\xi)$ *и* $\mathbf{M}(\xi^*) \neq \mathbf{M}(\xi)$ *тогда и только тогда, когда* $\mu_r(\xi^*) = \mu_r(\xi)$ *при* $0 \leqslant r \leqslant 2n-1$ *и* $\mu_{2n}(\xi^*) > \mu_{2n}(\xi)$.

Доказательство. Если $\mathbf{M}(\xi^*) \geqslant \mathbf{M}(\xi)$ *и* $\mathbf{v} = (v_0, \ldots, v_n)$, то необходимо

$$(\mathbf{v}, (\mathbf{M}(\xi^*) - \mathbf{M}(\xi))\,\mathbf{v}) = \sum_{i,j=1}^{n} v_i v_j (\mu^*_{i+j} - \mu_{i+j}) \geqslant 0.$$

Положим $v_0 = 1$, $v_j = v$ и $v_i = 0$, $i \neq 0, j$, для каждого фиксированного $j = 0, 1, \ldots, n$. Тогда, так как $\mu_0^* = \mu_0$, то выписанная выше квадратичная форма сводится к

$$2v(\mu_j^* - \mu_j) + v^2(\mu_{2j}^* - \mu_{2j}) \geqslant 0;$$

это неравенство справедливо при любом вещественном $v$. Отсюда следует, что $\mu_j^* = \mu_j$, $j = 0, 1, \ldots, n$.

Повторяя то же самое рассуждение при $v_p = 1$, $v_q = v$ $(v_i = 0$, $i \neq p$ или $q)$ и $p < q$ для последовательных значений $p = 1, \ldots$ $\ldots, n-1$, мы получим, что $\mu_j^* = \mu_j$ при $j = 0, \ldots, 2n-1$. Так как $\mathbf{M}(\xi^*) \neq \mathbf{M}(\xi)$, то заключаем, что $\mu_{2n}^* > \mu_{2n}$. Обратное утверждение очевидно.

Обращаясь к теореме 7.2 и результатам предыдущих глав (ср. теоремы 2.1 гл. II и 1.1 гл. III), мы можем вывести следующий результат.

Теорема 7.3. *Класс допустимых планов эксперимента состоит из мер, сосредоточенных на множестве индекса, не большего чем* $n - 1/2$, *и мерах индекса* $n$, *которые соответствуют верхним главным представлениям для некоторой моментной точки* $(\mu_0, \mu_1, \ldots, \mu_{2n-1})$.

Доказательство. Вектор $(\mu_0, \mu_1, \ldots, \mu_{2n-1})$ должен быть граничной точкой пространства моментов, порожденного $\{t_i\}_0^{2n-1}$ на $[a, b]$, или это есть внутренняя точка, а $(\mu_0, \mu_1, \ldots, \mu_{2n})$ лежит на верхней границе пространства, порожденного $\{t_i\}_0^{2n}$.

В первом случае соответствующая единственная мера имеет индекс, не больший чем $n - 1/2$, а во втором — мера соответствует верхнему главному представлению.

Ротатабельные планы. Пусть $X$ обозначает единичную $m$-мерную сферу, образованную векторами $x = (x_1, \ldots, x_m)$, для которых $\sum_1^m x_i^2 \leqslant 1$, и пусть $f_i$, $i = 0, 1, \ldots, n$, есть функции $\prod_{j=1}^m x_j^{\alpha_j}$, где $\alpha_j$ — неотрицательные целые числа и $\sum_1^m \alpha_j \leqslant d$. Из комбинаторики известно, что $n + 1 = \binom{d+m}{m}$. Задача построения оптимального плана при оценивании регрессионной функции в этом случае будет называться задачей регрессии степени $d$ на $m$-мерном шаре.

Пусть $\mathcal{G}$ — ортогональная группа линейных преобразований в $R^m$. Если $\mathbf{f}(\mathbf{x})$ представляет собой $(n+1)$-вектор-функцию с компонентами $f_i$, то $\mathbf{f}(g\mathbf{x}) = \mathbf{U}_q \mathbf{f}(\mathbf{x})$, где $\mathbf{U}_g$ есть $(n+1)\times(n+1)$-матрица, зависящая от $g \in G$ и не зависящая от $\mathbf{x}$ и $\mathbf{f}$. Более того, так как $\mathbf{f}(\mathbf{x}) = \mathbf{f}(gg^{-1}\mathbf{x}) = \mathbf{U}_g \mathbf{U}_g^{-1} \mathbf{f}(\mathbf{x})$ и функции $f_i$ линейно независимы, то отсюда следует, что матрица $\mathbf{U}_g$ неособенная и $\mathbf{U}_g^{-1} = \mathbf{U}_{g-1}$.

Пользуясь определением $\mathbf{M}(\xi)$, можно проверить, что $M(\bar{g}\xi) = \mathbf{U}_g^{-1} \mathbf{M}(\xi) \mathbf{U}_g'^{-1}$. Следовательно, $d(g\mathbf{x}, \xi) = d(\mathbf{x}, \bar{g}\xi)$ (см. (6.6) и (6.7)). Теорема 6.3 теперь гарантирует нам существование ортогонально-инвариантного оптимального плана для регрессии степени $d$ на $m$-мерном шаре.

Всякая ортогонально-инвариантная мера $\xi$ может быть записана в виде

$$\xi(A) = \int_0^1 \xi_1 (r^{-1}[A \cap S_r]) \rho(dr),$$

где $S_r$ — $(m-1)$-мерная сфера радиуса $r$, $\xi_1$ — равномерная вероятностная мера на $S_1$, $\rho$ — вероятностная мера на $[0,1]$ и подынтегральное выражение при $r = 0$ берется равным единице, если $0 \in A$, и нулю, если $0 \notin A$.

Если $\mathcal{M}_{2d}$ определяется как

$$\mathcal{M}_{2d} = \left\{(1, \mu_2, \mu_4, \ldots, \mu_{2d}) \mid \mu_{2j} = \int_0^1 r^{2j} \rho(dr), \int_0^1 \rho(dr) = 1\right\},$$

то имеем следующую теорему.

Теорема 7.4. *Существует единственный ортогонально-инвариантный план* $\xi_0$, *который максимизирует* $|\mathbf{M}(\xi)|$. *Радиальная компонента* $\rho_0$ *плана* $\xi_0$ *сосредоточена на множестве индекса* $d/2$ *и соответствует верхнему главному представлению некоторой точки в* $\mathcal{M}_{2(d-1)}$, *т. е.* $\rho_0$ *приписывает положительную массу радиусам* $\lambda_1 < \lambda_2 < \ldots < \lambda_{q+1} = 1$, *где* $q = [d/2]$. *Кроме того,* $\lambda_1 = 0$, *если* $d$ *четно, и* $\lambda_1 > 0$, *если* $d$ *нечетно.*

Замечание 7.1. Так как $\mathbf{M}(\xi)$ — одно и то же для любого оптимального плана $\xi$, то $d(\mathbf{x}, \xi) = d(\mathbf{x}, \xi_0) = d(\mathbf{x}, \bar{g}\xi_0) = d(g\mathbf{x}, \xi_0) =$

$= d(g\mathbf{x}, \xi)$, и, таким образом, для всякого оптимального плана $\xi$, $d(\mathbf{x}, \xi)$ зависит только от $r$.

Доказательство. Покажем сначала, что радиальная часть $\rho_0$ максимизирующей инвариантной меры $\xi_0$ не может быть сосредоточена на множестве индекса, меньшего чем $d/2$. Если $d$ — нечетное, скажем $d = 2l + 1$, и $\rho_0$ сосредоточено на множестве индекса, меньшего чем $d/2$, то мы можем принять без потери общности, что этот индекс равен $l$. Мера $\rho_0$ сосредоточена на $l$ внутренних точках $r_1, \ldots, r_l$ или $l - 1$ внутренних точках $r_1, \ldots, r_{l-1}$ и на обоих концах 0 и 1. В первом случае мы определяем $a_1, \ldots, a_n$ уравнением

$$\prod_1^l (r^2 - r_i^2) = \prod_1^l (x_1^2 + \ldots + x_m^2 - r_i^2) = \sum_0^n a_i f_i(x)$$

и в последнем случае уравнением

$$x_1 (r^2 - 1) \prod_1^{l-1} (r^2 - r_1^2) = \sum_0^n a_i f_i(x).$$

В обоих случаях получаем, что $\sum_{i=0}^n a_i m_{ij}(\xi_0) = 0$ для всех $j$, так что $|\mathbf{M}(\xi_0)| = 0$. Случай четного $d$ исследуется подобным образом. Мы доказали, таким образом, что $\rho_0$ сосредоточено на множестве индекса не меньше $d/2$.

Так как матрицы $\mathbf{M}(\xi_0)$ совпадают для всех оптимальных $\xi_0$, радиальные части $\rho_0$ с необходимостью приводят к одному и тому же множеству моментов $\mu_{2i}^0 = \int_0^1 r^{2i} \rho_0(dr)$, $i = 1, 2, \ldots, d$. Так как $\rho_0$ сосредоточено на множестве индекса не менее $d/2$, то нам известно из теоремы 2.1 гл. II, что моментная точка $(1, \mu_2^0, \mu_4^0, \ldots$ $\ldots, \mu_{2(d-1)}^0)$ есть внутренняя точка $\mathcal{M}_{2(d-1)}$. Мы теперь покажем, что $|\mathbf{M}(\xi)|$ есть возрастающая функция момента $\mu_{2d}$ при условии, что радиальная часть $\rho$ меры $\xi$ имеет фиксированное число $d - 1$ моментов $\int_0^1 r^{2i} \rho(dr) = \mu_{2i}^0$, $i = 1, \ldots, d - 1$. Если это доказано, то, применяя теорему 1.1 гл. III, заключаем, что мера $\rho_0$ должна быть единственным верхним главным представлением точки $(1, \mu_2^0, \ldots$ $\ldots, \mu_{2d-2}^0)$ в $\mathcal{M}_{2(d-1)}$, и доказательство теоремы закончено.

Осталось показать, что $|\mathbf{M}(\xi)|$ возрастает по $\mu_{2d}$. Определим $c(\alpha_1 + \beta_1, \ldots, \alpha_m + \beta_m)$ для любой инвариантной меры $\xi$ с помощью уравнения

$$\int x_1^{\alpha_1} \ldots x_m^{\alpha_m} x_1^{\beta_1} \ldots x_m^{\beta_m} d\xi = c(\alpha_1 + \beta_1, \ldots, \alpha_m + \beta_m)\, \mu_{\Sigma(\alpha_i + \beta_i)}$$

и рассмотрим квадратичную форму $(\mathbf{z}, \mathbf{M}(\xi)\mathbf{z})$, где $\mathbf{z}$ обозначает $(n+1)$-вектор с компонентами $z(\alpha_1, \ldots, \alpha_m)$, упорядоченными тем же способом, что и функции $f_i(x) = \prod_j x_j^{\alpha_j}$. Имеем

$$(\mathbf{z}, \mathbf{M}(\xi)\mathbf{z}) = \sum_{j=0}^{d-1} \mu_{2j}^0 p_{2j} + \mu_{2d} p_{2d},$$

где

$$p_{2j} = \Sigma z(\alpha_1, \ldots, \alpha_n) z(\beta_1, \ldots, \beta_m) c(\alpha_1 + \beta_1, \ldots, \alpha_m + \beta_m)$$

и сумма распространяется по всем $\alpha_i$, $\beta_i$ таким образом, что $\Sigma\alpha_i \leqslant \leqslant d$, $\Sigma\beta_i \leqslant d$ и $\Sigma(\alpha_i + \beta_i) = 2j$. Используя тот факт, что $|\mathbf{M}(\xi)| > > 0$, получаем, что коэффициенты $p_{2d}$ при $\mu_{2d}$ задают положительно определенную форму относительно переменных $z(\beta_1, \ldots, \beta_m)$, где $(\beta_1, \ldots, \beta_m)$ принадлежат множеству

$$A_0 = \left\{(\beta_1, \ldots, \beta_m) \,\Big|\, \sum_i (\alpha_i + \beta_i) = 2d\right.$$

для некоторого $(\alpha_1, \ldots, \alpha_m)\}$. Следовательно, $|\mathbf{M}(\xi)|$ строго возрастает по $\mu_{2d}$, что заканчивает доказательство.

Доказанная теорема показывает, что при $d = 1$ оптимальный инвариантный план равномерно распределен на сфере единичного радиуса. При $d = 2$ оптимальный план сосредоточивает массу $(1 - - \delta)$ в начале координат и $\delta$ — на сфере единичного радиуса. Стандартные вычисления приводят к соотношению $1 - \delta = 2/(m + + 1)(m + 2)$.

Мы теперь определим $\mathscr{B}$ как класс ортогонально-инвариантных мер $\xi$, у которых радиальная часть либо сосредоточена на множестве индекса, не большего чем $(d - 1)/2$, либо сосредоточена на множестве индекса $d/2$ и соответствует верхнему главному представлению некоторой точки в $\mathscr{M}_{2(d-1)}$. Применяя метод доказательства теоремы 7.3, мы получаем следующую теорему.

Теорема 7.5. *В классе ортогонально-инвариантных планов множество допустимых планов совпадает с* $\mathscr{B}$.

Полиномиальная регрессия. Планы эксперимента для случая, когда регрессионные функции $f_0, f_1, \ldots, f_n$ в (7.1) имеют вид $\sqrt{w(x)}\, x^i$, $i = 0, 1, \ldots, n$ $(w(x) \geqslant 0)$, представляют значительный практический интерес. При специальном выборе $w(x)$ оптимальные планы (оптимальность понимается в том смысле, что план $\xi$ максим зирует $|\mathbf{M}(\xi)|$) получаются с помощью теорем § 3. Например, если $w(x) = (1 - x)^{\alpha+1}(1 + x)^{\beta+1}$ $(\alpha > -1,\ \beta > -1)$, а $X$ есть отрезок $[-1, 1]$, то вид оптимального плана определяется на основе теоремы 3.2. А именно, план $\xi$ состоит из $n + 1$ равных точечных масс, помещенных в $n + 1$ нулях многочлена Якоби $P_{n+1}^{(\alpha,\beta)}(x)$.

Мы закончим данный параграф рассмотрением следующей задачи.

Пусть весовая функция $w(x) \equiv 1$ и $X = [-1, 1]$. Мы хотим минимизировать $d(x_0, \xi)$ в фикс рованной точке $x_0$, где $|x_0| > 1$ и охарактеризовать соответствующий оптимальный план. Это значит, что дисперсия $N^{-1}d(x_0, \xi)$ ($N$ есть число наблюдений) предсказанного значения полиномиальной регрессионной кривой $\sum_{i=0}^{n} \theta_i x_0^i$ должна быть минимизирована в точке $x_0$, вне интервала $[-1, 1]$, на котором выбраны наблюдения. Эта задача называется задачей экстраполяции.

Полное решение задачи дано в работе Хоэла и Левина [1964]. Их метод отличается от того, который дается ниже.

Теорема 7.6. *Если* $|x_0| > 1$ *и* $\xi$ *обозначает вероятностную меру на* $[-1, 1]$, *то*

$$d(x_0, \xi) \geqslant T_n^2(x_0), \tag{7.3}$$

*где* $T_n(x)$ *обозначает* $n$*-й многочлен Чебышева первого рода. Кроме того, если*

$$s_\nu = -\cos\frac{\nu\pi}{n}, \qquad \nu = 0, 1, \ldots, n, \tag{7.4}$$

*и* $L_\nu(x)$, $\nu = 0, 1, \ldots, n$,*—интерполяционные многочлены Лагранжа степени* $n$ *для точек* $s_\nu$, $\nu = 0, 1, \ldots, n$ (*т. е.* $L_\nu(s_j) = \delta_{\nu j}$, $\nu, j = 0, 1, \ldots, n$), *то равенство достигается в* (7.3) *тогда и только тогда, когда* $\xi$ *сосредоточивает массу*

$$\lambda_\nu^0 = \frac{|L_\nu(x_0)|}{\sum_{i=0}^{n} |L_i(x_0)|} \tag{7.5}$$

*в точках* $s_\nu$, $\nu = 0, 1, \ldots, n$.

Доказательство этой теоремы опирается на следующую лемму.

Лемма 7.1. *Пусть* $\xi$ *обозначает вероятностную меру на* $[-1, 1]$, *и пусть* $P_n$ *обозначает многочлен степени не более* $n$ *такой, что* $|P_n(x_0)|^2 = 1$ $(|x_0| > 1)$. *Тогда*

$$\sup_\xi \inf_{P_n} \int_{-1}^{1} |P_n(x)|^2 \xi(dx) = \frac{1}{|T_n(x_0)|^2}. \tag{7.6}$$

*Кроме того,*

$$\inf_{P_n} \int_{-1}^{1} |P_n(x)|^2 \xi(dx) = \frac{1}{|T_n(x_0)|^2}$$

*в том и только в том случае, когда* $\xi$ *такое, как в теореме* 7.6.

Замечание 7.2. Приведенная лемма также имеет приложения в связи с методом сопряженных градиентов для решения линейных уравнений и методом минимальной итерации для вычисления наибольшего собственного числа симметричной матрицы (см. работу Каниэля [1966]).

Доказательство теоремы 7.6. Пусть лемма 7.1 справедлива. Мы хотим минимизировать $d(x_0, \xi) = (\mathbf{f}(x_0), \mathbf{M}^{-1}(\xi)\mathbf{f}(x_0))$ над классом вероятностных мер на $[-1, 1]$, где $\mathbf{M}(\xi) = \|m_{ij}(\xi)\|$,

$$m_{ij} = \int_{-1}^{1} x^{i+j}\xi(dx) \text{ и } \mathbf{f}(x) = (1, x, \ldots, x^n).$$

Так как

$$\begin{gathered} d(x_0, \xi) = \mathbf{f}(x_0)\mathbf{M}^{-1}(\xi)\mathbf{f}(x_0)) = \sup_{b \neq 0} \frac{(\mathbf{f}(x_0), \mathbf{b})^2}{(\mathbf{b}, \mathbf{M}(\xi)\mathbf{b})}, \\ \mathbf{b} = (b_0, \ldots, b_n), \end{gathered} \tag{7.7}$$

мы получаем, что

$$\inf_{\xi} d(x_0, \xi) = \inf_{\xi} \sup_{P_n \neq 0} \frac{|P_n(x_0)|^2}{\int_{-1}^{1} |P_n(x)|^2 \xi(dx)}, \tag{7.8}$$

где $P_n(x) = \sum_{i=0}^{n} b_i x^i$. Однако правая часть (7.8), очевидно, равна

$$\left(\sup_{\xi} \inf_{P_n \neq 0} \frac{\int_{-1}^{1} |P_n(x)|^2 \xi(dx)}{|P_n(x_0)|^2}\right)^{-1}.$$

Отсюда на основе леммы 7.1 следует, что единственный оптимальный план $\xi$ таков, как описано в теореме 6.7, и $\min_{\xi} d(x_0, \xi) = T_n^2(x_0)$.

Прежде чем приступить к доказательству леммы 7.1, выведем вспомогательную формулу для $d(x, \xi)$.

Сначала отметим, что элементы матрицы $\mathbf{M}(\xi)$ можно рассматривать как точку в $(2n+1)$-мерном евклидовом пространстве, порожденном $T$-системой $1, x, x^2, \ldots, x^{2n}$. В этом случае для всякого плана $\xi$, для которого $|\mathbf{M}(\xi)| > 0$, существует план $\tilde{\xi}$, помещающий массу в $n+1$ точках, для которых $\mathbf{M}(\xi) = \mathbf{M}(\tilde{\xi})$ (см. следствие 3.1 гл. II). Следовательно, мы можем всегда ограничить внимание планами, сосредоточенными в $n+1$ точках.

Как в (7.7) и (7.8), мы запишем

$$d(x, \xi) = \max_{P_n \neq 0} \frac{|P_n(x)|^2}{\int_{-1}^{1} |P_n(y)|^2 \xi(dy)} = \max_{P_n} \frac{|P_n(x)|^2}{\sum_{\nu=0}^{n} \lambda_\nu |P_n(t_\nu)|^2},$$

где $P_n$ обозначает ненулевой многочлен степени не больше $n$, $\{t_\nu\}_{\nu=0}^{n}$ — $(n+1)$ значений, выбранных в интервале $[-1, 1]$ и $\lambda_\nu$ — положительные числа, сумма которых равна единице.

Пусть $L_\nu(x)$, $\nu = 0, 1, \ldots, n$, — интерполяционные многочлены Лагранжа, ассоциированные с $t_\nu$, $\nu = 0, 1, \ldots, n$ (т. е. $L_\nu(x)$ — многочлен степени $n$, удовлетворяющий $L_\nu(t_j) = \delta_{\nu j}$).

Тогда

$$P_n(x) = \sum_{\nu=0}^{n} P_n(t_\nu) L_\nu(x).$$

По неравенству Шварца имеем

$$|P_n(x)|^2 = \left| \sum_{\nu=0}^{n} \sqrt{\lambda_\nu} P_n(t_\nu) \frac{L_\nu(x)}{\sqrt{\lambda_\nu}} \right|^2 \leqslant$$

$$\leqslant \sum_{\nu=0}^{n} \lambda_\nu P_n^2(t_\nu) \sum_{\nu=0}^{n} \frac{L_\nu^2(x)}{\lambda_\nu}. \qquad (7.9)$$

Поскольку равенство для некоторого многочлена достигается (фиксируем $x$ и просто определим $P_n(t)$ так, что $P_n(t_\nu) = c(L_\nu(x)/\lambda_\nu)$, где $c$ — константа), то мы получаем, что

$$d(x, \xi) = \sum_{\nu=0}^{n} \frac{L_\nu^2(x)}{\lambda_\nu}, \qquad (7.10)$$

где $\xi$ обозначает план, приписывающий массу $\lambda_\nu$ точке $x_\nu$, $\nu = 0, 1, \ldots, n$.

Доказательство леммы 7.1. Рассмотрим игру (см. замечание 2.1 к § 2) $(\Xi, \Theta, K)$, где $\Xi$ обозначает класс вероятностных мер $\xi$ на $[-1, 1]$, $\Theta$ обозначает класс неотрицательных многочленов степени $2n$, нормированных так, что $P_{2n}(x_0) = 1$, а ядро $K$ определено на $\Xi \times \Theta$ посредством выражения

$$K(\xi, P_{2n}) = \int_{-1}^{1} P_{2n}(x)\, \xi(dx).$$

Очевидно, что $K$ билинейно по двум своим аргументам. Обычное доказательство компактности аппроксимации $\Theta$ с помощью $\Theta_N$, состоящего из всех многочленов $P_{2n}$, коэффициенты которых ограничены $N$ по абсолютной величине (см. теорему 2.1), показывает, что игра определена и

$$v = \min_{P_{2n}} \max_{-1 \leqslant x \leqslant 1} P_{2n}(x) = \max_{\xi} \min_{P_{2n}} \int_{-1}^{1} P_{2n}(x)\, \xi(dx). \qquad (7.11)$$

Мы утверждаем, что единственный оптимальный минимизирующий многочлен есть $T_n^2(x)/T_n^2(x_0)$. В противном случае мы легко выводим, что многочлен $P_{2n}(x) - T_n^2(x)/T_n^2(x_0)$ имеет нуль в точке $x_0$ и $2n$ нулей в $[-1, 1]$, что невозможно. Итак,

$$v = \min_{P_{2n}} \max_{-1 \leqslant x \leqslant 1} P_{2n}(x) = \max_{-1 \leqslant x \leqslant 1} \frac{T_n^2(x)}{T_n^2(x_0)} = \frac{1}{T_n^2(x_0)}$$

и $T_n^2(x)/T_n^2(x_0)$ — единственная оптимальная для игрока II стратегия. В частности,

$$\frac{\int_{-1}^{1} P_n^2(x)\,\xi(dx)}{P_n^2(x_0)} \leqslant \frac{\int T_n^2(x)\,\xi(dx)}{T_n^2(x_0)} \leqslant \frac{1}{T_n^2(x_0)} = v$$

при всех $\xi \in \Xi$.

Ясно, что оптимальный план $\xi^*$ сосредоточен в точках, где $T_n^2(x) = 1$, т. е. в точках $s_\nu = -\cos \nu\pi/n$, $\nu = 0, 1, \ldots, n$. Если $\lambda_\nu$, $\nu = 0, 1, \ldots, n$, обозначают соответствующие веса, то оптимальность плана $\xi^*$ влечет, что

$$\sum_{\nu=0}^{n} P_n^2(s_\nu)\,\lambda_\nu \geqslant \frac{1}{T_n^2(x_0)} \tag{7.12}$$

для всех многочленов $P_n$ таких, что $P_n^2(x_0) = 1$. Пусть $L_\nu$ обозначают интерполяционные многочлены Лагранжа, ассоциированные с точками $\{s_\nu\}_{\nu=0}^{n}$.

Тогда из (7.9) получаем

$$1 = P_n^2(x_0) \leqslant \left(\sum_{\nu=0}^{n} \lambda_\nu P_n^2(s_\nu)\right)\left(\sum_{\nu=0}^{n} \frac{L_\nu^2(x_0)}{\lambda_\nu}\right),$$

и, следовательно,

$$\sum_{\nu=0}^{n} \lambda_\nu P_n^2(s_\nu) \geqslant \left(\sum_{\nu=0}^{n} \frac{L_\nu^2(x_0)}{\lambda_\nu}\right)^{-1}. \tag{7.13}$$

Максимум по $\lambda_\nu$ $\left(\lambda_\nu \geqslant 0,\ \sum_{\nu=0}^{n} \lambda_\nu = 1\right)$ правой части (7.13) легко оценивается с помощью множителей Лагранжа и достигается в единственном случае при

$$\lambda_\nu^0 = \frac{|L_\nu(x_0)|}{\sum_{i=0}^{n} |L_i(x_0)|}, \qquad \nu = 0, 1, \ldots, n. \tag{7.14}$$

Отметим, что

$$\sum_{\nu=0}^{n} \frac{L_\nu^2(x_0)}{\lambda_\nu^0} = \left(\sum_{\nu=0}^{n} |L_\nu(x_0)|\right)^2.$$

Кроме того, $T_n(s_\nu) = (-1)^{n-\nu}$, так что $T_n(x_0) = \sum_{\nu=0}^{n} (-1)^{n-\nu} \times$ $\times L_\nu(x_0)$ и, так как знаки $L_\nu(x_0)$, $\nu = 0, 1, \ldots, n$, чередуются, имеем $|T_n(x_0)|^2 = \left(\sum_{\nu=0}^{n} |L_\nu(x_0)|\right)^2$. Таким образом, при $\lambda_\nu = \lambda_\nu^0$ равенство в (7.12) достигается, и, следовательно, множество $\{\lambda_\nu^0\}_0^n$ образует оптимальное множество весов для $\xi^*$.

При любом другом задании $\lambda_\nu$ имеем

$$\left(\sum_{\nu=0}^{n} \frac{L_\nu^2(x_0)}{\lambda_\nu}\right)^{-1} < \frac{1}{T_n^2(x_0)},$$

откуда видно, что оптимальные веса $\lambda_\nu^0$ определены единственным образом. Доказательство леммы закончено.

Главным и неожиданным аспектом теоремы 7.6 является то, что оптимальный план для оценивания линейной формы $(\theta, f(x_0)) = \sum_{i=0}^{n} \theta_i x_0^i$ сосредоточен на фиксированном множестве $n+1$ точек $-1 = s_0 < s_1 < \ldots < s_n = 1$ независимо от выбора $x_0$ ($|x_0| > 1$). Точки $\{s_i\}_0^n$ удовлетворяют $|T_n(s_i)| = 1$, $i = 0, 1, \ldots, n$, а многочлен Чебышева $T_n(x)$ характеризуется (с точностью до постоянного множителя) тем свойством, что он минимизирует величину

$$\sup_{-1 \leqslant x \leqslant 1} \left| x^n + \sum_{i=0}^{n-1} a_i x^i \right|.$$

В работе Кифера и Вольфовица [1965] теорема 7.6 была обобщена заменой обычных степеней $1, x, \ldots, x^n$ на $T$-систему функций $f_0, f_1, \ldots, f_n$. При подходящих ограничениях на функции $f_0, f_1, \ldots$ $\ldots, f_n$ наилучшая аппроксимация $f_n$ многочленами $\sum_{i=0}^{n-1} a_i f_i$ снова достигает максимума своего абсолютного значения в точности в $n+1$ точке $\{s_i^*\}_0^n$.

Ниже в теореме 7.7 мы сформулируем ту часть результатов Кифера и Вольфовица, которая является прямым обобщением теоремы 7.6.

Пусть $\{f_i\}_0^n$ образуют $T$-систему на $[-1, 1]$, и пусть $\underline{u}$ обозначает единственнный многочлен (см. теорему 10.2 гл. II), обладающий свойствами:

(1) $|\underline{u}(x)| \leqslant 1, \quad x \in [-1, 1]$,

(2) существуют $n+1$ точек $-1 \leqslant s_0 < s_1 < \ldots < s_n \leqslant 1$ таких, что $\underline{u}(s_{n-i}) = (-1)^i, \; i = 0, 1, \ldots, n$.

Далее пусть $\mathcal{C}$ обозначает множество векторов $\mathbf{c} = (c_0, c_1, \ldots, c_n)$, для которых

$$\begin{vmatrix} f_0(x_1) & \cdot\ \cdot & f_0(x_n) & c_0 \\ f_1(x_1) & \ldots & f_1(x_n) & c_1 \\ \cdot & \cdot\ \cdot\ \cdot & \cdot & \cdot \\ f_n(x_1) & \ldots & f_n(x_n) & c_n \end{vmatrix} \neq 0 \tag{7.15}$$

для всех $\{x_i\}_1^n$, удовлетворяющих $-1 \leqslant x_1 < x_2 < \ldots < x_n \leqslant 1$. В дальнейшем будет удобно связывать с каждым $\mathbf{c} \in \mathcal{C}$ точку $z$ и положить $f_i(z) = c_i, \quad i = 0, 1, \ldots, n$. Для любого многочлена $u = \sum_{i=0}^n a_i f_i$ мы можем теперь написать $u(z) = \sum_{i=0}^n a_i f_i(z) = \sum_{i=0}^n a_i c_i$. Отметим, что определители (7.15) необходимо сохраняют постоянный знак, так что условия принадлежности $\mathbf{c}$ к $\mathcal{C}$ определяют просто то обстоятельство, что функции $f_0, f_1, \ldots, f_n$ образуют $T$-систему на $[-1, 1] \cup \{z\}$.

Теорема 7.7. *Пусть $\{f_i\}_0^n$ — $T$-система на $[-1, 1]$, и пусть $\underline{u}$ и $\mathcal{C}$ таковы, как определено выше.*

(a) *Если $\mathbf{c} \in \mathcal{C}$ и $\xi$ обозначает вероятностную меру на $[-1, 1]$, то*

$$(\mathbf{c}, \mathbf{M}^{-1}(\xi)\,\mathbf{c}) \geqslant \underline{u}^2(z) \tag{7.16}$$

$(f_i(z) = c_i, \quad i = 0, 1, \ldots, n)$.

(b) *Пусть $s_0^* < s_1^* < \ldots < s_n^*$ — $n+1$ точка множества $B = \{x \mid \underline{u}^2(x) = 1\}$, причем знаки величин $\underline{u}(s_0^*), \ldots, \underline{u}(s_n^*)$ чередуются, и пусть $L_\nu(x) = \sum_{i=0}^n a_i^{(\nu)} f_i(x)$ обозначают интерполяционные многочлены Лагранжа для точек $\{s_i^*\}_0^n$, т. е. $L_\nu(s_j^*) = \delta_{\nu j}$, $\nu, j = 0, 1, \ldots, n$. Тогда равенство достигается в (7.16), если $\xi = \xi^*$ приписывает точкам $s_i$ массу*

$$\lambda_i^* = |L_i(z)| \Big/ \sum_{j=0}^n |L_j(z)|, \qquad i = 0, 1, \ldots, n.$$

(c) *Если $\sum_{i=0}^n a_i f_i(x) \equiv 1, \quad x \in [-1, 1]$ для некоторого множества коэффициентов $\{a_i\}_0^n$, то $B$ состоит в точности из $n+1$*

*точек и равенство в* (7.16) *достигается тогда и только тогда, когда* $\xi^*$ *определено как в части* (b).

Рассуждения, используемые при доказательстве теоремы 7.7, аналогичны рассуждениям, уже проделанным для теоремы 7.6 и, поэтому мы их опускаем.

## § 8. Обобщения

Исследования этого параграфа обобщают работу Кифера [1961]. Пусть $f = (f_0, f_1, \ldots, f_n)$ — векторнозначная функция, образованная из $n+1$ непрерывных функций, определенных на компактном пространстве $X$. В § 2 мы рассматривали две функции $\mathbf{M}(\xi) = \|m_{ij}\|$ и $d(x, \xi)$, где

$$m_{ij} = \int\limits_X f_i f_j \, d\xi, \qquad i, j = 0, 1, 2, \ldots, n,$$

$$d(x, \xi) = (f(x), \mathbf{M}^{-1}(\xi) f(x)).$$

Теорема 2.1 показывает, что всякая мера $\xi$, максимизирующая $|\mathbf{M}(\xi)|$, минимизирует также $\max\limits_x d(x, \xi)$ и наоборот. Обращениие к этой теореме было продиктовано желанием отождествить оптимальные планы, соответствующие, по-видимому, различным статистическим критериям. В статистической терминологии § 7 $\mathbf{M}(\xi)$ — информационная матрица, ассоциированная с экспериментом $\xi$, а $d(x, \xi)$ при ф ксированном $\xi$ соответствует (с точностью до постоянного множителя) дисперсии наилучшей линейной несмещенной оценки регрессионной функции $\sum\limits_{i=0}^{n} \theta_i f_i(x)$.

Функции $\mathbf{M}(\xi)$ и $d(x, \xi)$ соответствуют задаче оценивания полного множества $n+1$ параметров $\theta_0, \theta_1, \ldots, \theta_n$. Для рассмотрения оценивания только первых $s+1$ из $n+1$ параметров мы вводим две соответствующие функции. Если $\mathbf{M}(\xi)$ — неособенная матрица, то аналогом $|\mathbf{M}(\xi)|$ будет

$$|\mathbf{M}_s^*(\xi)| = |\mathbf{M}_1(\xi) - \mathbf{M}_2'(\xi)\mathbf{M}_3^{-1}(\xi)\mathbf{M}_2(\xi)|, \tag{8.1}$$

где

$$\mathbf{M}(\xi) = \begin{Vmatrix} \mathbf{M}_1(\xi) & \mathbf{M}_2'(\xi) \\ \mathbf{M}_2(\xi) & \mathbf{M}_3(\xi) \end{Vmatrix} \tag{8.2}$$

($\mathbf{M}_1(\xi)$ — матрица размера $(s+1) \times (s+1)$, штрих означает транспонирование).

Аналогом $d(x, \xi)$ является

$$d_s(x, \xi) = \operatorname{tr} \mathbf{P}(\xi)(\mathbf{f}^{(1)}(x) - \mathbf{D}'(\xi)\mathbf{f}^{(2)}(x))(\mathbf{f}^{(1)}(x) - \mathbf{D}'(\xi)\mathbf{f}^{(2)}(x))' =$$

$$= (\mathbf{f}^{(1)}(x) - \mathbf{D}'(\xi)\mathbf{f}^{(2)}(x), \mathbf{P}(\xi)(\mathbf{f}^{(1)}(x) - \mathbf{D}'(\xi)\mathbf{f}^{(2)}(x))), \tag{8.3}$$

где

$$\mathbf{f}^{(1)}(x) = (f_0(x), \ldots, f_s(x)), \quad \mathbf{f}^{(2)}(x) = (f_{s+1}(x), \ldots, f_n(x)),$$

$$\mathbf{P}(\xi) = [\mathbf{M}_s^*(\xi)]^{-1} \text{ и } \mathbf{D}(\xi) = \mathbf{M}_3^{-1}(\xi)\,\mathbf{M}_2(\xi).$$

Наша цель в этом параграфе — распространить принцип эквивалентности (теорема 7.1) так, чтобы показать, что мера $\xi$ максимизирует $|\mathbf{M}_s^*(\xi)|$ тогда и только тогда, когда она минимизирует $\max\limits_x d_s(x, \xi)$.

Сначала мы хотим распространить определение $\mathbf{M}_s^*(\xi)$ на случай, когда матрица $\mathbf{M}_3(\xi)$ особенная. Для этой цели нам потребуются следующие элементарные факты, которые вытекают из утверждения о том, что $\mathbf{M}(\xi)$ в (8.2) — положительно полуопределенная матрица.

Лемма 8.1. $\operatorname{rang} \mathbf{M}_2(\xi) \subset \operatorname{rang} \mathbf{M}_3(\xi)$.

Доказательство. Утверждение леммы эквивалентно соотношению $[\operatorname{rang} \mathbf{M}_2(\xi)]^{\perp} \supset [\operatorname{rang} \mathbf{M}_3(\xi)]^{\perp}$ ($\perp$ обозначает ортогональное дополнение). Но для произвольной матрицы $\mathbf{C}$ нуль — пространство $\mathcal{N}(\mathbf{C}')$ от $\mathbf{C}'$ совпадает с $[\operatorname{rang} \mathbf{C}]^{\perp}$. Следовательно, нам нужно показать, что $\mathcal{N}(\mathbf{M}_2'(\xi)) \supset \mathcal{N}(\mathbf{M}_3'(\xi))$.

Так как матрица $\mathbf{M}(\xi)$ неотрицательно определена, то если $\mathbf{x} = \{\varepsilon \mathbf{x}_1, \mathbf{x}_2\}$ ($\mathbf{x}_1$ — вектор из $s+1$ координат и $\mathbf{x}_2$ — вектор из $n-s$ координат), то для всех вещественных $\varepsilon$

$$(\mathbf{x}, \mathbf{M}(\xi)\,\mathbf{x}) = \varepsilon^2 (\mathbf{x}_1, \mathbf{M}_1(\xi)\,\mathbf{x}_1) + 2\varepsilon (\mathbf{x}_1, \mathbf{M}_2'(\xi)\,\mathbf{x}_2) + (\mathbf{x}_2, \mathbf{M}_3(\xi)\,\mathbf{x}_2) \geqslant 0.$$

Пусть $x_2$ содержится в $\mathcal{N}(\mathbf{M}_3(\xi)) = \mathcal{N}(\mathbf{M}_3'(\xi))$. Тогда $(\mathbf{x}_2, \mathbf{M}_3(\xi)\,x_2) = 0$, и из предшествующего неравенства следует, что $(x_1, \mathbf{M}_1'(\xi)\,x_2) = 0$ для всех $\mathbf{x}_1$. Следовательно, $\mathbf{x}_2 \in \mathcal{N}(\mathbf{M}_2'(\xi))$, что влечет $\mathcal{N}(\mathbf{M}_2'(\xi)) \supset \mathcal{N}(\mathbf{M}_3'(\xi))$.

Следствие 8.1. *Решение* $\mathbf{X}$ *матричного уравнения* $\mathbf{M}_3(\xi)\,\mathbf{X} = \mathbf{M}_2(\xi)$ *всегда существует.*

Лемма 8.2. *Матрица* $\mathbf{X}'\mathbf{M}_3(\xi)\,\mathbf{X}$ *не зависит от* $\mathbf{X}$, *где* $\mathbf{X}$ — *любое решение матричного уравнения* $\mathbf{M}_3(\xi)\,\mathbf{X} = \mathbf{M}_2(\xi)$.

Доказательство. Если $\mathbf{M}_3(\xi)\,\mathbf{D} = 0$, то, очевидно, $(\mathbf{X}' + \mathbf{D}')\,\mathbf{M}_3(\mathbf{X} + \mathbf{D}) = \mathbf{X}'\mathbf{M}\mathbf{X}$.

Положим

$$\mathbf{M}_s^*(\xi) = \mathbf{M}_1(\xi) - \mathbf{X}(\xi)'\,\mathbf{M}_3(\xi)\,\mathbf{X}(\xi),$$

где $\mathbf{X}(\xi)$ — решение уравнения $\mathbf{M}_3(\xi)\,\mathbf{X}(\xi) = \mathbf{M}_2(\xi)$. По лемме 8.2 $\mathbf{M}_s^*(\xi)$ определяется корректным образом независимо от выбранного частного решения. Мы определим также $d_s(x, \xi)$, как в (8.3), с заменой $\mathbf{D}(\xi)$ на $\mathbf{X}(\xi)$ при условии, что $\mathbf{M}_s^*(\xi)$ — неособенная матрица.

Величина $\mathbf{M}_s^*(\xi)$ обладает следующими свойствами.

Лемма 8.3.

(1) *Матрица* $\mathbf{M}_s^* = \mathbf{M}_s^*(\xi)$ *неотрицательно определена.*

(2) $\mathbf{M}_s^*(\xi) = \mathbf{M}_1(\xi) - \mathbf{X}'(\xi)\mathbf{M}_2(\xi)\mathbf{X}(\xi) \leqslant \mathbf{M}_1(\xi)$, *т. е.* $\mathbf{M}_1(\xi) - \mathbf{M}_s^*(\xi)$ *неотрицательно определена.*

Доказательство.

(1) Если $\mathbf{D}$ — произвольная матрица размера $(n-s)\times(s+1)$ и $\mathbf{I}$ — единичная матрица размера $(s+1)\times(s+1)$, то

$$\|\mathbf{I}, -\mathbf{D}'\| \begin{Vmatrix} \mathbf{M}_1 & \mathbf{M}_2' \\ \mathbf{M}_2 & \mathbf{M}_3 \end{Vmatrix} \begin{Vmatrix} \mathbf{I} \\ \mathbf{D} \end{Vmatrix} = \mathbf{M}_1 - \mathbf{M}_2'\mathbf{D} - \mathbf{D}'\mathbf{M}_2 + \mathbf{D}'\mathbf{M}_3\mathbf{D}$$

есть, очевидно, неотрицательно определенная матрица. В частности, пусть $\mathbf{D}$ удовлетворяет равенству $\mathbf{M}_3\mathbf{D} = \mathbf{M}_2$. Тогда указанная выше матрица принимает вид

$$\mathbf{M}_1 - \mathbf{M}_2'\mathbf{D} - \mathbf{D}'\mathbf{M}_2 + \mathbf{D}'\mathbf{M}_3\mathbf{D} = \mathbf{M}_1 - \mathbf{M}_2'\mathbf{D} = \mathbf{M}_1 - \mathbf{D}'\mathbf{M}_3\mathbf{D} = \mathbf{M}_s^*.$$

Таким образом, матрица $\mathbf{M}_s^* = \mathbf{M}_s^*(\xi)$ всегда неотрицательно определена.

(2) Этот результат получается непосредственно, так как матрица $\mathbf{X}'(\xi)\mathbf{M}_3(\xi)\mathbf{X}(\xi)$, очевидно, неотрицательно определена.

Нам потребуются также две следующие элементарные леммы.

Лемма 8.4. *Если* $\mathbf{X}(\xi)$ — *решение уравнения* $\mathbf{M}_3(\xi)\mathbf{X}(\xi) = \mathbf{M}_2(\xi)$, *то матрица*

$$\mathbf{X}'\mathbf{M}_3\mathbf{X} - \mathbf{X}'\mathbf{M}_3\mathbf{D} - \mathbf{D}'\mathbf{M}_3\mathbf{X} + \mathbf{D}'\mathbf{M}_3\mathbf{D} = \\ = \mathbf{X}'\mathbf{M}_3\mathbf{X} - \mathbf{M}_2'\mathbf{D} - \mathbf{D}'\mathbf{M}_2 + \mathbf{D}'\mathbf{M}_3\mathbf{D}$$

*неотрицательно определена. Кроме того, эта матрица является нулевой матрицей тогда и только тогда, когда* $\mathbf{D}$ *также удовлетворяет уравнению* $\mathbf{M}_3\mathbf{D} = \mathbf{M}_2$.

Доказательство. Выписанная в лемме матрица совпадает с матрицей

$$(\mathbf{X}' - \mathbf{D}')\mathbf{M}_3(\mathbf{X} - \mathbf{D}), \tag{8.4}$$

которая, очевидно, неотрицательно определена и обращается в нулевую матрицу, если $\mathbf{D}$ удовлетворяет равенству $\mathbf{M}_3\mathbf{D} = \mathbf{M}_2$. Более того, если $\mathbf{M}_3(\mathbf{X} - \mathbf{D}) \neq 0$, то $\mathbf{A} = \mathbf{M}_3^{1/2}(\mathbf{X} - \mathbf{D}) \neq 0$, и (8.4) принимает вид $\mathbf{A}\mathbf{A}'$, где $\mathbf{A} \neq 0$. Это выражение становится нулевой матрицей тогда и только тогда, когда $\mathbf{A} = 0$ или, что то же самое, тогда и только тогда, когда $\mathbf{D}$ удовлетворяет равенству $\mathbf{M}_3\mathbf{D} = \mathbf{M}_2$.

Лемма 8.5. *Если* $\mathbf{C}$ — *произвольная неотрицательно определенная матрица размера* $(s+1)\times(s+1)$, *то*

$$\inf_{|\mathbf{P}|=1} \operatorname{tr}\mathbf{PC} = (s+1)\,|\mathbf{C}|^{1/(s+1)}, \tag{8.5}$$

*где нижняя грань берется по множеству положительно определенных матриц* $\mathbf{P}$ *(размера* $(s+1)\times(s+1)$*), определитель которых равен единице. В случае, когда* $|\mathbf{C}| > 0$, *равенство в* (8.5) *достига-*

*ется тогда и только тогда, когда матрица* $\mathbf{P}$ *пропорциональна* $\mathbf{C}^{-1}$.

Это есть иная форма неравенства между средним арифметическим и средним геометрическим (см. стр. 325).

Мы теперь подготовлены к исследованию функций $\mathbf{M}_s^*(\xi)$ и $d_s(x, \xi)$. Так как функции $f_i$, $i = 0, 1, \ldots, n$, непрерывны, то из леммы 8.3 (2) следует, что

$$0 < \sup_{\xi} |\mathbf{M}_1(\xi) - \mathbf{X}'(\xi)\mathbf{M}_3(\xi)\mathbf{X}(\xi)| = a < \infty. \tag{8.6}$$

Пусть $\{\mathbf{M}(\xi_k)\}$ — последовательность информационных матриц, для которых $|\mathbf{M}_1(\xi_k) - \mathbf{X}'(\xi_k)\mathbf{M}_3(\xi_k)\mathbf{X}(\xi_k)| \to a$, и без потери общности мы можем считать последовательность $\xi_k$, слабо сходящейся к некоторому $\xi_0$. Тогда

$$\lim_{k\to\infty} \mathbf{M}(\xi_k) = \mathbf{M}(\xi_0). \tag{8.7}$$

Пока еще не ясно, сходятся ли матрицы $\mathbf{X}(\xi_k)$; однако ниже будет показано, что $a = |\mathbf{M}_s^*(\xi_0)| = |\mathbf{M}_1(\xi_0) - \mathbf{X}_1'(\xi_0)\mathbf{M}_3(\xi_0)\mathbf{X}(\xi_0)|$.

Введем функцию

$$\beta(\mathbf{P}, \mathbf{D}; \xi) = \operatorname{tr}\mathbf{P}(\mathbf{M}_1 - \mathbf{D}'\mathbf{M}_2 - \mathbf{M}_2'\mathbf{D} + \mathbf{D}'\mathbf{M}_3\mathbf{D}) =$$
$$= \int \operatorname{tr}\mathbf{P}(\mathbf{f}^{(1)} - \mathbf{D}'\mathbf{f}^{(2)})(\mathbf{f}^{(1)} - \mathbf{D}'\mathbf{f}^{(2)})'\,d\xi. \tag{8.8}$$

Здесь $\mathbf{D}$ — произвольная матрица размера $r \times (s+1)$, где $r = n - s$, а $\mathbf{P}$ — положительно определенная матрица, определитель которой равен $a^{-1}$. Очевидно,

$$\beta(\mathbf{P}, \mathbf{D}; \xi) = \operatorname{tr}\mathbf{P}(\mathbf{M}_1(\xi) - \mathbf{D}'\mathbf{M}_2(\xi) - \mathbf{M}_2'(\xi)\mathbf{D} + \mathbf{D}'\mathbf{M}_3(\xi)\mathbf{D}) =$$
$$= \operatorname{tr}\mathbf{P}(\mathbf{M}_1(\xi) - \mathbf{X}'(\xi)\mathbf{M}_3(\xi)\mathbf{X}(\xi) + \mathbf{X}'(\xi)\mathbf{M}_3(\xi)\mathbf{X}(\xi) -$$
$$- \mathbf{D}'\mathbf{M}_2(\xi) - \mathbf{M}_2'(\xi)\mathbf{D} + \mathbf{D}'\mathbf{M}_3(\xi)\mathbf{D}),$$

так что по лемме 8.4

$$\beta(\mathbf{P}, \mathbf{D}; \xi) \geqslant \operatorname{tr}\mathbf{P}(\mathbf{M}_1(\xi) - \mathbf{X}'(\xi)\mathbf{M}_3(\xi)\mathbf{X}(\xi)), \tag{8.9}$$

и равенство в (8.9) достигается тогда и только тогда, когда $\mathbf{D}$ удовлетворяет уравнению $\mathbf{M}_3(\xi)\mathbf{D} = \mathbf{M}_2(\xi)$. Поэтому

$$\inf_{\substack{\mathbf{P},\mathbf{D}\\ |\mathbf{P}|=a^{-1}}} \beta(\mathbf{P}, \mathbf{D}; \xi) = \inf_{\substack{\mathbf{P}\\ |\mathbf{P}|=a^{-1}}} \operatorname{tr}\mathbf{P}(\mathbf{M}_1(\xi) - \mathbf{X}'(\xi)\mathbf{M}_3(\xi)\mathbf{X}(\xi)). \tag{8.10}$$

Обращаясь теперь к лемме 8.5, мы выводим, что

$$\inf_{\substack{\mathbf{P}\\ |\mathbf{P}|=a^{-1}}} \operatorname{tr}\mathbf{P}(\mathbf{M}_1(\xi) - \mathbf{X}'(\xi)\mathbf{M}_3(\xi)\mathbf{X}(\xi)) =$$
$$= (s+1)\frac{|\mathbf{M}_1(\xi) - \mathbf{X}'(\xi)\mathbf{M}_3(\xi)\mathbf{X}(\xi)|^{1/(s+1)}}{a^{1/(s+1)}}, \tag{8.11}$$

где $a$ определено в (8.6). Соединяя (8.10) и (8.11), мы получаем, что для всех вероятностных мер $\xi$ выполняется равенство

$$\inf_{\substack{\mathbf{P},\mathbf{D}\\ |\mathbf{P}|=a^{-1}}} \beta(\mathbf{P}, \mathbf{D}; \xi) = (s+1)\frac{|\mathbf{M}_1(\xi) - \mathbf{X}'(\xi)\mathbf{M}_3(\xi)\mathbf{X}(\xi)|^{1/(s+1)}}{a^{1/(s+1)}}, \tag{8.12}$$

$$\sup_{\xi} \inf_{\substack{\mathbf{P},\mathbf{D}\\ |\mathbf{P}|=a^{-1}}} \beta(\mathbf{P}, \mathbf{D}; \xi) = s+1. \tag{8.13}$$

Если план $\xi$ таков, что $|\mathbf{M}_s^*(\xi)| = |\mathbf{M}_1(\xi) - \mathbf{X}'(\xi)\mathbf{M}_3(\xi)\mathbf{X}(\xi)| > 0$, то из леммы 8.5 следует, что равенство в (8.12) достигается тогда и только тогда, когда $\mathbf{D}$ удовлетворяет $\mathbf{M}_3(\xi)\mathbf{D} = \mathbf{M}_2(\xi)$ и матрица $\mathbf{P}$ пропорциональна $(\mathbf{M}_1(\xi) - \mathbf{X}'(\xi)\mathbf{M}_3(\xi)\mathbf{X}(\xi))^{-1}$, причем коэффициент пропорциональности определяется из условия $|\mathbf{P}| = a^{-1}$.

Наконец, так как

$$\beta(\mathbf{P}, \mathbf{D}; \xi_k) \geqslant (s+1)\frac{|\mathbf{M}_1(\xi_k) - \mathbf{X}'(\xi_k)\mathbf{M}_3(\xi_k)\mathbf{X}(\xi_k)|^{1/(s+1)}}{a^{1/(s+1)}},$$

то из сходимости последовательности $\{\xi_k\}$ следует, что

$$\beta(\mathbf{P}, \mathbf{D}; \xi_0) \geqslant s+1. \tag{8.14}$$

Поэтому

$$(s+1)\frac{|\mathbf{M}_1(\xi_0) - \mathbf{X}'(\xi_0)\mathbf{M}_3(\xi_0)\mathbf{X}(\xi_0)|^{1/(s+1)}}{a^{1/(s+1)}} = \inf_{\substack{\mathbf{P},\mathbf{D}\\ |\mathbf{P}|=a^{-1}}} \beta(\mathbf{P}, \mathbf{D}; \xi_0) \geqslant s+1$$

и, следовательно,

$$|\mathbf{M}_1(\xi_0) - \mathbf{X}'(\xi_0)\mathbf{M}_3(\xi_0)\mathbf{X}(\xi_0)| = a. \tag{8.15}$$

Проделанное исследование будет полезно при доказательстве следующей теоремы.

Теорема 8.1. *Пусть $\mathcal{P}$ обозначает множество всех положительно определенных матриц (размера $(s+1)\times(s+1)$), причем матрицы нормированы так, что $|\mathbf{P}| = a^{-1}$, и пусть $\mathcal{D}$ — множество всех вещественных матриц размера $(n-s)\times(s+1)$. Ядро, определенное в* (8.8), *удовлетворяет соотношению*

$$\max_{\xi} \min_{\mathbf{P},\mathbf{D}} \beta(\mathbf{P}, \mathbf{D}; \xi) = s+1 = \min_{\mathbf{P},\mathbf{D}} \max_{\xi} \beta(\mathbf{P}, \mathbf{D}; \xi).$$

*Кроме того, $\xi_0$, определенное в* (8.15), *удовлетворяет неравенству*

$$\beta(\mathbf{P}, \mathbf{D}; \xi_0) \geqslant s+1$$

*для всех $(\mathbf{P}, \mathbf{D}) \in \mathcal{P}\times\mathcal{D}$. Если $(\mathbf{P}_0, \mathbf{D}_0) \in \mathcal{P}\times\mathcal{D}$ удовлетворяет неравенству $\beta(\mathbf{P}_0, \mathbf{D}_0; \xi) \leqslant s+1$ для всех вероятностных мер $\xi$ на $X$, то $\mathbf{D}_0$ удовлетворяет равенству $\mathbf{M}_3(\xi)\mathbf{D}_0 = \mathbf{M}_2(\xi_0)$ и $\mathbf{P}_0$ равно $(\mathbf{M}_1(\xi_0) - \mathbf{D}_0'\mathbf{M}_3(\xi_0)\mathbf{D}_0)^{-1}$, умноженному на постоянный множитель.*

Доказательство. Рассмотрим последовательность игр (см. замечание 2.1) с ядрами

$$\Psi_N(\mathbf{Q}, \mathbf{E}; \xi) =$$
$$= \operatorname{tr}(\mathbf{Q}'\mathbf{M}_1(\xi)\mathbf{Q} + \mathbf{Q}'\mathbf{M}_2'(\xi)\mathbf{E} + \mathbf{E}'\mathbf{M}_2(\xi)\mathbf{Q} + \mathbf{E}'\mathbf{M}_3(\xi)\mathbf{E}) =$$
$$= \operatorname{tr}\left(\| \mathbf{Q}'\mathbf{E}' \| \begin{Vmatrix} \mathbf{M}_1(\xi) & \mathbf{M}_2'(\xi) \\ \mathbf{M}_2(\xi) & \mathbf{M}_3(\xi) \end{Vmatrix} \begin{Vmatrix} \mathbf{Q} \\ \mathbf{E} \end{Vmatrix}\right),$$

где $\mathbf{E}$ обозначает произвольную вещественную матрицу размера $(n-s) \times (s+1)$, элементы которой равномерно ограничены посредством $N$, а $\mathbf{Q}$ — положительно определенная матрица размера $(s+1) \times (s+1)$, нормированная так, что $|\mathbf{Q}| \geqslant a^{-1/2}$, а наименьшее собственное число $\mathbf{Q} \geqslant 1/N$. Переменная $\xi$, как и прежде, пробегает множество $\Xi$ всех вероятностных мер на $X$.

Обозначим пространство стратегий минимизирующего игрока через $\mathcal{L}_N \times \mathcal{E}_N$. Воспользовавшись известным неравенством

$$|\mathbf{Q}_1 + \mathbf{Q}_2|^{1/(s+1)} \geqslant |\mathbf{Q}_1|^{1/(s+1)} + |\mathbf{Q}_2|^{1/(s+1)},$$

где матрицы $\mathbf{Q}_1$ и $\mathbf{Q}_2$ положительно полуопределены, легко установить, что $\mathcal{L}_N$ выпукло и компактно и, очевидно, то же самое имеет место для $\mathcal{E}_N$. Ясно, что $\Xi$ выпукло и компактно, так как $X$ компактно. Кроме того, отметим, что ядро $\Psi_N(\mathbf{Q}, \mathbf{E}; \xi)$ выпукло по отношению к $(\mathbf{Q}, \mathbf{E})$ и линейно по $\xi$. Основная теорема теории игр (см. замечание 2.1) гарантирует существование $(\mathbf{Q}_N, \mathbf{E}_N) \in \mathcal{L}_N \times \mathcal{E}_N$ и $\xi_N \in \Xi$, удовлетворяющих соотношениям

$$\Psi_N(\mathbf{Q}_N, \mathbf{E}_N; \xi) \leqslant U_N \tag{8.16}$$

для всех $\xi \in \Xi$, и

$$\Psi_N(\mathbf{Q}, \mathbf{E}; \xi_N) \geqslant v_N \tag{8.17}$$

для всех $(\mathbf{Q}, \mathbf{E}) \in \mathcal{L}_N \times \mathcal{E}_N$ для некоторой константы $v_N$. Ясно, что $v_N$ убывает при возрастании $N$, так как $\mathcal{L}_N \times \mathcal{E}_N$ увеличивается. Это показывает, в частности, что $v_N$ равномерно ограничено.

Теперь, выбирая $\xi$ так, что матрица $\mathbf{M}(\xi)$ неособенная, мы получим из (8.16), что элементы матриц $\mathbf{Q}_N$ и $\mathbf{E}_N$ равномерно ограничены, как функции от $N$. Взяв пару предельных матриц $\mathbf{Q}_0$ и $\mathbf{E}_0$ для $\{\mathbf{Q}_N\}$ и $\{\mathbf{E}_N\}$ соответственно и слабый предел $\tilde{\xi}$ мер $\{\xi_N\}$ и осуществив очевидный предельный переход в (8.16) и (8.17), приходим к неравенствам

$$\Psi(\mathbf{Q}_0, \mathbf{E}_0; \xi) \leqslant v \tag{8.18}$$

для всех $\xi \in \Xi$, и

$$\Psi(\mathbf{Q}, \mathbf{E}; \tilde{\xi}) \geqslant v \tag{8.19}$$

для всех $(\mathbf{Q}, \mathbf{E}) \in \mathcal{L} \times \mathcal{E}$, где $v = \lim\limits_{N \to \infty} v_N$. Очевидно, $\mathbf{Q}_0$ — положительно определенная матрица размера $(s+1)\times(s+1)$, удовлетворяющая условию $|\mathbf{Q}_0|\geqslant a^{-1/2}$. Так как $|\mathbf{Q}_0|\geqslant a^{-1/2}$, то мы можем ввести матрицу $\mathbf{D}_0=-\mathbf{E}_0\mathbf{Q}_0^{-1}$ и $\mathbf{D}=-\mathbf{E}\mathbf{Q}^{-1}$ для произвольных $(\mathbf{Q}, \mathbf{E})$ в $\mathscr{L}\times\mathscr{E}$. Тогда (8.18) и (8.19) принимают вид

$$\beta(\mathbf{P}_0, \mathbf{D}_0; \xi)\leq v \tag{8.20}$$

для всех $\xi\in\Xi$ и

$$\beta(\mathbf{P}, \mathbf{D}; \tilde{\xi})\geqslant v \tag{8.21}$$

для всех $(\mathbf{P}, \mathbf{D})\in\mathscr{P}\times\mathscr{D}$.

Из определения непосредственно следует, что

$$\beta(\lambda\mathbf{P}, \mathbf{D}; \xi)=\lambda\beta(\mathbf{P}, \mathbf{D}; \xi)$$

для всех $\lambda>0$. Поэтому мы можем предполагать, что $\mathbf{P}_0$ удовлетворяет условию нормировки $|\mathbf{P}_0|=a^{-1}$; в противном случае заменим $\mathbf{P}_0$ на $\lambda\mathbf{P}_0$ при подходящем $\lambda$ $(0<\lambda<1)$, и соотношения (8.20) и (8.21), очевидно, сохранят силу.

Неравенства (8.20) и (8.21) влекут тождество

$$\max_{\xi}\min_{\mathbf{P},\mathbf{D}}\beta(\mathbf{P}, \mathbf{D}; \xi)=v=\min_{\mathbf{P},\mathbf{D}}\max_{\xi}\beta(\mathbf{P}, \mathbf{D}; \xi). \tag{8.22}$$

Сравнение (8.22), (8.13) и (8.14) показывает, что $v=s+1$ и что мы можем заменить $\tilde{\xi}$ в (8.21) на $\xi_0$. Так как $\mathbf{P}_0$, $\mathbf{D}_0$ минимизируют $\beta(\mathbf{P}, \mathbf{D}; \xi_0)$, то характеризация $\mathbf{P}_0$ и $\mathbf{D}_0$, даваемая в настоящей теореме, следует из рассуждений, сопровождающих (8.14). Доказательство теоремы закончено.

Используя эту теорему, мы можем получить следующее обобщение следствия к теореме 2.1 (ср. с работой Кифера [1962]).

Теорема 8.2. *Пусть $f_0, \ldots, f_n$ — линейно независимые непрерывные функции на компактном пространстве $X$. Тогда при $0\leq s\leq n$ существуют вещественные числа $a_{ij}$, $0\leq i\leq s$, $0\leq j\leq n$, с неособенной матрицей $\|a_{ij}\|$ $(0\leq i,\ j\leq s)$ и вероятностная мера $\xi_0$ на $X$ с конечным спектром такие, что*

(a) *функции $g_i=\sum_{j=0}^{n}a_{ij}f_j$, $0\leq i\leq s$, ортонормальны по отношению к $\xi_0$ и ортогональны $f_j$ при $s<j\leq n$,*

(b) $\max\limits_{x}\sum_{i=0}^{s}g_i^2(x)=\int\sum_{i=0}^{s}g_i^2(x)\,\xi(dx)=s+1.$

Доказательство. По лемме 2.1 можно выбрать меру $\xi_0$ с конечным спектром, удовлетворяющую (8.14). Пусть $\mathbf{D}_0$ и $\mathbf{P}_0$ такие, как в теореме 8.1.

Определим

$$(g_0, \ldots, g_s)=g=\mathbf{P}_0^{1/2}\{\mathbf{f}^{(1)}-\mathbf{D}_0'\mathbf{f}^{(2)}\}.$$

Так как $\mathbf{P}_0[(\mathbf{M}_1(\xi_0)-\mathbf{D}_0'\mathbf{M}_2(\xi_0)-\mathbf{M}_2'(\xi_0)\mathbf{D}_0+\mathbf{D}_0'\mathbf{M}_3(\xi_0)\mathbf{D}_0)]=\mathbf{I}$

($\mathbf{I}$ — единичная матрица), мы заключаем, что функции $g_0, \ldots, g_s$ ортонормальны по отношению к $\xi_0$.

Кроме того, так как

$$\beta(\mathbf{P}_0, \mathbf{D}_0 + \varepsilon\mathbf{D}; \xi_0) \geqslant \beta(\mathbf{P}_0, \mathbf{D}_0; \xi_0)$$

для всех $\mathbf{D} \in \mathcal{D}$ и произвольного вещественного $\varepsilon$, то отсюда следует, что

$$-2\varepsilon \int \operatorname{tr} g(x) f^{(2)}(x)' \mathbf{D} \mathbf{P}_0^{1/2} d\xi_0 + \varepsilon^2 \operatorname{tr} \mathbf{P}_0 \mathbf{D}' \mathbf{M}_3(\xi_0) \mathbf{D} \geqslant 0.$$

Однако это неравенство выполняется, только если коэффициент при $\varepsilon$ равен нулю. Теперь, выбирая $\mathbf{D}$ таким, что $\mathbf{D}\mathbf{P}_0^{1/2}$ — матрица размера $(n-s) \times (s+1)$ с нулями всюду, кроме $j$-го столбца и $i$-й строки, получим

$$\int g_i f_{s+j} d\xi_0 = 0, \qquad 0 \leq i \leq s, \quad 0 < j \leq n-s.$$

Этим завершается доказательство пункта (a). Чтобы доказать (b), заметим, что

$$s+1 = \beta(\mathbf{P}_0, \mathbf{D}_0; \xi_0) = \sup_{\xi} \beta(\mathbf{P}_0, \mathbf{D}_0; \xi) =$$

$$= \sup_{\xi} \int \operatorname{tr} g(x) g'(x) d\xi(x) = \sup_{x} \sum_{i=0}^{s} g_i^2(x),$$

откуда $\max\limits_{x} \sum\limits_{i=0}^{s} g_i^2(x) = s+1$ и $\int \sum\limits_{i=0}^{s} g_i^2(x) d\xi_0 = s+1$. Доказательство теоремы 8.2 закончено.

Получим теперь аналог теоремы 7.1.

Для любой меры такой, что $|\mathbf{M}_s^*(\xi)| > 0$, определим (см. (8.3))

$$d_s(x, \xi) = \operatorname{tr} \mathbf{P}(\xi)(f^{(1)}(x) - \mathbf{D}'(\xi) f^{(2)}(x))(f^{(1)}(x) - \mathbf{D}'(\xi) f^{(2)}(x))', \quad (8.23)$$

где

$$\mathbf{P}(\xi) = [\mathbf{M}_s^*(\xi)]^{-1} = [\mathbf{M}_1(\xi) - \mathbf{X}'(\xi)\mathbf{M}_3(\xi)\mathbf{X}(\xi)]^{-1},$$
$$\mathbf{M}_3(\xi)\mathbf{X}(\xi) = \mathbf{M}_2(\xi).$$

В случае, когда матрица $\mathbf{M}(\xi)$ неособенная, можно проверить (используя очевидные соотношения $\mathbf{M}_i$ и $\mathbf{M}^{(i)}$), что

$$\left( \mathbf{M}(\xi)^{-1} = \left\| \begin{matrix} \mathbf{M}^{(1)}(\xi) & \mathbf{M}^{(2)}(\xi)' \\ \mathbf{M}^{(2)}(\xi) & \mathbf{M}^{(3)}(\xi) \end{matrix} \right\| \right),$$

и $d_s(x, \xi)$ сводится к выражению

$$d_s(x, \xi) = (\mathbf{f}(x), \mathbf{M}^{-1}(\xi)\mathbf{f}(x)) - (\mathbf{f}^{(2)}(x), \mathbf{M}_3^{-1}(\xi)\mathbf{f}^{(2)}(x)).$$

Условия теоремы 7.1 в данном случае будут иметь вид:

$$
\begin{aligned}
&(1)\ \xi^* \text{ максимизирует } |\mathbf{M}_s^*(\xi)| = |\mathbf{M}_1(\xi) - \mathbf{X}'(\xi)\,\mathbf{M}_3(\xi)\,\mathbf{X}(\xi)|,\\
&(2)\ \xi^* \text{ минимизирует } \max_x d_s(x, \xi), \qquad\qquad (8.24)\\
&(3)\ \max_x d_s(x, \xi^*) = s + 1.
\end{aligned}
$$

Аналогом теоремы 7.1 является следующая теорема.

Теорема 8.3. *Множество мер* $\xi^*$, *удовлетворяющих* (1), (2) *или* (3) *из* (8.24) *совпадают.*

Доказательство. Если $\xi^*$ удовлетворяет (2), то

$$\max_x d_s(x, \xi^*) = \min_\xi \max_x d_s(x, \xi) = \min_\xi \max_\eta \int d_s(x, \xi)\, d\eta(x) =$$

$$= \min_\xi \max_\eta \beta(\mathbf{P}(\xi),\ \mathbf{D}(\xi);\ \eta) \leq \max_\eta \beta(\mathbf{P}_0,\ \mathbf{D}_0;\ \eta) \leq s + 1.$$

Но так как $\beta(\mathbf{P}(\xi),\ \mathbf{D}(\xi);\ \xi) = s + 1$, то отсюда следует, что $\max_x d_s(x, \xi^*) \geqslant s + 1$, и мы можем заключить, что $\xi^*$ удовлетворяет (3). Условие (3), очевидно, влечет (2), так как всегда $\max_x d_s(x, \xi) \geqslant s + 1$.

Если $\max_x d(x, \xi^*) = s + 1$, то, принимая во внимание (8.12), получим, что

$$s + 1 = \max_\xi \beta(\mathbf{P}(\xi^*), \mathbf{D}(\xi^*);\ \xi) =$$

$$= \max_\xi \lambda^{-1} \beta(\lambda\mathbf{P}(\xi^*),\ \mathbf{D}(\xi^*);\ \xi) \geqslant \frac{(s+1)}{\lambda},$$

где $\lambda$ таково, что $|\mathbf{P}(\xi^*)\,\lambda| = a^{-1}$ $(a = \sup_\xi |\mathbf{M}_s^*(\xi)|)$. Поэтому $a\mathbf{P}(\xi^*)| \leq 1$ или $a \leq |\mathbf{M}_s^*(\xi^*)|$ и, следовательно, условие (3) влечет условие (1).

Из уравнения (8.11) получаем, что всякий план $\xi^*$, максимизирующий $|\mathbf{M}_s^*(\xi)|$, удовлетворяет неравенству $\beta(\mathbf{P}, \mathbf{D};\ \xi^*) \geqslant s + 1$. Однако в этом случае оптимальная пара $\mathbf{P}_0$ и $\mathbf{D}_0$, на которой достигается минимум, характеризуется свойствами: $\mathbf{P}_0 = \mathbf{P}(\xi^*)$ и $\mathbf{M}_3(\xi^*)\,\mathbf{D}_0 = \mathbf{M}_2(\xi^*)$, так что $\beta(\mathbf{P}(\xi^*), \mathbf{D}(\xi^*);\ \xi) \leq s + 1$ для всех $\xi$ и, следовательно, $\max_x d_s(x, \xi^*) = s + 1$. Доказательство закончено.

Замечание 8.1. Выражение $\beta(\mathbf{P}, \mathbf{D};\ \xi)$ может быть записано в виде

$$\beta(\mathbf{P}, \mathbf{D};\ \xi) = \operatorname{tr} \overline{\mathbf{P}}\mathbf{M}(\xi)\,\overline{\mathbf{P}}',$$

где $\overline{\mathbf{P}}$ есть матрица размера $(s + 1) \times (n + 1)$, определенная выражением $\overline{\mathbf{P}} = \|\mathbf{Q},\ -\mathbf{Q}\mathbf{D}'\|$, где $\mathbf{Q}$ — положительный квадратный корень из $\mathbf{P}$. Если матрица $\mathbf{M}(\xi)$ — неособенная, то мы можем выразить $\beta(\mathbf{P}, \mathbf{D}; \xi)$ в виде

$$\beta(\mathbf{P}, \mathbf{D}; \xi) = \sup_{\mathbf{E}\neq 0} \frac{\operatorname{tr}^2 \overline{\mathbf{P}}\mathbf{E}}{\operatorname{tr}\mathbf{E}'\mathbf{M}(\xi)^{-1}\mathbf{E}}, \tag{8.25}$$

где $\mathbf{E}$ — произвольная матрица размера $(n+1)\times(s+1)$. Равенство достигается при $\mathbf{E} = \mathbf{M}(\xi)\overline{\mathbf{P}}'$. Для доказательства (8.25) можно воспользоваться неравенством $\operatorname{tr}^2\overline{\mathbf{P}}\mathbf{E} \leq (\operatorname{tr}\mathbf{E}'\mathbf{M}(\xi)\mathbf{E})(\operatorname{tr}\overline{\mathbf{P}}\mathbf{M}(\xi)\overline{\mathbf{P}}')$, которое легко выводится с помощью диагонализации $\mathbf{M}(\xi)$ и использования подходящей формы неравенства Шварца. В данном случае имеем неравенство

$$\inf_{\mathbf{P},\mathbf{D}} \beta(\mathbf{P}, \mathbf{D}; \xi) = \inf_{\mathbf{P},\mathbf{D}} \sup_{\mathbf{E}\neq 0} \frac{\operatorname{tr}^2(\overline{\mathbf{P}}\mathbf{E})}{\operatorname{tr}\mathbf{E}'\mathbf{M}(\xi)^{-1}\mathbf{E}}.$$

Замечание 8.2. При $|\mathbf{M}(\xi)| > 0$ легко показать, что

$$|\mathbf{M}_1(\xi) - \mathbf{X}'(\xi)\mathbf{M}_3(\xi)\mathbf{X}(\xi)| = |\mathbf{M}_1(\xi) - \mathbf{M}_2'(\xi)\mathbf{M}_3^{-1}(\xi)\mathbf{M}_2(\xi)|$$

сводится к $|\mathbf{M}(\xi)||\mathbf{M}_3(\xi)|^{-1}$, так что (ср. (8.11))

$$\inf_{\mathbf{P},\mathbf{D}} \beta(\mathbf{P}, \mathbf{D}; \xi) = \frac{s+1}{a^{1/(s+1)}} \left(\frac{|\mathbf{M}(\xi)|}{|\mathbf{M}_3(\xi)|}\right)^{1/(s+1)}.$$

Замечание 8.3. В случае $s = 0$ $\beta(\mathbf{P}, \mathbf{D}; \xi)$ принимает (с точностью до постоянного множителя) вид

$$\int_X \left(f_0(x) - \sum_{j=1}^n d_j f_j(x)\right)^2 \xi(dx).$$

Поэтому задача оценивания $\inf\limits_{\mathbf{P},\mathbf{D}} \beta(\mathbf{P}, \mathbf{D}; \xi)$ и определения $\mathbf{D}$, доставляющего минимум, эквивалентна нахождению линейной комбинации $f^*$ от функций $f_1, \ldots, f_n$, для которой $f_0 - f^*$ имеет минимальную норму в метрик $L_2(\xi)$.

Если мы рассматриваем функции $f_i$ как случайные величины, то задача состоит в определении $f^*$, которая минимизирует в существенном дисперсию $f_0 - f^*$. Эта линейная комбинация совпадает с линейной комбинацией, дающей максимальную корреляцию между $f_0$ и $f^*$; квадрат корреляции равен

$$\frac{\left(\int f_0 f^* d\xi\right)^2}{\int f_0^2 d\xi \int (f^*)^2 d\xi}.$$

В случае $(f_0(x), \ldots, f_n(x)) = (x^n, \ldots, x, 1)$ при $x \in [0, 1]$ результат теоремы 8.1 уточняется и имеет вид

$$\sup_{\xi} \inf_{d} \int_0^1 \left(x^n - \sum_{i=0}^{n-1} d_i x^i\right)^2 \xi(dx) = \inf_{d} \sup_{x} \left|x^n - \sum_{i=0}^{n-1} d_i x_i\right|^2,$$

и нижняя и верхняя грани в правой части достигаются, когда многочлен равен константе, умноженной на $n$-й многочлен Чебышева первого рода (ср. § 1 гл. IX).

Дальнейшее развитие темы этого параграфа и работ Элвинга [1952] и [1959] можно найти у Карлина и Стаддена [1966].

## § 9. Открытые вопросы

В этом параграфе мы сформулируем ряд родственных задач, которые остаются нерешенными.

На протяжении данной главы максимизация определителя $|\mathbf{M}(\xi)|$ играла важную роль в решении задачи Фейера и в нахождении устойчивых наиболее экономичных интерполяционных систем для различных весовых функций $w(x)$. В каждом случае, когда точное решение было получено, мера, максимизирующая $|\mathbf{M}(\xi)|$, приписывала массы $n+1$ точкам. Поэтому мы выдвигаем следующие задачи.

Задача 9.1. При каких условиях на функции $f_0, f_1, \ldots, f_n$ верхняя грань $|\mathbf{M}(\xi)|$ достигается для меры с конечным спектром, состоящим из $n+1$ точек?

Задача 9.2. В случае $f_i(x) = \sqrt{w(x)}\, x^i \quad i = 0, 1, \ldots, n$, как можно охарактеризовать весовые функции $w(x)$, для которых максимизация $|\mathbf{M}(\xi)|$ приводит к мере, сосредоточенной в $n+1$ точках?

Вместо максимизации $|\mathbf{M}(\xi)|$ по всем вероятностным мерам, определенным на $X$, можно рассматривать более узкую задачу.

Задача 9.3. Опр делить $\max|\mathbf{M}(\xi)|$, где максимум берется по всем мерам, сосредоточенным в $n+1$ точках.

В случае, когда $\xi$ приписывает массы $\lambda_0, \lambda_1, \ldots, \lambda_n$ точкам $x_0, x_1, \ldots, x_n$, имеем

$$|\mathbf{M}(\xi)| = \left\{\prod_{l=0}^{n} \lambda_l\right\} \{\det \| f_i(x_j) \|_{i,j=0}^{n}\}^2. \tag{9.1}$$

При максимизации (9.1) $\lambda_i$ необходимо равны друг другу, т. е.

$$\lambda_0 = \lambda_1 = \ldots = \lambda_n = \frac{1}{n+1}.$$

Задача 9.4. При каких условиях на $T$-систему $\{f_0, f_1, \ldots \ldots, f_n\}$ максимум (9.1) достигается на единственном множестве точек $x_0, x_1, \ldots, x_n$?

В случае, когда максимум (9.1) не единствен, максимум $|\mathbf{M}(\xi)|$ по всем $\xi$ не достигается обычно для меры, сосредоточенной в $n+1$ точках. Поэтому решение задачи 9.4 является шагом к решению задачи 9.1.

В § 8 мы исследовали (см. 8.8) величину

$$\beta(\mathbf{P}, \mathbf{D}; \xi) = \int_X \operatorname{tr} \mathbf{P}(\mathbf{f}^{(1)} - \mathbf{D}'\mathbf{f}^{(2)})(\mathbf{f}^{(1)} - \mathbf{D}'\mathbf{f}^{(2)})'\xi(dx),$$

где $\mathbf{f}^{(1)} = (f_0, \ldots, f_s)$, $\mathbf{f}^{(2)} = (f_{s+1}, \ldots, f_n)$, $\mathbf{D}$ — матрица размера $(n-s) \times (s+1)$, а $\mathbf{P}$ — положительно определенная матрица размера $(s+1) \times (s+1)$, определитель которой имеет заданное значение. Теорема 8.1 утверждает, что

$$\sup_{\xi} \inf_{\mathbf{P},\mathbf{D}} \beta(\mathbf{P}, \mathbf{D}; \xi) = s + 1 = \inf_{\mathbf{P},\mathbf{D}} \sup_{\xi} \Phi(\mathbf{P}, \mathbf{D}; \xi). \tag{9.2}$$

Далее, в случае $s = 0$ и $f_i(x) = x^i$, $i = 0, 1, \ldots, n$; $X = [-1, 1]$,

$$\sup_{\xi} \inf_{d_0,\ldots,d_{n-1}} \int \left( x^n - \sum_{i=0}^{n-1} d_i x^i \right)^2 d\xi = \inf_{d_0,\ldots,d_{n-1}} \sup_{x} \left| x^n - \sum_{i=0}^{n-1} d_i x^i \right|^2. \tag{9.3}$$

Нижняя грань в правой части достигается, когда $(1/2^{n-1}) \left( x^n - \sum_{i=0}^{n} d_i x^i \right)$ — $n$-й многочлен Чебышева первого рода (ср. § 1 гл. IX).

При $s = 1$ величина, соответствующая $\left( x^n - \sum_{i=0}^{n-1} d_i x^i \right)^2$, заменяется на

$$a(x^n - Q(x))^2 + 2b(x^n - Q(x))(x^{n-1} - R(x)) + c(x^{n-1} - R(x))^2, \tag{9.4}$$

где $Q(x)$ и $R(x)$ — многочлены степени $n-2$ и матрица $\mathbf{P} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ положительно определена и нормирована так, что определитель $ac - b^2$ равен фиксированному положительному числу.

Задача 9.5. Определить $\mathbf{P}$, $Q$ и $R$, подчиненные описанным выше ограничениям, для которых верхняя грань (9.4) при $x \in [-1, 1]$ минимальна, и охарактеризовать минимизирующие многочлены $Q_0(x)$ и $R_0(x)$. Более общим образом, определить условия на $f_0, f_1, \ldots, f_n$ так, чтобы оптимальный выбор $\mathbf{P}_0$, $\mathbf{D}_0$ и $\xi_0$ мог быть охарактеризован если не точно, то хотя бы качественно.

Глава XI

# ВЫПУКЛЫЕ В ОБОБЩЕННОМ СМЫСЛЕ ФУНКЦИИ, ИНДУЦИРОВАННЫЕ $ET$-СИСТЕМАМИ

## § 1. Предварительные замечания

Пусть $u_0, u_1, \ldots, u_n$ — $n+1$ функций таких, что $\{u_i\}_0^k$ — $ET$-система на замкнутом интервале $[a, b]$ для любого $k = 0, 1, \ldots$ $\ldots, n$. Система, обладающая такими свойствами, называется *обобщенной полной системой Чебышева* или *$ECT$-системой* (ср. определения 1.1 и 2.4 гл. I).

Определение 1.1. Говорят, что функция $\varphi$, определенная на открытом интервале $(a, b)$, является *выпуклой по отношению к системе функций* $\{u_i\}_0^n$, если

$$U\begin{pmatrix} u_0, u_1, \ldots, u_n, \varphi \\ t_0, t_1, \ldots, t_n, t_{n+1} \end{pmatrix} = \begin{vmatrix} u_0(t_0) & u_0(t_1) & \ldots & u_0(t_{n+1}) \\ u_1(t_0) & u_1(t_1) & \ldots & u_1(t_{n+1}) \\ \cdot & \cdot & \cdot & \cdot \\ u_n(t_0) & u_n(t_1) & \ldots & u_n(t_{n+1}) \\ \varphi(t_0) & \varphi(t_1) & \ldots & \varphi(t_{n+1}) \end{vmatrix} \geqslant 0 \qquad (1.1)$$

для всех наборов $\{t_i\}_0^{n+1}$, удовлетворяющих условию $a < t_0 <$ $< t_1 < \ldots < t_{n+1} < b$. Символически запишем это так:

$$\varphi \in C(u_0, u_1, \ldots, u_n). \qquad (1.2)$$

Из определения 1.1 видно, что условия (1.1) или (1.2) выполняются тогда и только тогда, когда $\{u_0, u_1, \ldots, u_n, \varphi\}$ — $WT$-система на открытом интервале $(a, b)$. Ясно, что класс функций, удовлетворяющий условию (1.1), является выпуклым конусом, который замкнут в топологии поточечной сходимости. Подчеркнем, что не делается каких-либо предположений относительно функции $\varphi$ в граничных точках $a$ и $b$, т. е. требования (1.1) относятся только к $t_i$, упорядоченным на открытом интервале $(a, b)$.

Замечание 1.1. Конус $C(u_0, u_1, \ldots, u_n)$ может быть определен условиями (1.1) без каких-либо дополнительных требований к системе функций $\{u_i\}_0^n$. Существенно, однако, установить в качестве минимального требования, чтобы функции $u_0, u_1, \ldots, u_n$ были линейно независимы на $(a, b)$, так как в противном случае конус включает все функции на $(a, b)$. Конус $C(u_0, u_1, \ldots, u_n)$ обладает исключительно богатой структурой, когда функции $u_0, u_1, \ldots, u_n$ образуют $T$-систему или $ECT$-систему. Примеры, приведенные ниже, свидетельствуют о широкой сфере применения, заключенной в вышеуказанной формулировке.

Пример 1.1. Если $u_0(t) \equiv 1$, то $C(u_0)$ — множество всех возрастающих функций на $(a, b)$.

Пример 1.2. Пусть $u_0(t) \equiv 1$, $u_1(t) = ct + d$, $d$ вещественно и $c > 0$. Тогда $C(u_0, u_1)$ — множество выпуклых функций на $(a, b)$ (вогнутых, если $c < 0$).

Пример 1.3. Пусть $u_0(t) \equiv 1$, и предположим, что функция $u_1(t) = \psi(t)$ такова, что $\psi'(t) > 0$ для всех $t \in (a, b)$. Тогда $C(1, \psi(t))$ включает множество всех функций, которые выпуклы по отношению к $\psi(t)$ на $(a, b)$.

Пример 1.4. Пусть $u_i(t) = t^i$, $i = 0, 1, \ldots, n$. Тогда $C(1, t, t^2, \ldots, t^n)$ состоит из всех пределов функций в классе $C^n$, для которых $\varphi^{(n+1)}(t) \geqslant 0$ на $(a, b)$.

Понятие выпуклости в обобщенном смысле, сформулированное здесь, по-видимому, было впервые введено Хопфом [1926] и получило дальнейшее развитие у Поповичиу [1936], [1961] и [1962]. Поповичиу главным образом имел дело со случаем $T$-системы $\{1, t, t^2, \ldots, t^n\}$. Другие вопросы обобщенной теории выпуклости были исследованы Беккенбахом [1937], Бонсаллом [1950] и другими авторами. Работа Беккенбаха и Беллмана [1961] содержит обзор исследований по обобщенной теории выпуклости, который дополняет замечательную монографию Поповичиу [1945], посвященную этому вопросу. Наше первоначальное исследование в этой области содержится в работе Карлина и Новикова [1963]. Основной целью этой статьи было описание дуального конуса и конуса выпуклых в обобщенном смысле функций.

Для того чтобы описать структуру конуса $C(u_0, u_1, \ldots, u_n)$, мы посвятим оставшуюся часть этого параграфа доказательству некоторых свойств $ECT$-систем. Первым основным результатом в этом направлении является следующая теорема.

Теорема 1.1. *Пусть* $u_0, u_1, \ldots, u_n$ *принадлежат классу* $C^n[a, b]$. *Тогда* $\{u_i\}_0^n$ — *$ECT$-система на* $[a, b]$ *в том и только в том случае, когда для* $k = 0, 1, \ldots, n$ $W(u_0, \ldots, u_k) > 0$ *на* $[a, b]$, *где* $W(u_0, \ldots, u_k)$ *обозначает определитель Вронского функций* $u_0, u_1, \ldots, u_k$, *т. е.*

$$W(u_0, u_1, \ldots, u_k)(t) = \begin{vmatrix} u_0(t) & u_0'(t) & \ldots & u_0^{(k)}(t) \\ u_1(t) & u_1'(t) & \ldots & u_1^{(k)}(t) \\ \cdot & \cdot & \cdot & \cdot \\ u_k(t) & u_k'(t) & \ldots & u_k^{(k)}(t) \end{vmatrix}.$$

Доказательство. Необходимость является непосредственным следствием определения $ET$-системы.

Доказательство достаточности разобьем на ряд отдельных предложений.

a) Вспомним определение $U^*\begin{pmatrix}0, 1, \ldots, n\\ t_0, t_1, \ldots, t_n\end{pmatrix}$ (см. § 2 гл. I). Если $t_0, t_1, \ldots, t_n$ есть последовательность вида

$$\underbrace{\xi_0, \ldots, \xi_0}_{r_0}, \quad \underbrace{\xi_1, \ldots, \xi_1}_{r_1}, \ldots, \underbrace{\xi_k, \ldots, \xi_k}_{r_k},$$

где $a \leq \xi_0 < \xi_1 < \ldots < \xi_k \leq b$, то столбцы $r_j$ этого определителя, включающие $\xi_j$, совпадают с первыми столбцами определителя $W(u_0, u_1, \ldots, u_n)$, вычисленными при $\xi_j$.

b) Очень часто будет удобно использовать *редуцированную систему функций* $\{v_i\}_0^{n-1}$, которая определяется как

$$v_i(t) = \frac{d}{dt}\left(\frac{u_{i+1}(t)}{u_0(t)}\right), \qquad i = 0, \ldots, n-1. \tag{1.3}$$

Важное тождество

$$W(u_0, u_1, \ldots, u_n) = u_0^{n+1} W(v_0, v_1, \ldots, v_{n-1}), \tag{1.4}$$

которое понижает порядок определителя Вронского на единицу, легко устанавливается с помощью правила Лейбница при дифференцировании в форме

$$\left(\frac{u(t)}{u_0(t)}\right)^{(m)} = \sum_{p=0}^{m}\binom{m}{p} u^{(m-p)}(t)\left(\frac{1}{u_0(t)}\right)^{(p)}.$$

Действительно, разделим каждый столбец определителя $W(u_0, u_1, \ldots, u_n)(t)$ на $u_0(t)$ и затем добавим к каждому столбцу соответствующую линейную комбинацию всех предшествующих ему столбцов, действуя справа налево, с тем, чтобы получить

$$W(u_0, u_1, \ldots, u_n) = u_0^{n+1} W\left(1, \frac{u_1}{u_0}, \ldots, \frac{u_n}{u_0}\right).$$

Тождество (1.4) теперь следует из разложения определителя, стоящего в правой части полученного равенства, по элементам первой строки.

c) В случае, когда $a \leq t_0 < t_1 < \ldots < t_n \leq b$, положительность определителя $U\begin{pmatrix}0, 1, \ldots, n\\ t_0, t_1, \ldots, t_n\end{pmatrix}$ уже доказана в лемме 5.3 гл. VIII.

d) Общий случай, когда $t_0 \leq t_1 \leq \ldots \leq t_n$ образуют последовательность

$$\underbrace{\xi_0, \ldots, \xi_0}_{r_0}, \quad \underbrace{\xi_1, \ldots, \xi_1}_{r_1}, \ldots, \underbrace{\xi_k, \ldots, \xi_k}_{r_k} \qquad (r_0 + \ldots + r_k = n+1),$$

$\xi_0 < \xi_1 < \ldots < \xi_k$ доказывается индукцией по $n$ с помощью редуцированной системы (1.3). Для каждого $j$ выносим за знак определителя $u_0(\xi_j)>0$ из столбцов $r_j$, содержащих $\xi_j$, и затем вычитаем из каждого столбца линейную комбинацию предшествующих ему столбцов, чтобы получить из этих $r_j$ столбцов определителя $U^*\begin{pmatrix}0, & 1, & \ldots, & n\\ t_0, & t_1, & \ldots, & t_n\end{pmatrix}$ первые $r_j$ столбцов определителя $W(1, u_1/u_0, \ldots, u_n/u_0)$, вычисленных при $\xi_j$. Затем вычитаем из каждого столбца, начинающегося с единицы, предшествующий ему столбец такого же вида, действуя справа налево, и применяем теорему о среднем. Элементы столбца с номером $r_0+\ldots+r_j+1$ с точностью до положительного множителя $\xi_{j+1}-\xi_j$ соответственно равны при некотором $\eta_j$

$$(0, v_0(\eta_j), \ldots, v_{n-1}(\eta_j))', \qquad \xi_j<\eta_j<\xi_{j+1}, \qquad j=0, 1, \ldots, k-1.$$

Разложение по элементам первой строки, которая теперь состоит из нулевых компонент, за исключением первой, дает

$$\operatorname{sign} U^*\begin{pmatrix}0, & 1, & \ldots, & n\\ t_0, & t_1, & \ldots, & t_n\end{pmatrix}=\operatorname{sign} V^*\begin{pmatrix}0, & 1, & \ldots, & n-1\\ \hat{t}_0, & \hat{t}_1, & \ldots, & \hat{t}_{n-1}\end{pmatrix},$$

где $\hat{t}_0, \hat{t}_1, \ldots, \hat{t}_{n-1}$ есть последовательность вида

$$\underbrace{\xi_0, \ldots, \xi_0}_{r_0-1}, \eta_0, \underbrace{\xi_1, \ldots, \xi_1}_{r_1-1}, \eta_1, \ldots, \underbrace{\xi_k, \ldots, \xi_k}_{r_k-1}.$$

Так как функции $v_0, v_1, \ldots, v_{n-1}$ удовлетворяют условию $W(v_0, \ldots, v_k)>0$, $k=0, 1, \ldots, n-1$ (см. (1.4)), то теорема сводится к случаю $n-1$. Поскольку случай $n=0$ тривиален, то доказательство можно считать законченным.

Исходя из $n+1$ функций $w_0, w_1, \ldots, w_n$, которые строго положительны на $[a, b]$ и таковы, что $w_k$ принадлежит классу непрерывных функций $C^{n-k}[a, b]$, образуем следующую систему функций:

$$\begin{aligned}
&u_0(t)=w_0(t),\\
&u_1(t)=w_0(t)\int_a^t w_1(\xi_1)\,d\xi_1,\\
&u_2(t)=w_0(t)\int_a^t w_1(\xi_1)\int_a^{\xi_1} w_2(\xi_2)\,d\xi_2 d\xi_1,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&u_n(t)=w_0(t)\int_a^t w_1(\xi_1)\int_a^{\xi_1} w_2(\xi_2)\ldots\int_a^{\xi_{n-1}} w_n(\xi_n)\,d\xi_n\ldots d\xi_1.
\end{aligned}\tag{1.5}$$

Ниже доказывается (теореме 1.2), что функции $u_0, u_1, \ldots, u_n$, определенные в (1.5), образуют $ECT$-систему на $[a, b]$, которая подчиняется граничным условиям

$$u_k^{(p)}(a)=0, \qquad p=0, 1, \ldots, k-1; \qquad k=1, 2, \ldots, n. \tag{1.6}$$

Следующая теорема раскрывает фундаментальный характер представления (1.5) в случае $ECT$-систем.

Теорема 1.2. *Пусть функции* $u_0, u_1, \ldots, u_n$ *из класса* $C^n[a, b]$ *удовлетворяют начальным условиям*

$$u_k^{(p)}(a) = 0, \qquad p = 0, 1, \ldots, k-1; \qquad k = 1, 2, \ldots, n. \quad (1.7)$$

*Следующие три условия эквивалентны:*

1) *система* $\{u_i\}_0^n$ *представима в виде* (1.5),

2) $\{u_i\}_0^n$ *является ECT-системой на* $[a, b]$,

3) $W(u_0, \ldots u_k) > 0$ *на* $[a, b]$ *для* $k = 0, 1, \ldots, n$.

Замечание 1.2. Если функции $u_0, u_1, \ldots, u_n$ не удовлетворяют условиям (1.7) в граничной точке $t = a$ и $\{u_i\}_0^n$ образуют $ECT$-систему, то можно вычесть из каждой функции $u_k$ соответствующую линейную комбинацию предшествующих функций, чтобы получить систему $\tilde{u}_k = u_k - \sum_{i=0}^{k-1} d_{ki}u_i$, $k = 1, 2, \ldots, n$, которая удовлетворяет (1.7). Очевидно, $W(u_0, u_1, \ldots, u_k) = W(u_0, \tilde{u}_1, \ldots, \tilde{u}_k)$.

Доказательство. Эквивалентность условий 2) и 3) является содержанием теоремы 1.1. Для того чтобы доказать, что из 3) следует 1), воспользуемся индукцией. Если $n = 0$, теорема тривиальна. Предположим, что она доказана для систем из $n - 1$ или меньшего числа функций. Пусть $\{v_i\}_0^{n-1}$ обозначает редуцированную систему (1.3), т. е.

$$v_i(t) = \frac{d}{dt}\left(\frac{u_{i+1}(t)}{u_0(t)}\right), \qquad i = 0, 1, \ldots, n-1. \quad (1.8)$$

Так как $u_0^{k+1}W(v_0, \ldots, v_{k-1}) = W(u_0, \ldots, u_k)$ (см. (1.4)), то $W(v_0, \ldots$ $\ldots, v_{k-1}) > 0$ на $[a, b]$ для $k = 1, 2, \ldots, n$, следовательно, $\{v_i\}_0^n$ удовлетворяет 3). Кроме того, из соотношений (1.7) вытекает, что

$$v_k^{(p)}(a) = 0, \qquad p = 0, 1, \ldots, k-1; \qquad k = 1, \ldots, n-1. \quad (1.9)$$

Следовательно, по предположению индукции существуют функции $w_1, w_2, \ldots, w_n$, которые положительны и таковы, что $w_k$ принадлежит классу $C^{n-k}[a, b]$ и

$$\begin{aligned}
&v_0(t) = w_1(t),\\
&v_1(t) = w_1(t)\int_a^t w_2(\xi_2)\,d\xi_2,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \qquad (1.10)\\
&v_{n-1}(t) = w_1(t)\int_a^t w_2(\xi_2)\int_a^{\xi_2} w_3(\xi_3)\ldots\int_a^{\xi_{n-1}} w_n(\xi_n)\,d\xi_n\ldots d\xi_2.
\end{aligned}$$

Полагая $w_0(t) = u_0(t)$, находим, что

$$u_k(t) = w_0(t)\int_a^t v_{k-1}(\xi)\,d\xi + c_k w_0(t), \qquad k = 1, \ldots, n. \quad (1.11)$$

Но (1.7) означает, что $c_k = 0$. Следовательно, (1.11) сводится к требуемому представлению (1.5).

Обратное утверждение доказывается еще проще. Предположим, что $u_0, u_1, \ldots, u_n$ допускают представление вида (1.5). Утверждается, что

$$W(u_0, u_1, \ldots, u_k) = w_0^{k+1} w_1^k \ldots w_k, \qquad k = 0, 1, \ldots, n.$$

Проверка этого утверждения проводится индукцией с использованием редуцированной системы $\{v_i\}_0^{n-1}$ и того факта, что если $\{u_i\}_0^n$ имеет представление (1.5), то редуцированная система $\{v_i\}_0^{n-1}$ имеет представление (1.10). Таким образом,

$$W(u_0, u_1, \ldots, u_k) = u_0^{k+1} W(v_0, v_1, \ldots, v_{k-1}) = w_0^{k+1} w_1^k \ldots w_k, \quad (1.12)$$

что и завершает доказательство.

З а м е ч а н и е 1.3. Рассмотрение (1.12) показывает, что функции $w_0, \ldots, w_n$ выражаются как отношение некоторых определителей Вронского, а именно

$$w_0 = u_0, \quad w_1 = \frac{W(u_0, u_1)}{u_0^2},$$

$$w_k = \frac{W(u_0, \ldots, u_k)\, W(u_0, \ldots, u_{k-2})}{[W(u_0, \ldots, u_{k-1})]^2}, \qquad k = 2, \ldots, n. \quad (1.13)$$

## § 2. Свойства конуса $C(u_0, u_1, \ldots, u_n)$

Пусть $\{u_i\}_0^n$ — $ECT$-система, заданная на $[a, b]$ в явном виде

$$\begin{aligned}
u_0(t) &= w_0(t),\\
u_1(t) &= w_0(t)\int_a^t w_1(\xi_1)\,d\xi_1,\\
u_2(t) &= w_0(t)\int_a^t w_1(\xi_1)\int_a^{\xi_1} w_2(\xi_2)\,d\xi_2\,d\xi_1,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
u_n(t) &= w_0(t)\int_a^t w_1(\xi_1)\int_a^{\xi_1} w_2(\xi_2)\ldots\int_a^{\xi_{n-1}} w_n(\xi_n)\,d\xi_n\ldots d\xi_1,
\end{aligned} \quad (2.1)$$

где $w_i(t)$ — положительные функции из класса $C^{n-i}[a, b]$. Пусть $D_j$, $j = 0, 1, \ldots, n$, обозначает дифференциальный оператор первого порядка

$$(D_j\varphi)(t) = \frac{d}{dt}\left(\frac{\varphi(t)}{w_j(t)}\right). \tag{2.2}$$

Заметим, что функция $w_n$ является только непрерывной, так что область определения оператора $D_n$ сильно ограничена. Непосредственные выкладки из (2.1) показывают, что

$$\begin{gathered} D_jD_{j-1}\ldots D_0u_{j+1} = w_{j+1} > 0, \qquad j = 0, 1, \ldots, n-1, \\ D_jD_{j-1}\ldots D_0u_j = 0, \qquad j = 0, 1, \ldots, n. \end{gathered} \tag{2.3}$$

Если функция $\varphi$ принадлежит конусу $C(u_0, \ldots, u_n)$, то она обладает существенными свойствами непрерывности и дифференцируемости. Перечислим некоторые из этих свойств для удобства последующих ссылок. Подробное изложение доказательства этих свойств довольно трудоемко и будет отложено до § 11.

A) Если функция $\varphi \in C(u_0, u_1, \ldots, u_n)$ и $n \geqslant 1$, то $\varphi \in C^{n-1}(a, b)$. Далее, $\varphi^{(n-1)}(t)$ имеет производную справа $\varphi_R^{(n)}$ на $(a, b)$, которая непрерывна справа, и производную слева $\varphi_L^{(n)}$, которая непрерывна слева. Для случая $n = 0$ можно лишь утверждать, что если $\varphi \in C(u_0)$, то $\varphi/u_0$ не убывает на $(a, b)$.

B) Пусть $D_{n-1}^R g$ обозначает производную справа функции $g/w_{n-1}$. Если $\varphi \in C(u_0, u_1, \ldots, u_n)$, то функция

$$\frac{(D_{n-1}^R D_{n-2}\ldots D_0\varphi)(t)}{w_n(t)} = \rho(t) \tag{2.4}$$

не убывает и непрерывна справа на $(a, b)$.

Свойства регулярности, присущие функциям из $C(u_0, u_1, \ldots, u_n)$, которые описаны выше и доказаны в § 11, могут быть неизвестны. Соответствующие результаты для случая, когда $u_i(t) = t^i$, получены Поповичиу [1945]. Этот случай значительно проще в в том отношении, что конус выпуклых в обобщенном смысле функций здесь инвариантен по отношению к преобразованиям, которые допускают использование техники возмущений с применением подходящих операций осреднения.

Утверждение B) устанавливает, что любая функция $\varphi$ из конуса $C(u_0, u_1, \ldots, u_n)$ индуцирует вполне аддитивную меру $d\rho(t)$ на интервале $(a, b)$. Установим теперь обратное утверждение. В частности, установим, что если функция $\varphi$, определенная на $(a, b)$, удовлетворяет условиям регулярности A) и такова, что функция, определяемая (2.4), непрерывна справа и не убывает на $(a, b)$, то $\varphi$ принадлежит конусу $C(u_0, u_1, \ldots, u_n)$. Следующая лемма будет необходима в дальнейшем.

Л е м м а 2.1. *Функция*

$$\varphi(t) = \varphi_n(t; x) =$$
$$= \begin{cases} w_0(t) \int\limits_x^t w_1(\xi_1) \int\limits_x^{\xi_1} w_2(\xi_2) \ldots \int\limits_x^{\xi_{n-1}} w_n(\xi_n)\, d\xi_n \ldots d\xi_1, & x \le t \le b, \\ 0, \ t < x \text{ или } t > b, \end{cases} \tag{2.5}$$

*содержится в конусе* $C(u_0, u_1, \ldots, u_n)$. *При* $n = 0$ $\varphi_0(t; x) = w_0(t)$ *для* $x \le t \le b$ *и равна нулю в противном случае.*

*Далее, когда* $a \le x \le t_0 < t_1 < \ldots < t_{n+1} \le b$, *имеем*

$$U\begin{pmatrix} u_0, u_1, \ldots, u_n, \varphi \\ t_0, t_1, \ldots, t_n, t_{n+1} \end{pmatrix} = 0. \tag{2.6}$$

(Заметим, что $\varphi_n(t; a) = u_n(t)$.)

Д о к а з а т е л ь с т в о. Доказательство производится индукцией по $n$. Случай $n = 0$ очевиден.

Предположим, что лемма верна для $n - 1$ при любой $ECT$-системе $\{v_i\}_0^{n-1}$, и рассмотрим определитель

$$U\begin{pmatrix} u_0, \ldots, u_n, \varphi \\ t_0, \ldots, t_n, t_{n+1} \end{pmatrix} = \begin{vmatrix} u_0(t_0) & u_0(t_1) & \ldots & u_0(t_{n+1}) \\ u_1(t_0) & u_1(t_1) & \ldots & u_1(t_{n+1}) \\ \cdot & \cdot & \cdot & \cdot \\ u_n(t_0) & u_n(t_1) & \ldots & u_n(t_{n+1}) \\ \varphi(t_0) & \varphi(t_1) & \ldots & \varphi(t_{n+1}) \end{vmatrix}.$$

Используя представление (2.1), можно вынести за знак определителя $u_0(t_i)$ из $i$-го столбца, $i = 0, 1, \ldots, n + 1$. Далее вычтем из $i$-го столбца предшествующий ему столбец и затем разложим определитель по элементам первой строки. (Заметим, что когда $a \le x \le t_0 < t_1 < \ldots < t_{n+1} \le b$, последние две строки идентичны, так что (2.6) выполняется.)

После элементарных преобразований результирующее выражение записывается в виде

$$\int\limits_{t_0}^{t_1} \int\limits_{t_1}^{t_2} \ldots \int\limits_{t_n}^{t_{n+1}} \left( \prod_{k=0}^{n+1} u_0(t_k) \right) U\begin{pmatrix} v_0, \ldots, v_{n-1}, \tilde{\varphi}_n \\ \eta_0, \ldots, \eta_{n-1}, \eta_n \end{pmatrix} d\eta_n d\eta_{n-1} \ldots d\eta_0,$$

где $v_i = D_0 u_{i-1}$, $i = 0, 1, \ldots, n-1$, и $\tilde{\varphi}_n = D_0 \varphi_n$. Так как подынтегральное выражение неотрицательно по предположению индукции, то доказательство завершено.

З а м е ч а н и е 2.1. Функция $\varphi_n(t; x)$, $n \geqslant 1$, обладает следующими свойствами:

a) Она есть единственное решение уравнения

$$\frac{(D^R_{n-1} D_{n-2} \cdot\ \cdot D_0\varphi)(t)}{w_n(t)} = \rho_x(t),$$

где $\rho_x(t) = \begin{cases} 1, & t \geqslant x, \\ 0, & t < x, \end{cases}$ удовлетворяющее начальным условиям

$$\begin{gathered} \varphi(a) = 0, \\ (D_k D_{k-1} \ldots D_0 \varphi)(a) = 0, \qquad k = 0, 1, \ldots, n-1. \end{gathered} \tag{2.7}$$

b) В § 4 будет доказано, что функции $\varphi_n(t; x)$ образуют крайние лучи конуса $C(u_0, u_1, \ldots, u_n)$ по модулю линейного пространства многочленов $\sum\limits_0^n a_i u_i$.

Л е м м а 2.2.

a) *Пусть $\varphi(t)$ — любая функция на $(a, b)$, обладающая свойствами регулярности* A) *и такая, что функция*

$$\rho(t) = \frac{(D^R_{n-1} D_{n-2} \ldots D_0\varphi)(t)}{w_n(t)} \tag{2.8}$$

*непрерывна справа и не убывает на $(a, b)$. Тогда $\varphi(t)$ допускает представление вида*

$$\varphi(t) = \int\limits_c^b \varphi_n(t; x)\, d\rho(x) +$$

$$+ \sum_{i=0}^n \frac{(D_{i-1} \ldots D_1 D_0 \varphi)(c)}{w_i(c)} \varphi_i(t; c), \qquad a < c \leq t < b, \tag{2.9}$$

*где $\varphi_i(t; x)$ определена в (2.5). (Коэффициент при $\varphi_0(t; c)$ по соглашению выбирается равным $\varphi(c)/w_0(c)$.) Кроме того, правая часть равенства (2.9) не зависит от $c \in (a, t]$.*

b) *Если $\lim\limits_{\substack{c \to a \\ c > a}} \rho(c) = \rho(a+)$ существует и конечен, то значение*

$$\frac{(D_{i-1} \ldots D_0\varphi)(a+)}{w_i(a)} \tag{2.10}$$

*существует для $i = 0, 1, \ldots, n$ и*

$$\varphi(t) = \int\limits_a^t \varphi_n(t; x)\, d\rho(x) + \sum_{i=0}^n \frac{(D_{i-1} \ldots D_0\varphi)(a+)}{w_i(a)} u_i(t).$$

Замечание 2.2. В дальнейшем оператор $D_{n-1}$ будет использоваться для обозначения $D_{n-1}^R$. Отметим, что интервал в (2.9) распространяется только до $t$, так как $\varphi_n(t; x) = 0$ для $x > t$.

Доказательство. На основании определения $\varphi_n(t; x)$, данного в (2.5), путем непосредственных вычислений получаем, что

$$\frac{d}{dx}\varphi_n(t; x) = -w_n(x)\varphi_{n-1}(t; x). \tag{2.11}$$

Очевидно, когда $\rho(t)$ непрерывна справа и не убывает на $(a, b)$, величина

$$\int_c^t \varphi_n(t; x)\, d\rho(x) \tag{2.12}$$

определена для $c \in (a, b)$. Интегрируя (2.12) по частям и принимая во внимание (2.11), получим

$$\int_c^t \varphi_n(t; x)\, d\rho(x) = \rho(x)\varphi_n(t; x)\Big|_c^t - \int_c^t \rho(x)\frac{d}{dx}\varphi_n(t; x)\, dx =$$

$$= -\frac{(D_{n-1}^R \ldots D_0\varphi)(c)\,\varphi_n(t; c)}{w_n(c)} + \int_c^t [(D_{n-1}^R \ldots D_1 D_0\varphi)(x)]\,\varphi_{n-1}(t; x)\, dx.$$

Неоднократно повторенное интегрирование по частям приводит к искомому представлению (2.9).

Для того чтобы доказать (2.10), заметим прежде всего, сравнивая (2.5) и (2.1), что

$$\varphi_k(t; a) = u_k(t), \qquad k = 0, 1, \ldots, n. \tag{2.13}$$

Кроме того, если $\rho(a+)$ существует, то при выполнении последовательных интегрирований, находим, что $(D_{i-1} \ldots D_0\varphi)(a+)/w_i(a)$ существует для каждого $i = 0, 1, \ldots, n$. Теперь представление (2.10) следует из (2.9) при $c \to a$.

Представление в виде (2.9) обобщает обычную формулу Тейлора с остаточным членом. Эта формула находит применение в численном анализе и теории интерполяции. (В случае функций вида $u_i(t) = t^i$ см., например, Сард [1949].)

Лемма 2.3. *Если функция $\rho(t)$ возрастает и непрерывна справа на $(a, \beta) \subset (a, b)$, то любые два решения уравнения*

$$d\left[\frac{(D_{n-1}^R D_{n-2} \ldots D_0\varphi)(t)}{w_n(t)}\right] = d\rho(t) \tag{2.14}$$

*различаются на $(\alpha, \beta)$ на многочлен, т. е. на функцию вида* $\sum_{k=0}^{n} c_k u_k(t)$.

Доказательство. Достаточно показать, что единственными решениями уравнения

$$d\left[\frac{(D^R_{n-1}D_{n-2}\dots D_0\varphi)(t)}{w_n(t)}\right]=0, \qquad t\in(\alpha,\beta) \tag{2.15}$$

служат многочлены. Пусть $\varphi$ — решение уравнения (2.15), для фиксированного $t_0\in(\alpha,\beta)$ пусть $u$ — многочлен, удовлетворяющий условиям

$$\begin{gathered} u(t_0)=\varphi(t_0), \\ (D_i\dots D_1D_0u)(t_0)=(D_i\dots D_1D_0\varphi)(t_0), \qquad i=0,1,\dots,n-1. \end{gathered} \tag{2.16}$$

Задание многочлена $u$, удовлетворяющего условиям (2.16), возможно, так как соответствующая система линейных уравнений является треугольной с определителем, равным $\prod_{j=0}^{n} w_j(t_0)>0$.

Теперь $\bar\varphi=u-\varphi$ удовлетворяет уравнению (2.15), и

$$\begin{gathered} \bar\varphi(t_0)=0, \\ (D_i\dots D_1D_0\bar\varphi)(t_0)=0, \qquad i=0,1,\dots,n-1. \end{gathered} \tag{2.17}$$

Поэтому последовательные интегрирования тождества

$$d\left[\frac{(D^R_{n-1}\dots D_1D_0\bar\varphi)(t)}{w_n(t)}\right]\equiv 0$$

с нижним пределом $t_0$, приводят к выводу, что $\bar\varphi(t)=u(t)-\varphi(t)=0$ для всех $t\in(\alpha,\beta)$.

Теорема 2.1. *Если* $n\geqslant 1$, *то функция* $\varphi$, *определенная на* $(a,b)$, *содержится в конусе* $C(u_0,u_1,\dots,u_n)$ *тогда и только тогда, когда* $\varphi$ *удовлетворяет свойствам регулярности* A), *а связанная с ней функция*

$$\rho(t)=\frac{(D^R_{n-1}D_{n-2}\dots D_0\varphi)(t)}{w_n(t)}$$

*непрерывна справа и не убывает на* $(a,b)$.

Доказательство. Необходимость в точности соответствует утверждениям A) и B), доказательство которых отложено до § 11. Чтобы доказать достаточность, необходимо проверить, что

$$U\begin{pmatrix}u_0, & \dots, & u_n, & \varphi\\ t_0, & \dots, & t_n, & t_{n+1}\end{pmatrix}\geqslant 0 \tag{2.18}$$

для всех наборов $\{t_i\}_0^{n+1}$, удовлетворяющих условию $a<t_0<t_1<\dots<t_{n+1}<b$.

На основании леммы 2.2 запишем

$$\varphi(t) = \varphi_1(t) + \varphi_2(t), \tag{2.19}$$

где

$$\varphi_1(t) = \int_c^b \varphi_n(t; x)\, d\rho(x), \qquad a < c < t_0, \tag{2.20}$$

$$\varphi_2(t) = \sum_{i=0}^{n} \frac{(D_{i-1} \ldots D_1 D_0 \varphi)(c)}{w_i(c)} \varphi_i(t; c). \tag{2.21}$$

Теперь по лемме 2.1 $\varphi_n(t; x) \in C(u_0, u_1, \ldots, u_n)$, поэтому

$$U\begin{pmatrix} u_0, \ldots, u_n, & \varphi_1 \\ t_0, \ldots, t_n, & t_{n+1} \end{pmatrix} = \int_c^b U\begin{pmatrix} u_0, \ldots, u_n, & \varphi_n(\cdot; x) \\ t_0, \ldots, t_n, & t_{n+1} \end{pmatrix} d\rho(x) \geqslant 0. \tag{2.22}$$

Далее, так как

$$d\left[\frac{(D_{n-1}^R \ldots D_1 D_0 \varphi_2)(t)}{w_n(t)}\right] = 0 \text{ для } t > c,$$

то $\varphi_2(t) = \sum_{i=0}^{n} a_i u_i(t)$ для $t \in (c, b)$. Но $t_0 > c$ за счет выбора $c$, поэтому

$$U\begin{pmatrix} u_0, \ldots, u_n, & \varphi_2 \\ t_0, \ldots, t_n, & t_{n+1} \end{pmatrix} = 0. \tag{2.23}$$

Неравенство (2.18) теперь немедленно следует из (2.22) и (2.23).

Следствие 2.1. *Для $n \geqslant 1$ и $a < c < b$ любая функция $\varphi \in C(u_0, u_1, \ldots, u_n)$ допускает представление вида*

$$\varphi(t) = \int_c^t \varphi_n(t; x)\, d\rho(x) + \sum_{i=0}^{n} \frac{(D_{i-1} \ldots D_1 D_0 \varphi)(c)}{w_i(x)} \varphi_i(t; c), \tag{2.24}$$

*справедливое для $t \in [c, b]$, где*

$$d\rho(x) = d\left[\frac{D_{n-1}^R \ldots D_1 D_0 \varphi(x)}{w_n(x)}\right]$$

*есть неотрицательная мера.*

*Если $\rho(a+) > -\infty$, то величина $c$ может быть заменена на $a+$. Полученное представление тогда справедливо на всем интервале $(a, b)$.*

Замечание 2.3. Для $n = 0$ представление (2.24) имеет место только для регуляризованной функции $\tilde{\varphi}$, полученной путем измерения $\varphi$ таким образом, что $\tilde{\varphi}/u_0$ непрерывна справа. Такое изменение очевидно возможно, так как $\varphi/u_0$ возрастает и, следовательно, допускает самое большее счетное число разрывов

$t_1, t_2, \ldots$ Если $\rho(x) = \tilde{\varphi}(x)/u_0(x)$, то представление функции $\varphi$ принимает вид

$$\varphi(t) = \int_c^b \varphi_0(t; x)\, d\rho(x) + \frac{\tilde{\varphi}(c)}{w_0(c)} \varphi_0(t; c) + \sum_{i=1}^{\infty} [\varphi(t_i) - \tilde{\varphi}(t_i)]\, \delta(t; t_i),$$

где $\delta(t; t_i) = \begin{cases} 1, & t = t_i, \\ 0, & t \neq t_i. \end{cases}$

Замечание 2.4. Параллельно с доказательством (2.24) мы могли бы работать с «сопряженными» функциями

$$\tilde{\varphi}_k(t; x) = \begin{cases} w_0(t) \int_t^x w_1(\xi_1) \int_{\xi_1}^x w_2(\xi_2) \ldots \int_{\xi_{k-1}}^x w_k(\xi_k)\, d\xi_k \ldots d\xi_1 & \text{при } t \leqslant x,\ k = 0, 1, \ldots, n, \\ 0 \text{ при } t > x. \end{cases} \tag{2.25}$$

Интегрирование по частям, как в лемме 2.2, тогда приводит к представлению вида

$$\varphi(t) = (-1)^{n+1} \int_t^c \tilde{\varphi}_n(t; x)\, d\rho(x) + \sum_{i=0}^{n} (-1)^{n-i} \frac{(D_{i-1} \ldots D_0 \varphi)(c)}{w_i(c)} \tilde{\varphi}_i(t; c), \tag{2.26}$$

которое справедливо для $t \in (a, c]$, $c \in b$. В случае, когда $\rho\ (b-) < \infty$, величина $c$ может быть заменена на $b-$, и данное представление будет справедливо на всем $(a, b)$.

Замечание 2.5. Стоит обратить особое внимание на представление (2.24) в специальном случае $u_i(t) = t^i$, $i = 0, 1, \ldots, n$. Для удобства примем $[a, b] = [0, 1]$. Следующие соотношения легко проверяются:

$$w_0(t) \equiv 1, \quad w_k(t) \equiv k, \qquad k = 1, 2, \ldots, n,$$

$$\varphi_n(t; x) = \begin{cases} (t - x)^n, & t \geqslant x, \\ 0, & t < x, \end{cases}$$

$$\frac{1}{w_k(t)} (D_{k-1} \ldots D_1 D_0 \varphi)(t) = \frac{1}{k!} \frac{d^k \varphi(t)}{dt^k} = \frac{1}{k!} \varphi^k(t), \qquad k = 1, 2, \ldots, n.$$

Представление (2.24) функции $\varphi \in C(1, t, \ldots, t^n)$ тогда приводится к виду

$$\varphi(t) = \int_c^t \frac{(t - x)^n}{n!}\, d\bar{\rho}(x) + \sum_{i=0}^{n} \frac{\varphi^{(i)}(c)}{i!} (t - c)^i \quad \text{для} \quad 0 < c < t < 1,$$

где $\bar{\rho}(x) = \varphi^{(n)}(x)$. Это, как легко видно, есть известная формула интегрального исчисления.

З а м е ч а н и е 2.6. Классический результат Бернштейна устанавливает, что если функция $f$, определенная на $[a, b]$, имеет непрерывную $(n+1)$-ю производную, то

$$\left\| f(t) - \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(t-a)^i \right\| \leqslant C \| f^{(n+1)} \|,$$

где $\|f\| = \sup\limits_{a \leqslant t \leqslant b} |f(t)|$ и $C$—константа, которая не зависит от $f$. Таким образом, каждая функция $f \in C^{n+1}$ может быть аппроксимирована многочленом степени $n$ с ошибкой, зависящей только от величины $\|f^{(n+1)}\|$. С целью обобщения этого результата на случай, когда мы аппроксимируем многочленами вида $\sum\limits_{i=0}^{n} a_i u_i$, предположим, что $D_n D_{n-1} \ldots D_0 f$ непрерывна на $[a, b]$. Повторное интегрирование по частям показывает, что

$$f(t) - \sum_{i=0}^{n} a_i^* u_i(t) = \int_a^b \varphi_n(t; x)(D_n \ldots D_0 f)(x)\,dx,$$

где $a_0^* = f(a)/w_0(a)$ и $a_i^* = [(D_{i-1} \ldots D_0 f)(a)]/w_i(a)$, $i = 1, 2, \ldots, n$. Следовательно,

$$\left\| f - \sum_{i=0}^{n} a_i^* u_i \right\| = \sup_{a \leqslant t \leqslant b} \left| \int_a^b \varphi_n(t; x)(D_n \ldots D_1 D_0 f)(x)dx \right| \leqslant$$

$$\leqslant \sup_{a \leqslant t \leqslant b} \sup_{a \leqslant x \leqslant b} |(D_n \ldots D_1 D_0 f)(x)| \int_a^b \varphi_n(t; x)\,dx = C_0 \| D_n \ldots D_1 D_0 f \|,$$

где $C_0 = \int_a^b \varphi_n(b; x)\,dx$. Так как мы все гда можем выбрать $f(x)$ так, что $(D_n \ldots D_1 D_0 f)(x) \equiv 1$, то константа $C_0$ в неравенстве является наилучшей возможной. Рассмотренное выше обобщение неравенства Бернштейна было представлено нам А. Зедеком, который предложил другое, более сложное доказательство.

Следующий результат является отчасти обратным по отношению к теореме 2.1.

Т е о р е м а 2.2. *Если функция* $\rho(x)$ *непрерывна справа, неубывающая и ограниченной вариации на* $(a, b)$, *то*

$$\varphi(t) = \int_a^b \varphi_n(t; x)d\rho(x) \tag{2.27}$$

*содержится в конусе* $C(u_0, u_1, \ldots u_n)$ *и*

$$d\left[\frac{(D_{n-1}^R \ldots D_0 \varphi)(t)}{w_n(t)}\right] = d\rho(t). \tag{2.28}$$

*В более частном случае, если* $d\rho(x) = f(x)\,dx$, *где* $f(x)$ *неотрицательна и непрерывна на* $[a, b]$, *то*

$$D_n \ldots D_1 D_0 \varphi(t) = f(t).$$

Доказательство. Функция $\varphi(t)$ вида (2.27), очевидно, принадлежит конусу $C(u_0, u_1, \ldots, u_n)$. Кроме того,

$$\frac{(D_{n-2} \ldots D_0\varphi)(t)}{w_{n-1}(t)} = \int_a^b \frac{D_{n-2} \ldots D_0\varphi_n(t; x)}{w_{n-1}(t)}\, d\rho(x) =$$

$$= \int_a^t \left( \int_x^t w_n(\xi)\, d\xi \right) d\rho(x) = \int_a^t w_n(\xi) \left( \int_a^\xi d\rho(x) \right) d\xi.$$

Значит,

$$(D_{n-1}^R \ldots D_0\varphi)(t) = w_n(t) \int_a^t d\rho(x),$$

поэтому

$$d\left[\frac{(D_{n-1}^R \ldots D_0\varphi)(t)}{w_n(t)}\right] = d\rho(t).$$

Теперь перейдем к вопросу об аппроксимации функции $\varphi$, содержащейся в конусе $C(u_0, u_1, \ldots, u_n)$ функциями из класса $C^n[a, b]$, также принадлежащими $C(u_0, u_1, \ldots, u_n)$. Для этой цели в дальнейшем будет необходима следующая лемма.

Лемма 2.4. *Пусть функции* $\varphi_1$ *и* $\varphi_2$ *принадлежат конусу* $C(u_0, u_1, \ldots, u_n)$. *Предположим также, что*

1) $d\rho_1 \geqslant d\rho_2$, *т.е.* $\rho_1(\beta) - \rho_1(\alpha) \geqslant \rho_2(\beta) - \rho_2(\alpha)$ *для всех* $\alpha \leqslant \beta$, $\alpha, \beta \in (a, b)$;

2) *для некоторого* $c \in (a, b)$

$$\varphi_1(c) = \varphi_2(c),$$

$$(D_i \ldots D_1 D_0 \varphi_1)(c) = (D_i \ldots D_1 D_0 \varphi_2)(c), \qquad i = 0, 1, \ldots, n-1. \tag{2.29}$$

*Тогда* $\varphi_1(t) - \varphi_2(t) \geqslant 0$ *для* $t \in (c, b)$ *и* $(-1)^{n+1}[\varphi_1(t) - \varphi_2(t)] \geqslant 0$ *для* $t \in (a, c)$.

Доказательство. Пусть $t \in [c, b)$. Так как $\varphi_n(t; x) \geqslant 0$, то из (2.24) получим, что

$$\varphi_1(t) = \int_c^b \varphi_n(t; x)\, d\rho_1(x) + \sum_{i=0}^{n} \frac{(D_{i-1} \ldots D_0\varphi_1)(c)}{w_i(c)} \varphi_i(t; c) \geqslant$$

$$\geqslant \int_c^b \varphi_n(t; x)\, d\rho_2(x) + \sum_{i=0}^{n} \frac{(D_{i-1} \ldots D_0\varphi_2)(c)}{w_i(c)} \varphi_i(t; c) = \varphi_2(t).$$

Следовательно, $\varphi_1(t) \geqslant \varphi_2(t)$ для $t \in [c, b)$. Соотношение $(-1)^{n+1}[\varphi_1(t) - \varphi_2(t)] \geqslant 0$ для $t \in (a, c)$ получается путем сравнения (2.26) и (2.29).

Теорема 2.3. *Для любой функции* $\varphi \in C(u_0, u_1, \ldots, u_n)$

a) *существует функция* $\varphi(t; \varepsilon) \in C(u_0, u_1, \ldots, u_n)$ *такая, что*

$$\frac{D_{n-1}^R \ldots D_1 D_0 \varphi(a+; \varepsilon)}{w_n(t)} > -\infty$$

(см. следствие 2.1),

$$\varphi(t; \varepsilon)\begin{cases} = \varphi(t), & t \in (a+\varepsilon, b-\varepsilon), \\ \leqslant \varphi(t), & t \in [b-\varepsilon, b) \end{cases}$$

*и*

$$(-1)^{n+1}[\varphi(t) - \varphi(t; \varepsilon)] \geqslant 0, \quad t \in (a, a+\varepsilon],$$

b) *существует последовательность функций* $\{\varphi_k\}_1^\infty$ *такая, что* $\varphi_k \in C^n[a, b] \cap C(u_0, u_1, \ldots u_n)$,

$$\lim_{k\to\infty} \varphi_k^{(i)}(t) = \varphi^{(i)}(t), \ i = 0, 1, \ldots, n-1 \ (\varphi^{(n-1)} = \varphi_R^{(n-1)})$$

*равномерно на каждом компактном подынтервале* $(a, b)$.

Замечание 2.7. Предположение относительно системы функций $\{u_i\}_0^n$ гарантирует, что связанные с ними функции $w_i$ (посредством формулы (2.1)) принадлежат классу $C^{n-i}$ на $[a, b]$. Если предположить несколько большее, а именно, что $w_i \in C^{n-i+1}$ на $[a, b]$, то доказательство может быть продолжено с тем, чтобы показать, что функции $\varphi_k$ в части b) могут быть выбраны из класса $C^{n+1}$.

Замечание 2.8. Когда $C(u_0, \ldots, u_n) = C(1, t)$, т. е. конус $C(u_0, \ldots, u_n)$ совпадает с классом выпуклых функций на $(a, b)$, функция $\varphi(t; \varepsilon)$, построенная в части a), соответствует замене $\varphi(t)$ подходящими линейными отрезками на $(a, a+\varepsilon]$ и $[b-\varepsilon, b)$.

Доказательство. Для каждой функции $\varphi \in C(u_0, u_1, \ldots, u_n)$, как известно, справедливо представление вида

$$\varphi(t) = \int_C^t \varphi_n(t; x)\, d\rho(x) + \sum_{i=0}^n \frac{(D_{i-1} \ldots D_0 \varphi)(c)}{w_i(c)} \varphi_i(t; c),$$

где $a < c < t < b$ и $\rho(t) = \dfrac{(D_{n-1}^R \ldots D_1 D_0 \varphi)(t)}{w_n(t)}$. Видоизменим меру $d\rho$, определенную на $(a, b)$, следующим образом

$$\rho(t; \varepsilon) = \begin{cases} \rho(t), & t \in [a+\varepsilon, b-\varepsilon), \\ \rho(a+\varepsilon), & t \in [a, a+\varepsilon), \\ \rho(b-\varepsilon), & t \in [b-\varepsilon, b). \end{cases}$$

Таким образом, $d\rho(t; \varepsilon)$ имеет нулевую меру около граничных точек $a$ и $b$. Пусть

$$g(t; \varepsilon) = \int_a^b \varphi_n(t; x)\, d\rho(x; \varepsilon).$$

Так как

$$d\left[\frac{D_{n-1}^R \ldots D_0 g(t;\ \varepsilon)}{w_n(t)}\right] = d\rho(t;\ \varepsilon) = d\left[\frac{D_{n-1}^R \ldots D_0 \varphi(t)}{w_n(t)}\right]$$

для $t \in (a+\varepsilon,\ b-\varepsilon)$, то по лемме 2.3 $g(t;\ \varepsilon)$ и $\varphi(t)$ различаются на $(a+\varepsilon,\ b-\varepsilon)$ на многочлен $\sum_{i=0}^{n} a_i(\varepsilon)\, u_i(t)$.

Следовательно, функция

$$\varphi(t;\ \varepsilon) = \int_a^b \varphi_n(t;\ x)\, d\rho(x;\ \varepsilon) + \sum_{i=0}^{n} a_i(\varepsilon)\, u_i(t) \tag{2.30}$$

удовлетворяет требованиям a); неравенства, содержащие $\varphi(t;\ \varepsilon)$ и $\varphi(t)$, следуют из леммы 2.4.

Перейдем теперь к доказательству утверждения b). Прежде всего изменим функцию $\varphi$, как при доказательстве утверждения a), чтобы получить функцию $\varphi(t;\ \varepsilon)$, определенную в (2.30). Далее обозначим через $\{\rho_m(t;\ \varepsilon)\}_{m=1}^{\infty}$ последовательность возрастающих функций из класса непрерывных функций $C^1[a,\ b]$, которая слабо сходится к $\rho(t;\ \varepsilon)$ и удовлетворяет условию

$$\rho_m(t;\ \varepsilon) = \rho(t;\ \varepsilon) \text{ для } t \in (a+\varepsilon,\ b-\varepsilon).$$

Поскольку

$$\varphi_n(t;\ x) = \begin{cases} w_0(t)\int\limits_x^t w_1(\xi_1)\ldots \int\limits_x^{\xi_{n-1}} w_n(\xi_n)\, d\xi_n \ldots d\xi_1, & t \geqslant x, \\ 0, & t < x, \end{cases} \tag{2.31}$$

и $w_i \in C^{n-i}[a,\ b]$, $i = 0,\ 1,\ \ldots,\ n$, то мы получаем, что:

1) производные

$$\frac{d^j}{dt^j}\varphi_n(t;\ x), \qquad j = 0,\ 1,\ \ldots,\ n-1,$$

непрерывны по $x \in [a,\ b]$ и

2) существует постоянная $C$ такая, что

$$\left|\frac{d^j}{dt^j}\varphi_n(t;x)\right| \leq C$$

для всех $t$, $x \in [a,\ b]$ и $j = 0,\ 1,\ldots,\ n-1$. Следовательно,

$$\frac{d^j}{dt^j}\int_a^b \varphi_n(t;x)\, d\rho(x) = \int_a^b \frac{d^j}{dt^j}\varphi_n(t;x)\, d\rho(x), \qquad j = 0, 1, \ldots, n-1 \tag{2.32}$$

для любой функции $\rho(x)$ ограниченной вариации. В частности, равенство (2.32) выполняется для функций $\rho(t;\ \varepsilon)$ и $\rho_m(t;\ \varepsilon)$.

Теперь положим $g_m(t;\ \varepsilon) = \int\limits_a^b \varphi_n(t;\ x)\, d\rho_m(x;\ \varepsilon)$.

Так как $\rho_m(t;\varepsilon)$ слабо сходится к $\rho(t;\varepsilon)$, то на основании свойств 1) и 2) заключаем, что для некоторого $m(\varepsilon)$ и любого $t\in(a,b)$ имеем

$$|g^{(j)}(t;\varepsilon)-g_m^{(j)}(t;\varepsilon)|=\left|\int_a^b \varphi_n^{(j)}(t;x)\,d\rho(x;\varepsilon)-\int_a^b \varphi_n^{(j)}(t;x)\,d\rho_m(x;\varepsilon)\right|\leqslant\varepsilon$$

для $m\geqslant m(\varepsilon)$, $j=0,1,\ldots,n-1$.

Для каждого $k=1,2,\ldots$ выберем $\varepsilon_k<1/k$, $m_k\geqslant m(\varepsilon_k)$ и положим

$$\varphi_k(t)=\int_a^b \varphi_n(t;x)\,d\rho_{m_k}(x;\varepsilon_k)+\sum_{i=0}^n a_i(\varepsilon_k)\,u_i(t), \tag{2.33}$$

где коэффициенты $a_i(\varepsilon_k)$ определены, как в (2.30). Тогда

$$|\varphi^{(j)}(t;\varepsilon_k)-\varphi_k^{(j)}(t)|=|g^{(j)}(t;\varepsilon_k)-g_{m_k}^{(j)}(t;\varepsilon_k)|\leqslant\varepsilon_k.$$

Для фиксированного $\delta>0$ выберем $k^{-1}<\delta$. Так как $\varepsilon_k<\frac{1}{k}<\delta$ и $\varphi(t;\varepsilon_k)=\varphi(t)$ для $t\in(a+\varepsilon_k, b-\varepsilon_k)$, то заключаем, что

$$|\varphi^{(j)}(t)-\varphi_k^{(j)}(t)|=|\varphi^{(j)}(t;\varepsilon_k)-\varphi_k^{(j)}(t)|\leqslant\varepsilon_k$$

для любого $t\in[a+\delta, b-\delta]$. Значит, последовательность $\{\varphi_k\}_1^\infty$ удовлетворяет требованиям b) при условии, что $\varphi_k$ принадлежит классу непрерывных функций $C^n[a,b]$. Кроме того, $n$-я производная от $\varphi_n(t;x)$ по $t$ непрерывна и ограничена, за исключением $x=t$. Следовательно, для любого $m$ и $\varepsilon$ функция

$$\frac{d^n}{dt^n}g_m(t;\varepsilon)=\int_a^b\frac{d}{dt^n}\varphi_n(t;x)\,d\rho_m(x;\varepsilon)$$

непрерывна по $t\in[a,b]$, и значит, $\varphi_k\in C^n[a,b]$, как и требовалось.

## § 3. Дальнейшие свойства конуса $C(u_0, u_1, \ldots, u_n)$

Пусть $\{u_i\}_0^n$ — $ECT$-система на $[a,b]$ вида (2.1). В § 1 и 2 была найдена удобная редуцированная система

$$v_i(t)=(D_0u_{i+1})(t)=\frac{d}{dt}\,\frac{u_{i+1}(t)}{w_0(t)},\qquad i=0,1,\ldots,n-1.$$

Из теоремы 1.2 и тождества (см. (1.4))

$$W(u_0,\ldots,u_n)=w_0^{n+1}\cdot W(D_0u_1, D_0u_2,\ldots,D_0u_n) \tag{3.1}$$

видно, что система $\{D_0u_{i+1}\}_0^{n-1}$ является $ECT$-системой на $[a,b]$. Это наводит на мысль рассмотреть конус $C(D_0u_1, D_0u_2,\ldots,D_0u_n)$. Заметим, что система $\{D_0u_{i+1}\}_0^{n-1}$ допускает представление вида (2.1), в котором $w_0, w_1,\ldots,w_n$ заменяются на $w_1, w_2,\ldots,w_n$.

Далее по теореме 2.2 $\varphi$ принадлежит $C(u_0, u_1, \ldots, u_n)$ тогда и только тогда, когда

$$\rho(t) = \frac{D^R_{n-1} D_{n-2} \ldots D_1 (D_0 \varphi(t))}{w_n(t)} \tag{3.2}$$

непрерывна справа и не убывает. Отсюда немедленно следует, что если $n \geqslant 2$, то функция $\varphi$ принадлежит $C(u_0, u_1, \ldots, u_n)$ тогда и только тогда, когда $D_0\varphi \in C(D_0u_1, \ldots, D_0u_n)$.

Очевидно, можно продолжать процесс «редуцирования», в котором система $\{D_0u_{i+1}\}_0^{n-1}$ называется первой редуцированной системой. Тогда $i$-я редуцированная система на основе системы $\{u_j\}_0^n$ состоит из функций

$$D_{i-1} \ldots D_0 u_j, \qquad j = i,\ i+1, \ldots, n. \tag{3.3}$$

Резюмируем вышеприведенное рассуждение в следующей теореме.

Теорема 3.1. *Если $\{u_i\}_0^n$ — ECT-система на $[a, b]$, то для $n \geqslant 2$ функция $\varphi$ принадлежит $C(u_0, \ldots, u_n)$ только и только тогда, когда $D_{i-1}D_{i-2} \ldots D_0\varphi$ содержится в конусе, образованном функциями* (3.3). *Кроме того, для $n \geqslant 1$ $D_{n-1}D_{n-2} \ldots D_0\varphi \in C(w_n)$ всякий раз, когда $\varphi \in C(u_0, \ldots, u_n)$.*

В добавление к конусам, связанным с редуцированными системами, определенными в (3.3), стоит также одновременно исследовать конусы $C(u_0, u_1, \ldots, u_k)$, $k = 0, 1, \ldots, n$.

Интерес к пересечению конусов типа $\bigcap\limits_{i=k}^{n} C(u_0, \ldots, u_i)$ или $\bigcap\limits_{i=0}^{k} C(u_0, \ldots, u_i)$ возникает в различных контекстах. Например, если $u_0(t) \equiv 1$ и $u_1(t) = t$ на $[a, b]$, то $C(u_0) \bigcap C(u_0, u_1)$ состоит из всех неубывающих выпуклых функций на $[a, b]$.

Предположим, что имеется бесконечная система функций $u_0, u_1, \ldots, u_n, \ldots$ такая, что $\{u_i\}_0^k$ является $ET$-системой для любого $k = 0, 1, \ldots$ Пусть $C^+$ обозначает конус неотрицательных функций на $(a, b)$. В случае, когда $u_i(t) = t^i$, $i = 0, 1, \ldots$, конус $C^+ \bigcap [\bigcap\limits_{i=0}^{\infty} C(u_0, \ldots$ $\ldots, u_i)]$ состоит из всех функций, удовлетворяющих условию $\varphi^{(i)}(t) \geqslant 0$, $i = 0, 1, \ldots$ (см. ниже теорему 3.2). Аналогично конус $C^+ \bigcap [\bigcap\limits_{i=0}^{\infty} C(u_0, \ldots, u_i)]$ в этом специальном случае точно совпадает с классом абсолютно монотонных функций на $(a, b)$. В § 8 будет исследован аналог результата для обобщенных $ET$-систем, согласно которому любая абсолютно монотонная функция допускает разложение в степенной ряд с неотрицательными коэффициентами.

С помощью теоремы 3.1 получаем следующий результат.

Теорема 3.2. *Пусть* $\{u_i\}_0^n$ — *ECT-система вида* (2.1). *Тогда для* $n \geqslant 1$ *и* $k = 0, 1, \ldots, n$ *функция* $\varphi$ *принадлежит* $\bigcap_{i=k}^{n} C(u_0, \ldots, u_i)$ *тогда и только тогда, когда*

$$D_i D_{i-1} \ldots D_0 \varphi \geqslant 0, \qquad i = k, k+1, \ldots, n-1, \tag{3.4}$$

*и функция* $\rho(t)$, *задаваемая соотношением*

$$\frac{D_{n-1}^R D_{n-2} \ldots D_0 \varphi(t)}{w_n(t)} = \rho(t), \tag{3.5}$$

*непрерывна справа и не убывает.*

Следствие 3.2. a) *Если* $\varphi_n(t; x)$ *определена посредством соотношения*

$$\varphi_n(t; x) = \begin{cases} w_0(t) \int\limits_x^t w_1(\xi_1) \int\limits_x^{\xi_1} w_2(\xi_2) \ldots \int\limits_x^{\xi_{n-1}} w_n(\xi_n)\, d\xi_n \ldots d\xi_1, & x \leq t, \\ 0, & t < x, \end{cases} \tag{3.6}$$

*то*

$$\varphi_n(t; x) \in \bigcap_{i=0}^{n} C(u_0, \ldots, u_i). \tag{3.7}$$

(b) *Для любого* $j = 0, 1, \ldots, n$ *имеем*

$$u_j \in \bigcap_{i=0}^{n} C(u_0, \ldots, u_i). \tag{3.8}$$

Доказательство. По лемме 2.1 $\varphi_n(t; x) \in C(u_0, \ldots, u_n)$, поэтому функция (3.5) непрерывна справа и не у ывает. Далее дифференцирование (3.6) показывает, что условие (3.4) выполняется. Следовательно, (3.7) справедливо. Утверждение (b) следует из аналогичного рассуждения.

Замечание 3.1 Заметим, что для функции $\varphi \in \bigcap_{i=k}^{n} C(u_0, \ldots, u_i)$, $k \leq n-1$, имеем $D_{n-1} \ldots D_0 \varphi(t) \geqslant 0$. В этом случае $\rho(t) \geqslant 0$, так что $\rho(a+) \geqslant 0$. Следовательно, применение леммы (2.2) (b) показывает, что функция $\varphi$ представима в виде

$$\varphi(t) = \int\limits_a^b \varphi_n(t; x)\, d\rho(x) + \sum_{i=0}^{n} \frac{D_{i-1} \ldots D_1 D_0 \varphi(a+)}{w_i(a)} u_i(t).$$

В §§ 3 и 4 большей частью основываются на работе Циглера [1966], за исключением некоторых усилений.

## § 4. Крайние лучи конусов

В этом параграфе мы определим крайние лучи конуса $C(u_0, u_1, \ldots, u_n)$, а также крайние лучи для пересечения конусов $C(u_0, u_1, \ldots, u_k)$, $k = 0, 1, \ldots, n$.

Начнем с исследования природы крайних лучей для конуса $C(u_0, u_1, \ldots, u_n)$. Очевидно, что этот конус не может иметь крайние лучи в обычном смысле. Действительно, так как $\pm u_i(t) \in C(u_0, u_1, \ldots, u_n)$, то любая функция $\varphi \in C(u_0, u_1, \ldots, u_n)$ может быть тривиально записана как

$$\varphi(t) = \frac{(\varphi(t) + u(t)) + (\varphi(t) - u(t))}{2},$$

где $u(t) = \sum_{i=0}^{n} a_i u_i(t)$ — любой нетривиальный многочлен.

Поэтому удобно ввести понятие крайних лучей по модулю линейного пространства многочленов $K$. В частности, говорят, что функция $\varphi$ является крайним лучом конуса $C(u_0, \ldots, u_n)$ по момодулю $K$ тогда и только тогда, когда $\varphi$ не может быть представлена в виде $\varphi = \varphi_1 + \varphi_2$, где $\varphi_1$ и $\varphi_2$ принадлежат конусу $C(u_0, u_1, \ldots, u_n)$ и $\beta\varphi_1 \neq \alpha\varphi_2 + u$ для некоторого многочлена $u$ и вещественных чисел $\alpha$ и $\beta$.

Теорема 4.1. *Крайние лучи конуса $C(u_0, u_1, \ldots, u_n)$ $(n \geqslant 1)$ по модулю линейного пространства, натянутого на систему функций $\{u_i\}_0^n$, состоят из решений уравнения*

$$d\left[\frac{(D_{n-1}^R D_{n-2} \ldots D_0 \varphi)(t)}{w_n(t)}\right] = \alpha\, d\rho_x(t), \qquad a < t < b,\ \alpha > 0, \tag{4.1}$$

*где* $\rho_x(t) = \begin{cases} 1, & t \geqslant x, \\ 0, & t < x. \end{cases}$

Замечание 4.1. Случай $n = 0$ заслуживает специального рассмотрения, так как при этом функция $\varphi(t)/w_0(t)$ не убывает, но не обязательно непрерывна справа. Если мы регуляризуем функцию $\varphi$ так, что функция $\varphi/w_0$ непрерывна справа, то крайними лучами конуса $C(u_0)$ являются функции, регуляризация которых удовлетворяет уравнению

$$d\left[\frac{\varphi(t)}{u_0(t)}\right] = \alpha\, d\rho_x(t). \tag{4.2}$$

В частности, крайние лучи конуса $C(u_0)$ совпадают почти всюду с решениями (4.2).

Доказательство. Предположим прежде всего, что функция удовлетворяет уравнению (4.1) (для определенности примем $\alpha = 1$), и

$$\varphi = \varphi_1 + \varphi_2, \tag{4.3}$$

где $\varphi_1$ и $\varphi_2$ содержатся в конусе $C(u_0, u_1, \ldots, u_n)$.

Так как $\varphi(t)$ удовлетворяет уравнению (4.1), имеем

$$d\rho_x(t) = d\rho_1(t) + d\rho_2(t), \tag{4.4}$$

где

$$d\rho_i(t) = d\left[\frac{(D^R_{n-1} \ldots D_0\varphi_i)(t)}{w_n(t)}\right], \qquad i = 1, 2,$$

— неотрицательные меры на $(a, b)$.

Из соотношения (4.4) видно, что носитель $d\rho_i(t)$ сосредоточен в точке $x$, поэтому

$$d\rho_i(t) = \lambda_i d\rho_x(t), \text{ где } 0 \leqslant \lambda_i \leqslant 1, \qquad \lambda_1 + \lambda_2 = 1, \qquad i = 1, 2.$$

В этом случае по лемме 2.3

$$\varphi_i(t) = \lambda_i\varphi(t) + \sum_{j=0}^{n} a_{ji}u_j(t), \quad i = 1, 2,$$

и, следовательно, функции $\varphi_1$ и $\varphi_2$ с точностью до постоянного множителя отличаются друг от друга на многочлен, т. е. $\beta\varphi_1 = \alpha\varphi_2 + u$ для некоторого многочлена $u$ и вещественных чисел $\alpha$ и $\beta$. Это доказывает, что решения уравнения (4.1) определяют крайние лучи конуса $C(u_0, u_1, \ldots, u_n)$.

Пусть теперь $\bar{\varphi}(t)$ — решение уравнения

$$d\left[\frac{(D^R_{n-1} \ldots D_0\bar{\varphi})(t)}{w_n(t)}\right] = d\rho(t) \geqslant 0,$$

где носитель $d\rho(t)$ содержит по крайней мере две точки $t_1$ и $t_2$ в $(a, b)$. Пусть $I_1$ и $I_2$ — непересекающиеся интервалы такие, что $t_j \in I_i$, $i = 1, 2$, $I_1 \cup I_2 = (a, b)$ и $I_1 < I_2$, т. е. $I_1$ лежит слева от $I_2$. Определим

$$d\rho_i(t) = \begin{cases} d\rho(t), & t \in I_i, \quad i = 1, 2, \\ 0 & \text{в противном случае,} \end{cases}$$

так что

$$d\rho(t) = d\rho_1(t) + d\rho_2(t).$$

Для $c \in (a, t_0)$, где $t_0 = \sup\{t \mid t \in I_1\}$, положим

$$\varphi^*(t; c) = \int_c^b \varphi_n(t; x)\, d\rho_1(x) + \sum_{i=0}^{n} \frac{(D_{i-1} \ldots D_0\bar{\varphi})(c)}{w_i(c)} \varphi_i(t; c). \tag{4.5}$$

Неоднократно повторенное интегрирование по частям, с использованием того, что

$$d\rho_1(t) = d\left[\frac{(D^R_{n-1} \ldots D_0\bar{\varphi})(t)}{w_n(t)}\right]$$

для $a < t < t_0$, приводит к равенству

$$\int_c^d \varphi_n(t; x)\, d\rho_1(x) = \sum_{i=0}^{n} \frac{(D_{i-1} \ldots D_0 \bar{\varphi})(d)}{w_i(d)} \varphi_i(t; d) - $$

$$- \sum_{i=0}^{n} \frac{(D_{i-1} \ldots D_0 \bar{\varphi})(c)}{w_i(c)} \varphi_i(t; c) \qquad (4.6)$$

для $a < c < d < t_0$. Равенства (4.5) и (4.6) означают, что $\varphi^*(t; c) = \varphi^*(t; d)$ для $a < c < d < t_0$, поэтому функция

$$\varphi_1(t) = \varphi^*(t; c), \qquad a < c < t < b \qquad (4.7)$$

полностью определена на $(a, b)$.

Так как функция $\bar{\varphi}$ имеет представление вида

$$\bar{\varphi}(t) = \int_c^b \varphi_n(t; x)\, d\rho(x) + \sum_{i=0}^{n} \frac{(D_{i-1} \ldots D_0 \bar{\varphi})(c)}{w_i(c)} \varphi_i(t; c), \qquad c < t < b,$$

то $\bar{\varphi}(t) = \varphi_1(t) + \varphi_2(t)$, где

$$\varphi_2(t) = \int_a^b \varphi_n(t; x)\, d\rho_2(x). \qquad (4.8)$$

Очевидно, $\varphi_j \in C(u_0, \ldots, u_n)$, $j = 1, 2$. Кроме того, $\beta\varphi_1 \neq \alpha\varphi_2 + u$ для любого многочлена $u$ и вещественных $\alpha$ и $\beta$, так как

$$d\left[\frac{D_{n-1}^R \ldots D_0 \varphi_j}{w_n(t)}\right] = d\rho_j(t), \qquad j = 1, 2.$$

Итак, показано, что любая функция $\varphi$, для которой соответствующая мера $d\rho$ имеет носитель, содержащий по крайней мере две точки, не может быть крайним лучом. На этом доказательство теоремы заканчивается.

Замечание 4.2. Дифференцирование функции $\varphi_n(t; x)$ (см. (2.5)) приводит к соотношению

$$d\left[\frac{D_{n-1}^R \ldots D_0 \varphi_n(t; x)}{w_n(t)}\right] = d\rho_x(t). \qquad (4.9)$$

Кроме того, функция $\varphi_n(t; x)$ однозначно определяется начальными условиями

$$(D_i \ldots D_0 \varphi)(x) = 0, \qquad i = 0, 1, \ldots, n-2, \qquad (4.10)$$

в совокупности с дифференциальным соотношением

$$\frac{(D_{n-1}^R \ldots D_0 \varphi)(t)}{w_n(t)} = \begin{cases} 1, & t \geqslant x, \\ 0, & t < x. \end{cases} \qquad (4.11)$$

Теорема 4.1 утверждает, что функция $\varphi_n(t; x)$ лежит на крайнем луче конуса $C(u_0, u_1, \ldots, u_n)$. Кроме того, из леммы 2.2 и теоремы 2.2 известно, что любая функция $\varphi \in C(u_0, \ldots, u_n)$, удовлетворяющая условию $\rho(a+) > -\infty$, представима в виде выражения от крайних функций $\varphi_n(t; x)$, а именно

$$\varphi(t) = \int_a^b \varphi_n(t; x)\, d\rho(x) + \sum_{i=0}^{n} \frac{(D_{i-1} \ldots D_0\varphi)(a+)}{w_i(a)} u_i(t). \tag{4.12}$$

В общем случае, когда возможно $\rho(a+) = -\infty$, имеем следующее представление:

$$\varphi(t) = \int_c^b \varphi_n(t; x)\, d\rho(x) + \sum_{i=0}^{n} \frac{(D_{i-1} \ldots D_0\varphi)(c)}{w_i(c)} \varphi_i(t; c), \tag{4.13}$$

которое применимо только для $t \in [c, b]$. Так как

$$d\left[\frac{D_{n-1} \ldots D_0\varphi_i(t; c)}{w_n(t)}\right] \equiv 0, \qquad i = 0, 1, \ldots, n, \ c < t < b,$$

то по лемме 2.3 второй член в правой части выражения (4.13) является многочленом $\sum_{i=0}^{n} a_i(c) u_i(t)$ на $(c, b)$. Следовательно, для $t \in (c, b)$ выражение (4.13) может быть переписано в виде

$$\varphi(t) = \int_a^b \varphi_n(t; x)\, d\rho(x) + \sum_{i=0}^{n} a_i(c) u_i(t). \tag{4.14}$$

Перейдем теперь к задаче характеризации крайних лучей конуса $\bigcap_{i=k}^{n} C(u_0, \ldots, u_i)$, где $0 \leq k \leq n$. В этом случае крайние лучи определяются по модулю линейного пространства, натянутого на функции $u_0, u_1, \ldots, u_k$.

Для дальнейшего будет необходима следующая вспомогательная лемма.

Лемма 4.1. *Пусть* $u(t) = \sum_{i=0}^{n} b_i u_i(t)$. *Тогда*

$$u \in \bigcap_{i=k}^{n} C(u_0, \ldots, u_i)$$

*в том и только том случае, когда* $b_i \geq 0$, $i = k+1, k+2, \ldots, n$. *Далее, если* $C^+$ *обозначает конус положительных функций на*

$(a, b)$, *то*

$$u \in C^{+}\left[\bigcap_{j=0}^{n} C(u_0, \ldots, u_j)\right]$$

*в том и только в том случае, когда* $b_i \geqslant 0$, $i = 0, 1, \ldots, n$.

Доказательство. Пусть $u$ принадлежит $\bigcap_{i=k}^{n} C(u_0, \ldots, u_i)$. Предположим противное, т. е. что $b_j < 0$ для некоторого $j \geqslant k+1$. Тогда

$$(D_{j-1} \ldots D_1 D_0 u)(t) = b_j w_j(t) + \sum_{i=j+1}^{n} b_i (D_{j-1} \ldots D_1 D_0 u_i)(t) \quad (4.15)$$

Вычисляя (4.15) в граничной точке $a$, получим

$$(D_{j-1} \ldots D_1 D_0 u)(a) = b_j w_j(a).$$

Следовательно, $(D_{j-1} \ldots D_0 u)(t) > 0$ в некоторой окрестности $a$ и $u \notin C(u_0, \ldots, u_{j-1})$. Так как $j-1 \geqslant k$, то $u \notin \bigcap_{i=k}^{n} C(u_0, \ldots, u_i)$, и этот вывод противоречит принятой гипотезе.

Второе утверждение теоремы следует из первого в сочетании с тем фактом, что если $u \in C^{+}$, то $u(a) = b_0 u_0(a) \geqslant 0$ и, следовательно, $b_0 \geqslant 0$.

Теорема 4.2. а) *Крайние лучи конуса* $\bigcap_{j=k}^{n} C(u_0, u_1, \ldots, u_j)$ *по модулю линейного пространства многочленов, натянутого на* $\{u_0, u_1, \ldots, u_k\}$, *образуются функциями* $u_{k+1}, \ldots, u_n$ *и решением уравнения*

$$d\left[\frac{(D_{n-1}^{R} \ldots D_1 D_0 \varphi(t)}{w_n(t)}\right] = d\rho_x(t), \qquad a < t < b,\ a < x < b, \quad (4.16)$$

*при граничных условиях*

$$(D_i \ldots D_1 D_0 \varphi)(a+) = 0, \qquad i = k, k+1, \ldots, n-1. \quad (4.17)$$

b) *Пусть* $C^{+}$ *обозначает конус функций, которые положительны на* $(a, b)$. *Тогда крайними лучами конуса* $C^{+} \cap \left[\bigcap_{j=0}^{n} C(u_0, \ldots, u_j)\right]$ *являются функции* $\varphi_n(t; x)$, *определенные посредством* (2.5) (*или* (4.16) *и* (4.17) *при* $k = 0$), *вместе с функциями* $u_0, u_1, \ldots, u_n$.

Замечание 4.3. Отметим, что в части b) теоремы 4.2 крайние лучи не выбираются по модулю какого-либо линейного пространства.

Доказательство. Прежде всего покажем, что если функция $\varphi$ удовлетворяет (4.16) и граничным условиям (4.17), то $\varphi$ определяет крайний луч по модулю многочленов вида $\sum_0^k a_i u_i$.

Предположим

$$\varphi = \varphi_1 + \varphi_2, \tag{4.18}$$

где функции $\varphi_1$ и $\varphi_2$ принадлежат конусу $\bigcap_{j=k}^n C(u_0, \ldots, u_j)$. В этом случае теорема 4.1 утверждает, что

$$\varphi_1 = \alpha\varphi_2 + \sum_0^n a_i u_i. \tag{4.19}$$

Теперь, так как $\varphi_1$ и $\varphi_2$ — элементы конуса $\bigcap_{j=k}^n C(u_0, \ldots, u_j)$, из теоремы 3.2 следует, что

$$\begin{aligned} &D_i D_{i-1} \ldots D_0 \varphi_1 \geqslant 0, \\ &D_i D_{i-1} \ldots D_0 \varphi_2 \geqslant 0, \qquad i = k, k+1, \ldots, n-1. \end{aligned} \tag{4.20}$$

Следовательно, так как функция $\varphi$ удовлетворяет гриничным условиям (4.17), заключаем с помощью (4.18), что (4.17) также справедливо и для функций $\varphi_1$ и $\varphi_2$.

Применяя теперь оператор $D_{n-1} \ldots D_0$ при $t = a$ к (4.19), получим $0 = 0 + a_n w_n(a)$. Так как $w_n(a) > 0$, то имеем $a_n = 0$. Продолжая действовать таким же образом с операторами более низкого порядка, получаем, что $a_i = 0$, $i = k+1$, $k+2, \ldots, n$.

Следовательно, если функция $\varphi$ удовлетворяет (4.16) и (4.17), то $\varphi$ определяет крайний луч по модулю линейной комбинации $u_0, u_1, \ldots, u_k$.

Теперь покажем, что любая функция $u_i$, $i = k+1, \ldots, n$, определяет крайний луч в том же смысле. Действительно, предположим противное, т. е. что для некоторого $i_0 \geqslant k+1$

$$u_{i_0} = \varphi_1 + \varphi_2, \tag{4.21}$$

где функции $\varphi_1$ и $\varphi_2$ принадлежат конусу $\bigcap_{j=k}^n C(u_0, \ldots, u_j)$. Применяя оператор $d[(D_{n-1}^R \ldots D_0)/w_n]$ к обеим частям равенства (4.21), приходим к соотношениям

$$d\left[\frac{(D_{n-1}^R \ldots D_0\varphi_i)(t)}{w_n(t)}\right] \equiv 0, \qquad i = 1, 2.$$

Далее из леммы 2.3 заключаем, что

$$\varphi_1(t) = \sum_{j=0}^{n} a_j u_j(t),$$
$$\varphi_2(t) = \sum_{j=0}^{n} b_j u_j(t). \tag{4.22}$$

Затем по лемме 4.1 получаем, что $a_j$ и $b_j$ неотрицательны для $j = k+1, \ldots, n$.

Объединяя (4.21) и (4.22) и используя линейную независимость $u_0, \ldots, u_n$, находим, что

$$a_j + b_j = 0, \qquad j = 0, 1, \ldots, n, \ j \neq i_0,$$
$$a_{i_0} + b_{i_0} = 1.$$

Теперь, так как $a_j \geqslant 0$ и $b_j \geqslant 0$, $j = k+1, \ldots, n$, следует, что $a_j = b_j = 0$, $j = k+1, \ldots, n$, $j \neq i_0$, и $b_{i_0} = 1 - a_{i_0}$. Следовательно,

$$\varphi_1 = a_{i_0} u_{i_0} + \sum_{i=0}^{k} a_i u_i, \qquad 0 \leq a_{i_0} \leq 1,$$

$$\varphi_2 = (1 - a_{i_0}) u_{i_0} - \sum_{i=0}^{k} a_i u_i,$$

что доказывает, что $u_i$ — крайний луч по модулю линейной комбинации функций $u_0, u_1, \ldots, u_k$.

Для доказательства обратного утверждения покажем, что любая функция

$$\bar{\varphi} \in \bigcap_{j=k}^{n} C(u_0, u_1, \ldots, u_j),$$

которая не равна $u_{k+1}, u_{k+2}, \ldots, u_n$ или не удовлетворяет уравнению (4.16) и условиям (4.17), не является крайней. Пусть

$$d\rho(t) = d\left[\frac{(D_{n-1}^{R} \ldots D_0 \bar{\varphi})(t)}{w_n(t)}\right] \geqslant 0.$$

Функция $\bar{\varphi}$ указанного типа должна удовлетворять одному из следующих трех условий:

1) $d\rho(t) \equiv 0$,

2) $d\rho(t)$ имеет носитель по крайней мере с двумя точками,

3) $d\rho(t) = d\rho_x(t)$ для некоторого $x \in (a, b)$ и $(D_{i_0} D_{i_0-1} \ldots D_0 \bar{\varphi}) \times \times (a+) > 0$ для некоторого $i_0 = k, k+1, \ldots, n-1$.

В случае (1) $\bar{\varphi}(t)=\sum_{i=0}^{n} a_i u_i(t)$, где $a_i \geqslant 0$, $i=k+1, \ldots, n$. Очевидно, $\bar{\varphi}(t)$ равна одному из многочленов $u_{k+1}, \ldots, u_n$ по модулю линейной комбинации $\sum_{i=0}^{k} b_i u_i$, или она не является крайней. В случае 2) мы продолжаем, как в доказательстве теоремы **4.1**, и записываем

$$\bar{\varphi}=\varphi_1+\varphi_2, \tag{4.23}$$

где функции $\varphi_1$ и $\varphi_2$ определены в (4.7) и (4.8).

Прямым дифференцированием проверяется, что функции $\varphi_1$ и $\varphi_2$ удовлетворяют условию

$$D_i \ldots D_0 \varphi \geqslant 0, \qquad i=k,\ k+1, \ldots, n-1,$$

так что из теоремы 3.2 получаем, что

$$\varphi_1, \varphi_2 \in \bigcap_{j=k}^{n} C(u_0, \ldots, u_j).$$

Теперь, если $d\rho(t)=d\rho_x(t)$, т. е. случай 3) выполняется, то

$$\bar{\varphi}(t)=\varphi_n(t; x)+\sum_{i=0}^{n} a_i u_i(t). \tag{4.24}$$

Если, кроме того, $(D_{i_0} \ldots D_0 \bar{\varphi})(a+)>0$ для некоторого $i_0=k$, $k+1, \ldots, n-1$, то из (4.24) следует

$$0<(D_{i_0} \ldots D_0 \bar{\varphi})(a+)=a_{i_0+1} W_{i_0+1}(a),$$

поэтому $a_{i_0+1}>0$. В этой ситуации мы записываем $\bar{\varphi}=\varphi_1+\varphi_2$, где

$$\varphi_i=\frac{1}{2} \varphi_n(t; x)+\frac{1}{2} \sum_{j \neq i_0+1} a_j u_j(t)+\alpha_i a_{i_0+1} u_{i_0+1}(t), \qquad i=1,2,$$

где $\alpha_1, \alpha_2>0$, $\alpha_1+\alpha_2=1$ и $\alpha_1 \neq 1/2$.

Выше было показано, что крайние лучи конуса $\bigcap_{j=k}^{n} C(u_0, u_1, \ldots$ $\ldots, u_j)$ по модулю линейного пространства многочленов $\sum_{0}^{k} a_i u_i$ образованы функциями $u_{k+1}, \ldots, u_n$ и решениями уравнения (4.16) при граничных условиях (4.17).

Для того чтобы доказать b), прежде всего заметим, что, так как

$$C^{+} \cap\left[\bigcap_{j=0}^{n} C(u_0, \ldots, u_j)\right] \subset \bigcap_{j=0}^{n} C(u_0, \ldots, u_j),$$

то функции

$$u_1, \ldots, u_n, \qquad \varphi_n(t;\, x) \tag{4.25}$$

определяют крайние лучи конуса $C^+ \cap \left[\bigcap_{j=0}^{n} C(u_0, \ldots, u_n)\right]$ по модулю, кратному $u_0$. Пусть $\varphi$ обозначает одну из функций в (4.25), и предположим, что

$$\varphi = \varphi_1 + \varphi_2,$$

где функции $\varphi_1$ и $\varphi_2$ содержатся в $C^+ \cap \left[\bigcap_{j=0}^{n} C(u_0, \ldots, u_j)\right]$. Очевидно, $\varphi_1 = \alpha\varphi_2 + \beta u_0$. Так как $\varphi_1$ и $\varphi_2$ содержатся в $C^+$ и так как $\varphi(a) = 0$, то получаем, что $\varphi_1(a) = \varphi_2(a) = 0$, поэтому $\beta u_0(a) = 0$ или $\beta = 0$. Таким образом, $u_1, \ldots, u_n$ и $\varphi_n(t;\, x)$ действительно определяют крайние лучи $C^+ \cap \left[\bigcap_{j=0}^{n} C(u_0, \ldots, u_j)\right]$. Тот факт, что $u_0$ также определяет крайний луч, и результат, согласно которому $u_0, \ldots, u_n$ и $\varphi_n(t;\, x)$ являются единственными функциями, определяющими крайние лучи, устанавливаются рассуждением того же типа, что и использовавшееся в части (а).

## § 5. Конус, дуальный к конусу $C(u_0, \ldots, u_n)$

Содержание настоящего параграфа связано с работами Карлина и Новикова [1963] и Циглера [1966]. Некоторые из рассматриваемых здесь вопросов в неявном виде встречаются у Поповичиу. Частные случаи описания конуса, дуального к конусу выпуклых функций, содержатся у Харди, Литтлвуда и Полиа [1952], а также у Левина и Стечкина [1948]. Соотношение упорядоченности, накладываемое дуальным конусом на конус выпуклых функций, оказывается также полезным в теории сравнения экспериментов, см., например, Блекуэлл [1950].

В данном параграфе $d\mu$ будет обозначать обобщенную меру на интервале $[a, b]$ такую, что

$$\int_a^b |d\mu| < \infty. \tag{5.1}$$

Каждая такая мера может быть представлена как $d\mu = d\mu_1 - d\mu_2$, где $d\mu_1$ и $d\mu_2$ есть конечные неотрицательные меры на $(a, b)$. Введем дополнительное ограничение для мер $d\mu$ такое, что для каждой функции $\varphi \in C(u_0, u_1, \ldots, u_n)$ интеграл $\int_a^b \varphi\, d\mu$ определен (при этом допускаются бесконечные значения). В частности, если

$\varphi^+(t) = \max(\varphi(t), 0)$ и $\bar{\varphi}(t) = \varphi^+(t) - \varphi(t)$, то можно записать

$$\int_a^b \varphi\, d\mu = \int_a^b \varphi^+ d\mu_1 + \int_a^b \varphi^- d\mu_2 - \left[\int_a^b \varphi^- d\mu_1 + \int_a^b \varphi^+ d\mu_2\right].$$

То, что интеграл $\int_a^b \varphi\, d\mu$ определен, означает, что по крайней мере одно из выражений

$$\int_a^b \varphi^+ d\mu_1 + \int_a^b \varphi^- d\mu_2, \quad \int_a^b \varphi^- d\mu_1 + \int_a^b \varphi^+ d\mu_2$$

является конечным.

Определение 5.1. *Конус, дуальный к конусу* $C(u_0, u_1, \ldots, u_n)$, *обозначаемый через* $C^*(u_0, u_1, \ldots, u_n)$, *представляет собой множество обобщенных мер* $d\mu$ *на интервале* $(a, b)$, *удовлетворяющих указанным выше требованиям интегрируемости, а также слгдующему соотношению*:

$$\int_a^b \varphi(t)\, \mu(dt) \geqslant 0 \tag{5.2}$$

для всех $\varphi \in C(u_0, u_1, \ldots, u_n)$. Перейдем к описанию свойств конуса $C^*(u_0, u_1, \ldots, u_n)$. Различные случаи применения этих свойств приводятся в §§ 6, 7.

Так как $u_i$, $i = 0, 1, \ldots, n$, ограничены на интервале $[a, b]$ и $\int |d\mu| < \infty$, то сразу же можно сделать вывод, что $\int |u_i|\, |d\mu| < \infty$. Следовательно, так как $\pm u_i \in C(u_0, \ldots, u_n)$ для $i = 0, 1, \ldots, n$, очевидно, что для того, чтобы $d\mu$ принадлежала конусу $C^*(u_0, \ldots, u_n)$, необходимо, чтобы выполнялись следующие условия на моменты:

$$\int_a^b u_i d\mu = 0, \qquad i = 0, 1, \ldots, n. \tag{5.3}$$

Кроме того, если $d\mu \in C^*(u_0, \ldots, u_n)$, то должны выполняться также и условия

$$\int_a^b \varphi_n(t; x)\, d\mu(t) \geqslant 0, \qquad a < x < b. \tag{5.4}$$

Теперь докажем, что условия (5.3) и (5.4) являются достаточными для того, чтобы $d\mu \in C^*(u_0, u_1, \ldots, u_n)$. Для любой функции $\varphi \in C(u_0, u_1, \ldots, u_n)$ построим функцию $\varphi(t; \varepsilon) \in C(u_0, u_1, \ldots, u_n)$, как в теореме 2.3, удовлетворяющую следующим условиям:

$$\varphi(t; \varepsilon) \begin{cases} = \varphi(t), & t \in (a + \varepsilon,\ b - \varepsilon), \\ \leq \varphi(t), & t \in [b - \varepsilon, b), \end{cases} \tag{5.5}$$

$$(-1)^{n+1}[\varphi(t) - \varphi(t; \varepsilon)] \geqslant 0, \qquad t \in (a, a + \varepsilon]. \tag{5.6}$$

Функция $\varphi(t;\varepsilon)$ в явном виде выражается следующим образом:

$$\varphi(t;\varepsilon)=\int_a^b \varphi_n(t;x)\,d\rho(x;\varepsilon)+\sum_0^n a_i(\varepsilon)\,u_i(t)$$

(см. (2.30)), где $\rho(x;\varepsilon)$ является функцией с ограниченной вариацией на интервале $(a, b)$. Так как функция $\varphi_n(t;x)$ ограничена на $[a, b]$, то $\varphi(t;\varepsilon)$ также ограничена, поэтому $\int_a^b |\varphi(t;\ \varepsilon)|\,|d\mu| \leq \infty$.

После обращения к выражениям (5.3) и (5.4) получим, что

$$\int_a^b \varphi(t;\varepsilon)\,d\mu(t)=\int_a^b\left(\int_a^b \varphi_n(t;x)\,d\mu(t)\right)d\rho(x;\varepsilon)\geqslant 0. \qquad (5.7)$$

Теперь покажем, что

$$\lim_{\varepsilon\to 0}\int_a^b \varphi(t;\varepsilon)\,d\mu(t)=\int_a^b \varphi(t)\,d\mu(t), \qquad (5.8)$$

причем здесь допускаются и значения, равные бесконечности. Для этого достаточно убедиться в том, что каждый из интегралов

$$\int_a^b \varphi^+ d\mu_1,\quad \int_a^b \varphi^- d\mu_2,\quad \int_a^b \varphi^- d\mu_1,\quad \int_a^b \varphi^+ d\mu_2$$

является пределом для соответствующих интегралов, где функция $\varphi(t)$ заменена на функцию $\varphi(t;\varepsilon)$. Эти предельные соотношения очевидны. Например, соотношение

$$\lim_{\varepsilon\to 0}\int_a^b \varphi^+(t;\varepsilon)\,d\mu_1(t)=\int_a^b \varphi^+(t)\,d\mu_1(t) \qquad (5.9)$$

является следствием условий (5.5) и (5.6) и теоремы о монотонной сходимости.

Учитывая (5.8) и (5.7), получаем неравенство: $\int_a^b \varphi\,d\mu\geqslant 0$.

На основе предшествующего анализа можно сформулировать следующую теорему.

Теорема 5.1. *Обобщенная мера* $d\mu$ *содержится в дуальном конусе* $C^*(u_0, u_1, \ldots, u_n)$ *тогда и только тогда, когда*

$$\int_a^b u_j\,d\mu=0,\qquad j=0, 1, \ldots, n, \qquad (5.10)$$

$$\int_a^b \varphi_n(t;x)\,d\mu(t)\geqslant 0,\qquad a<x<b. \qquad (5.11)$$

Теорема 5.2. *Обобщенная мера* $d\mu$ *содержится в конусе, дуальном по отношению к конусу* $\bigcap_{j=k}^{n} C(u_0, \ldots, u_j)$, $k \leq n-1$, *тогда и только тогда, когда*

$$1)\ \int_a^b u_j d\mu = 0, \qquad j = 0, 1, \ldots, k;$$

$$2)\ \int_a^b u_j d\mu \geqslant 0, \qquad j = k+1, k+2, \ldots, n; \tag{5.12}$$

$$3)\ \int_a^b \varphi_n(t; x)\, d\mu(t) \geqslant 0, \qquad a < x < b.$$

Доказательство. Условия $\int_a^b u_j d\mu = 0$, $j=0, 1, \ldots, k$, являются необходимыми, так как $\pm u_j \in \bigcap_{i=k}^{n} C(u_0, \ldots, u_i)$ для $j=0,1,\ldots,k$. Кроме того, согласно следствию 3.2 функции $u_{k+1}, u_{k+2}, \ldots, u_n$ и $\varphi_n(t; x)$ принадлежат конусу $\bigcap_{i=k}^{n} C(u_0, \ldots, u_i)$, поэтому условия 2) и 3) из (5.12) являются также необходимыми.

В соответствии с замечанием 3.1 каждая функция $\varphi$, принадлежащая $\bigcap_{i=k}^{n} C(u_0, \ldots, u_i)$, обладает следующим представлением:

$$\varphi(t) = \int_a^b \varphi_n(t; x)\, d\rho(x) + \sum_{i=0}^{n} \frac{(D_{i-1} \ldots D_1 D_0 \varphi)(a+)}{w_i(a)} u_i(t).$$

Следовательно,

$$\int_a^b \varphi(t)\, d\mu(t) =$$

$$= \int_a^b \int_a^b \varphi_n(t; x)\, d\rho(x)\, d\mu(t) + \sum_{i=k+1}^{n} \frac{(D_{i-1} \ldots D_0 \varphi)(a+)}{w_i(a)} \int_a^b u_i d\mu. \tag{5.13}$$

В соответствии с теоремой 3.2 коэффициенты при $\int_a^b u_i d\mu$, $i = k+1, \ldots, n$, неотрицательны. Поэтому на основании 2) и 3) в (5.12) можно заключить, что $\int_a^b \varphi\, d\mu \geqslant 0$.

Следствие 5.2. *Если $C^+$ обозначает конус положительных функций на интервале $(a, b)$, то необходимыми и достаточными условиями, для того чтобы $d\mu$ содержалась в конусе, дуальном к конусу $C^+ \cap \left[\bigcap_{i=0}^{n} C(u_0, \ldots, u_i)\right]$, являются условия*

$$\int_a^b u_i d\mu \geqslant 0, \qquad i = 0, 1, \ldots, n,$$

и

$$\int_a^b \varphi_n(t; x)\, d\mu(t) \geqslant 0, \qquad a < x < b.$$

Можно охарактеризовать дуальные конусы и для других типов пересечений. В большинстве случаев необходимые и достаточные условия, накладываемые на обобщенную меру $d\mu$, не так просты, как указано в теоремах 5.1 и 5.2. Рассмотрим один из интересных случаев.

Пусть $C(u_0)^-$ обозначает конус функций $\varphi$ таких, что $-\varphi \in C(u_0)$.

Теорема 5.3. *Обобщенная мера $d\mu$ принадлежит конусу, дуальному к конусу $C(u_0)^- \cap C(u_0, u_1)$ тогда и только тогда, когда*

$$\int_a^b u_0 d\mu = 0, \tag{5.14}$$

$$\int_a^b \tilde{\varphi}_1(t; x)\, d\mu(t) \geqslant 0, \tag{5.15}$$

*где (см. замечание 2.4)*

$$\tilde{\varphi}_1(t; x) = \begin{cases} w_0(t) \int_t^x w_1(\xi)\, d\xi, & t < x, \\ 0, & t \geqslant x. \end{cases}$$

Доказательство. Так как $\pm u_0 \in C(u_0)^- \cap C(u_0, u_1)$, то условие (5.14) является, очевидно, необходимым. Далее

$$-D_0\tilde{\varphi}_1(t; x) = \begin{cases} w_1(t), & t < x, \\ 0 & t \geqslant x. \end{cases}$$

поэтому $-\tilde{\varphi}_1(t; x)/u_0(t)$ — неубывающая функция по $t$. Следовательно, $\tilde{\varphi}_1(t; x) \in C(u_0)^-$. Кроме того, $D_0\tilde{\varphi}_1(t; x)/w_1$ — непрерывная справа и неубывающая функция, так что $\tilde{\varphi}_1(t; x) \in C(u_0, u_1)$. Таким образом условие (5.15) также является необходимым.

Докажем обратное. Если $\varphi \in C(u_0) \cap C(u_0, u_1)$, то функция $\rho(t) = (D_0\varphi(t))/w_1(t)$ является неубывающей и неположительной, так что $\rho(b-) < \infty$. В соответствии с замечанием 2.4 функция $\varphi$ может быть представлена следующим образом:

$$\varphi(t) = \int_b^t \tilde{\varphi}_1(t; x)\, d\rho(x) - \frac{D_0\varphi(b-)}{w_1(b)} \tilde{\varphi}_1(t; b) + \frac{\varphi(b-)}{w_0(b)} \tilde{\varphi}_0(t; b).$$

Следовательно, так как $\tilde{\varphi}_0(t; b) = w_0(t) = u_0(t)$ и $(D_0\varphi(b-))/w_1(b) \leqslant 0$ можно заключить, учитывая условия (5.14) и (5.15), что $\int_a^b \varphi\, d\mu \geqslant 0$.

Нас интересует задача получения простых достаточных условий для утверждения, что $d\mu \in C^*(u_0, \ldots, u_n)$. С этой целью необходимо формализовать понятие изменения знака меры.

Определение 5.2. Говорят, что обобщенная мера $d\mu$ имеет *$k$ изменений знака* на интервале $(a, b)$, если этот интервал можно разбить на непересекающиеся последовательные интервалы $J_0, \ldots, J_k$ таким образом, что $d\mu$ на $J_0, \ldots, J_k$ меняет знак и не обращается в нуль. (В случае, когда $d\mu = f(t)\, dt$ для некоторой непрерывной функции $f$, число изменений знака $d\mu$ соответствует числу обычных изменений знака функции $f$.)

Теорема 5.4. *Пусть $d\mu$ такова, что*

$$\int_a^b u_j d\mu = 0, \qquad j = 0, 1, \ldots, n. \tag{5.16}$$

*Если $d\mu$ имеет точно $n+1$ изменений знака на интервале $(a, b)$ и является неотрицательной и не равной нулю мерой на некотором интервале, простирающемся до конечной точки $b$, то $d\mu \in C^*(u_0, \ldots, u_n)$.*

Доказательство. Пусть $J_0, \ldots, J_{n+1}$ является подынтервалами интервала $(a, b)$, для которых $d\mu$ удовлетворяет условиям, сформулированным в определении 5.2. Заметим, что $d\mu$ есть неотрицательная мера на $J_{n+1}$. Определим $t_0, \ldots, t_n$ посредством соотношения $t_i = \sup\{t \mid t \in J_i\}$, $i = 0, 1, \ldots, n$, и пусть

$$\theta(t) = \begin{vmatrix} u_0(t_0) & \cdots & u_0(t_n) & u_0(t) \\ \cdot & \cdot & \cdot & \cdot \\ u_n(t_0) & \cdots & u_n(t_n) & u_n(t) \\ \varphi(t_0) & \cdots & \varphi(t_n) & \varphi(t) \end{vmatrix}, \tag{5.17}$$

где $\varphi \in C(u_0, \ldots, u_n)$.

Раскрыв определитель (5.17), видим, что $\theta(t)$ может быть записан в виде

$$\theta(t) = U\begin{pmatrix} u_0, \ldots, u_n \\ t_0, \ldots, t_n \end{pmatrix} \varphi(t) + \sum_0^n a_i u_i(t).$$

Используя условие ортогональности (5.16) для меры $d\mu$, получим

$$\int_a^b \theta(t)\,d\mu(t) = U\begin{pmatrix} u_0, \dots, u_n \\ t_0, \dots, t_n \end{pmatrix} \int_a^b \varphi(t)\,d\mu(t).$$

Однако из выражения (5.17) следует, что $\theta(t)\,d\mu(t)$ является неотрицательной мерой на всем интервале $(a, b)$, так что $\int_a^b \varphi(t)\,d\mu \geqslant 0$. Следовательно, $d\mu \in C^*(u_0, \dots, u_n)$.

Относительно количества изменений знака каждой меры $d\mu \in C^*(u_0, u_1, \dots, u_n)$ известно следующее.

Теорема 5.5. *Если мера* $d\mu$ *удовлетворяет соотношениям ортогональности*

$$\int_a^b u_j d\mu = 0, \qquad j = 0, 1, \dots, n, \tag{5.18}$$

*то* $d\mu$ *имеет по крайней мере* $n+1$ *перемену знака.*

Доказательство. Предположим, что мера $d\mu$ имеет $p$ перемен знака ($p \leqslant n$). Тогда существует такое разбиение на подынтервалы $J_0, \dots, J_p$, при котором $d\mu$ не обращается в нуль и меняет знак на $J_0, \dots, J_p$. Пусть $t_i = \sup\{t \mid t \in J_i\}$, $i = 0, 1, \dots, p-1$, и определим

$$u(t) = \mu(J_p)\,U\begin{pmatrix} u_0, \dots, u_{p-1}, u_p \\ t_0, \dots, t_{p-1}, t \end{pmatrix}.$$

Многочлен $u(t)$ удовлетворяет условию $u(t)\,d\mu(t) \geqslant 0$. Кроме того, так как носитель меры $d\mu$ не может быть ограничен конечным множеством $\{t_0, \dots, t_{p-1}\}$, то $\int_a^b u(t)\,d\mu(t) > 0$. Однако это неравенство не совместимо со свойствами ортогональности, которыми должна обладать, по предположению, мера $d\mu$. Отсюда следует, что $d\mu$ должна иметь по крайней мере $n+1$ перемену знака.

## § 6. Примеры и приложения теории дуальных конусов

В этом параграфе мы приведем ряд примеров и приложений теории, изложенной в § 5. Примеры в основном взяты у Карлина и Новикова [1963]. Кроме того, мы рассмотрим неравенство Стеффенсена [1947] и неравенства Фавара [1933] и Бервальда [1947]. Мы также включили приложение этой теории, впервые предложенное в работе Ханта [1955]. Соответствующие иллюстративные примеры к теории дуальных конусов можно найти у Брунка [1956], [1964], Фаня и Лорентца [1954], Хефдинга [1956], Сизельски [1957], [1958] и Левина и Стечкина [1948], см. также Беккенбаха и Беллмана [1961].

Пример 6.1. *Неравенство Стеффенсена.* В первом примере, рассматривается конус $C(1)$ (т. е. $n = 0$, $u_0(t) = 1$). Это функции $\varphi$ удовлетворяющие следующему условию:

$$\begin{vmatrix} 1 & 1 \\ \varphi(t_0) & \varphi(t_1) \end{vmatrix} \geqslant 0$$

для $a < t_0 < t_1 < b$, т. е. возрастающие функции. Неравенство Стеффенсена [1947] сформулируем следующим образом: если $0 \leq g(t) \leq 1$ на $[a, b]$ и $\varphi \in C(1)$ на $[a, b]$, то

$$\int_a^{a+c} \varphi(t)\,dt \leq \int_a^b \varphi(t)\,g(t)\,dt \leq \int_{b-c}^b \varphi(t)\,dt, \tag{6.1}$$

где $c = \int_a^b g(t)\,dt$.

Мы докажем только левое неравенство. Пусть $g_c(t)$ является характеристической функцией интервала $[a, a + c]$. Тогда функция $g(t) - g_c(t)$ удовлетворяет за счет выбора $c$ условию моментов (5.16) по отношению к функции $u_0(t) \equiv 1$, отрицательна вблизи $a$ и имеет точно одну смену знака (при $t = a + c$). Таким образом, может быть применена теорема 5.4, откуда и вытекает требуемое неравенство.

Пример 6.2. *Неравенство Фавара.* Пусть $f(x)$ является неотрицательной непрерывной вогнутой функцией на интервале $[a, b]$, не равной тождественно нулю, и пусть $\psi(y)$ является выпуклой функцией на интервале $[0, 2\bar{f}]$, где

$$\bar{f} = (b - a)^{-1} \int_a^b f(x)\,dx.$$

Тогда

$$\frac{1}{2\bar{f}} \int_0^{2\bar{f}} \psi(y)\,dy \geqslant \frac{1}{b-a} \int_a^b \psi[f(x)]\,dx. \tag{6.2}$$

Достаточно установить неравенство (6.2) при допущении, что $f$ принадлежит классу непрерывности $C^1$ и является строго вогнутой; общий случай далее следует при переходе к пределу.

Введем меру $d\mu_f$, определенную для всех непрерывных функций $\theta$ посредством соотношения

$$\frac{1}{b-a} \int_a^b \theta(f(x))\,dx = \int_0^\infty \theta(y)\,d\mu_f(y).$$

В рассматриваемом случае $d\mu_f(y) = (1/(b - a))\,m(y)\,dy$, где $m(y)$ определяется возрастающей функцией $\Sigma\,|f'(x)|^{-1}$ для $y \in [f_{\min}, f_{\max}]$ (и равна нулю в остальных случаях), причем сумма берется самое большее по двум значениям $x$, которые удовлетворяют равенству $f(x) = y$.

Рассмотрим теперь разность

$$\frac{1}{2\bar{f}}\int_0^{2\bar{f}} \psi(y)\,dy - \frac{1}{b-a}\int_a^b \psi[f(x)]\,dx = \int_0^\infty \psi(y)\,[d\nu(y) - d\mu_f(y)],$$

где $\nu(y) = \begin{cases} y/2\bar{f}, & 0 \leqslant y \leqslant 2\bar{f}, \\ 1, & 2\bar{f} \leqslant y. \end{cases}$

Из выбора $\nu(y)$ немедленно следует, что эта разность обращается в нуль, если положить $\psi(y) = u_0(y) \equiv 1$ и $\psi(y) = u_1(y) = y$, так что $u_0$ и $u_1$ оказываются ортогональными к $d\nu(y) - d\mu_f(y)$ на $[0, 2\bar{f}]$. По теореме 5.5 разность $d\nu(y) - d\mu_f(y)$ должна иметь по крайней мере два изменения знака (это означает, что $f_{\max} \leq 2\bar{f}$ — результат, который также был установлен Фаваром). Так как $m(y)$ возрастает, легко убедиться в том, что $d\nu(y) - d\mu_f(y)$ имеет ровно два изменения знака, причем в следующем порядке: $+$, $-$, $+$. По теореме 5.4 следует, что (6.2) выполняется.

С помощью аналогичных рассуждений доказывается, что

$$\frac{1}{2\bar{f} - 2c}\int_0^{2\bar{f}-c} \psi(y)\,dy \geqslant \frac{1}{b-a}\int_a^b \psi(f(x))\,dx \tag{6.3}$$

для любого $c$, удовлетворяющего условию $0 < c \leq f_{\min}$, где $\psi(y)$ — выпуклая функция на интервале $[c, 2\bar{f} - c]$.

Используя идею доказательства неравенства (6.3), можем получить вариант этого неравенства, применимый в случае, когда $f(x)$ — непрерывная, неотрицательная и выпуклая функция на отрезке $[a, b]$. Так, когда $f$ — выпуклая функция, $m(y)$ в выражении $d\mu_f(y) = (1/(b-a))\, m(y)\, dy$ убывает на интервале $[f_{\min}, f_{\max}]$. Пусть $d > f_{\min}$. Разность $d\nu(y) - d\mu_f(y)$ ($d\nu(y) = (2\bar{f} - 2d)^{-1}$ на $[d, 2\bar{f} - d]$ и 0 в остальных случаях) изменяет знак в следующем порядке: $-$, $+$, $-$, когда $y$ пересекает положительную ось. Отсюда, используя теорему 5.4, можно получить следующее неравенство:

$$\frac{1}{2\bar{f} - 2d}\int_d^{2\bar{f}-d} \psi(y)\,dy \leq \frac{1}{b-a}\int_a^b \psi(f(x))\,dx, \tag{6.4}$$

если $\psi(y)$ является выпуклой функцией на интервале $[d, 2\bar{f} - d]$. Заметим, что это неравенство противоположно неравенству (6.3). Следует также подчеркнуть, что $c$ в неравенстве (6.3) и $d$ в неравенстве (6.4) могут быть заданы произвольно при условии только, что $d > f_{\min} > c$. Неравенство (6.2), полученное Фаваром, является частным случаем, когда $c = 0$. Предпочтительнее формулировать это неравенство в общем виде (6.3) или (6.4).

При специальном выборе функции $\psi$ в неравенстве (6.2), как показал Фавар, получается ряд интересных неравенств. Например, если $\psi(y) = y^p$, $p \geqslant 1$, можно вывести, что если $f$ — неотрицательная вогнутая функция на интервале $[a, b]$, то

$$\frac{1}{b-a}\int_a^b [f(x)]^p\, dx \leq \frac{2^p}{p+1}\left[\frac{1}{b-a}\int_a^b f(x)\, dx\right]^p, \qquad p \geqslant 1. \tag{6.5}$$

Извлекая корень степени $p$ из обеих частей неравенства и устремив $p$ к бесконечности, получим

$$\frac{1}{b-a}\int_a^b f(x)\, dx \geqslant \frac{1}{2}\max_{a\leqslant x\leqslant b} f(x).$$

Для получения другого приложения неравенства (6.2) возьмем $\psi(y) = -\log y$. В результате получится следующее неравенство;

$$\exp\left\{\frac{1}{b-a}\int_a^b \mathrm{lo}\ \ f(x)\, dx\right\} \geqslant \frac{2}{e}\cdot\frac{1}{b-a}\int_a^b f(x)\, dx.$$

Эти неравенства, а также приведенное ниже неравенство (6.7) имеют направление, противоположное по сравнению с обычными неравенствами Гельдера или Иенсена.

Пример 6.3. *Неравенство Бервальда.*

Оно представляет собой обобщение неравенства Фавара (пример 6.2) при использовании конуса $C(1, X)$ (где функция $X(x)$ обязательно должна быть возрастающей) вместо $C(1, x)$. Пусть функция $f(x)$ является вогнутой или в более общем случае такой, что $d\mu_f$ возрастает (см. пример 6.2). Пусть $\bar{z}$ — единственный положительный корень уравнения

$$\frac{1}{\bar{z}}\int_0^{\bar{z}} X(y)\, dy = \frac{1}{b-a}\int_a^b X(f(x))\, dx = \int_0^{\infty} X(y)\, d\mu_f(y).$$

Результат Бервальда состоит в том, что

$$\frac{1}{\bar{z}}\int_0^{\bar{z}} \varphi(y)\, dy \geqslant \frac{1}{b-a}\int_a^b \varphi(f(x))\, dx \tag{6.6}$$

для $\varphi \in C(1, X)$. Доказательство аналогично доказательству неравенства (6.2); оно заключается в непосредственном применении теоремы 5.4. Из определения $\bar{z}$ следует, что для соответствующей меры $d\nu(y) - d\mu_f(y)$ ($d\nu(y) = 1/\bar{z}$ для $0 \leq y \leq \bar{z}$ и 0 в остальных случаях) выполнены условия, наложенные на моменты по отношению к функции $X(y)$.

В качестве одного из многочисленных полезных неравенств, которые можно получить из неравенства (6.6), приведем следующий результат. Если $X(y) = y^{\alpha}$, $\varphi(y) = y^{\beta}$, $0 < \alpha < \beta$, и $f$ — неотрицательная вогнутая функция на интервале $[a, b]$, то

$$\left[\frac{\beta+1}{b-a}\int_a^b [f(x)]^{\beta}\,dx\right]^{1/\beta} \leq \left[\frac{\alpha+1}{b-a}\int_a^b [f(x)]^{\alpha}dx\right]^{1/\alpha}. \qquad (6.7)$$

Заметим, что если положить $\beta = p \geqslant 1$ и $\alpha = 1$, то неравенство (6.7) сводится к неравенству (6.5).

Пример 6.4. Следующее непосредственное приложение теоремы 5.4 с использованием конуса $C(1, x)$ взято из работы Левина и Стечкина [1948].

Пусть функция $p(x)$ определена на интервале $(0, 1)$ и удовлетворяет условиям: 1) $p(x)$ — неубывающая функция при $0 \leq x \leq 1/2$ и 2) $p(x) = p(1-x)$. Простая проверка показывает, что функция

$$f(x) = 1 - \frac{p(x)}{\int_0^1 p(x)\,dx}$$

удовлетворяет условиям теоремы 5.4. Следовательно, неравенство $\int_0^1 f(x)\varphi(x)\,dx \geqslant 0$ или эквивалентное ему неравенство

$$\int_0^1 p(x)\varphi(x)\,dx \leq \int_0^1 p(x)\,dx \int_0^1 \varphi(x)\,dx$$

выполняется для любой выпуклой функции $\varphi$.

При определенном выборе функции $p(x)$ можно получить следующие полезные неравенства:

1) если $\varphi$ является выпуклой функцией на интервале $(0, 1)$, то

$$\int_0^1 x(1-x)\varphi(x)\,dx \leq \frac{1}{6}\int_0^1 \varphi(x)\,dx;$$

2) если $\varphi$ является выпуклой функцией на интервале $(0, \pi)$, то

$$\int_0^{\pi} \varphi(x)\sin x\,dx \leq \frac{2}{\pi}\int_0^{\pi} \varphi(x)\,dx.$$

Пример 6.5. Пусть $F$ и $G$ представляют собой функции распределения вероятностей положительных случайных величин с одним и тем же средним значением. Таким образом, имеем

$$\int_0^\infty dF = \int_0^\infty dG = 1, \tag{6.8}$$

$$\int_0^\infty (1 - F(x))\,dx = \int_0^\infty (1 - G(x))\,dx = \mu. \tag{6.9}$$

Предположим далее, что

$$1 - F(x) = e^{-\psi_1(x)}, \quad 1 - G(x) = e^{-\psi_2(x)}, \tag{6.10}$$

где $\psi_2$ принадлежит конусу $C(\psi_1)$ и $\psi_1$ строго возрастает. (Это эквивалентно тому, что $(1/y)\,\psi_2(\psi_1^{-1}(y))$ принадлежит конусу $C(1)$, т. е. возрастает.)

Мы утверждаем, что из выражений (6.8) — (6.10) следует, что

$$\int_0^\infty \varphi\,dF \leq \int_0^\infty \varphi\,dG$$

для любой выпуклой функции $\varphi$. Для простоты предположим, что функция $\varphi$ принадлежит классу $C^1$, так как общий случай можно легко получить путем аппроксимации. Кроме того, учитывая равенство (6.8), можно без потери общности принять $\varphi(0) = 0$. Заметим, что

$$\int_0^\infty \varphi\,dF = \int_0^\infty \varphi'(1 - F)\,dx = \int_0^\infty \varphi' e^{-\psi_1} dx,$$

и, следовательно, требуется установить, что

$$\int_0^\infty \varphi'(e^{-\psi_1} - e^{-\psi_2})\,dx \leq 0$$

для любой возрастающей функции $\varphi'$, т. е. что $e^{-\psi_2} - e^{-\psi_1}$ содержится в дуальном конусе к конусу $C(1)$. Отметим, что

$$\operatorname{sign}(e^{-\psi_2} - e^{-\psi_1}) = \operatorname{sign}(\psi_1(x) - \psi_2(x)) = \operatorname{sign}\left(1 - \frac{\psi_2(\psi_1^{-1}(y))}{y}\right),$$

где $y = \psi_1(x) > 0$ для $x > 0$. Согласно (6.10) эта разность должна по крайней мере один раз менять знак (если не равна тождественно нулю), а тогда применима теорема 5.4.

Пример 6.6. Теперь предположим, что относительно $\psi_1$ и $\psi_2$ делаются допущения, что $\psi_1$ — строго возрастающая функция, а $\psi_2$ — выпуклая функция по отношению к $\psi_1$ т. е. $\psi_2$ принадлежит конусу $C(1, \psi_1)$ (или, что эквивалентно, $\psi_2(\psi_1^{-1}(y))$ — выпуклая по $y$). Таким образом,

$$\begin{vmatrix} 1 & 1 & 1 \\ \psi_1(x_1) & \psi_1(x_2) & \psi_1(x_3) \\ \psi_2(x_1) & \psi_2(x_2) & \psi_2(x_3) \end{vmatrix} \geqslant 0$$

при $x_1 < x_2 < x_3$. Предположим, что $F(0) = G(0) = 0$, так что $\psi_1(0) = \psi_2(0) = 0$, а также выполняются условия (6.8) и (6.9) из примера 6.5. Имеем

$$e^{-\psi_2(x)} - e^{-\psi_1(x)} = e^{-\psi(\xi)} - e^{-\xi}, \qquad \psi_1(x) = \xi,$$

где $\psi(\xi) = \psi_2(\psi_1^{-1}(\xi))$ — по предположению выпуклая функция, а $\psi(0) = 0$. Кроме того, мы знаем, что $\psi(\xi) - \xi$ может иметь самое большее два интервала, на которых она положительна, но так как эта разность обращается в нуль при $\xi = 0$, она должна в действительности быть отрицательной вблизи нуля и в конечном счете положительной (возможность того, что $\psi(\xi) \geqslant \xi$ на интервале $(0, \infty)$, исключается согласно теореме 5.5). Таким образом, можно и в этом случае для выпуклой функции $\varphi$ считать, что

$$\int_0^\infty \varphi\, dF \leqslant \int_0^\infty \varphi\, dG.$$

Пример 6.7 (Хант [1965]). Пусть $Y$ — случайная величина с функцией распределения $F$ такая, что $\int_{-\infty}^{\infty} y\, dF(y) = 0$ и $\int_{|y| \leqslant 1} dF(y) = 1$. Если $X$ — симметричная случайная величина с функцией распределения $G$, удовлетворяющей условию $1 \leqslant \int_{-\infty}^{\infty} |x|\, dG(x) < \infty$, то

$$\int_{-\infty}^{\infty} \varphi(x)\, dG(x) \geqslant \int_{-\infty}^{\infty} \varphi(x)\, dF(x) \tag{6.11}$$

для всех выпуклых функций $\varphi$ на интервале $(-\infty, \infty)$. Для доказательства неравенства (6.11) достаточно показать, что

$$\int_{-\infty}^{\infty} \varphi(x)\, dG(x) \geqslant \int_{-\infty}^{\infty} \varphi(x)\, dG(x; a) \geqslant \int_{-\infty}^{\infty} \varphi(x)\, dF(x), \tag{6.12}$$

где $a = \int_{-\infty}^{\infty} |x|\, dG(x) \geqslant 1$ и

$$G(x; a) = \begin{cases} 0, & x < -a, \\ 1/2, & -a \leqslant x < a, \\ 1, & a \leqslant x. \end{cases}$$

Обобщенная мера

$$d\mu(x) = dG(x; a) - dF(x)$$

удовлетворяет условиям $\int d\mu = 0$, $\int x\, d\mu = 0$ и меняет знак в следующем порядке: $+$, $-$, $+$, если $d\mu \not\equiv 0$. Таким образом, правое неравенство в (6.12) является следствием теоремы 5.4.

Далее, для симметричной случайной величины с функцией распределения $H(x)$ выполняется равенство

$$\int_{-\infty}^{\infty} \varphi(x)\, dH(x) = \int_{0}^{\infty} \frac{\varphi(x) + \varphi(-x)}{2}\, dH_s(x),$$

где $H_s$ — распределение на $[0, \infty)$, полученное отражением всей массы $dH$ сосредоточенной на $(-\infty, 0)$ на интервал $(0, \infty)$. (Заметим, что $\psi(x) = (\varphi(x) + \varphi(-x))/2$ — выпуклая по $x$ функция). Чтобы доказать левое неравенство в (6.12), достаточно показать, что

$$\int_{0}^{\infty} \psi(x)\, dG_s(x) \geqslant \int_{0}^{\infty} \psi(x)\, dG_s(x; a) \tag{6.13}$$

для любой выпуклой функции $\psi$ на интервале $[0, \infty)$. Однако неравенство (6.13) непосредственно вытекает из теоремы 5.4.

Заметим, что неравенство (6.11) включает только функции распределения, так что в качестве основного вероятностного пространства можно выбрать единичный квадрат с обыкновенной лебеговой мерой. Далее мы будем предполагать, что случайные величины определены на соответствующем произведении пространств единичного интервала.

Допустим теперь, что $\varphi$ — функция от $x_1, \ldots, x_n$, удовлетворяющая следующим условиям:

1) $\varphi$ — непрерывная выпуклая функция в $R^n$,

2) $X_1, \ldots, X_n$ и $Y_1, \ldots, Y_n$ — совокупности независимых случайных величин,

3) существуют математические ожидания $EX_i$ и $EY_i$,

4) $E(\varphi(X_i)) \geqslant E(\varphi(Y_i))$, где $i = 1, 2, \ldots, n$, для всех выпуклых функций $\varphi$ на $R^1$.

При указанных допущениях сразу же следует, что

$$E(\varphi(X_1, \ldots, X_n)) \geqslant E(\varphi(Y_1, \ldots, Y_n)). \tag{6.14}$$

Доказательство проиллюстрируем для случая $n = 2$. В этом случае, если $F_i$ и $G_i$ — функции распределения $X_i$ и $Y_i$ $(i = 1, 2)$, то

$$\int dF_1(x_1) \int \varphi(x_1, x_2)\, dF_2(x_2) \geqslant \int dF_1(x_1) \int \varphi(x_1, x_2)\, dG_2(x_2) =$$

$$= \int dG_2(x_2) \int \varphi(x_1, x_2)\, dF_1(x_1) \geqslant \int dG_2(x_2) \int \varphi(x_1, x_2)\, dG_1(x_1).$$

Применяя неравенства (6.11) и (6.14), можно получить следующее обобщение неравенства Хинчина.

Пусть $\{Y_k\}_1^\infty$ представляет собой последовательность независимых случайных величин, удовлетворяющих условиям $EY_k = 0$ и $|Y_k| \leq 1$, и пусть $\{a_k\}_1^\infty$ — последовательность констант. Тогда

$$E\left\{\sup_n \left|\sum_{k=1}^n a_k Y_k\right|^{2q}\right\} \leq 2\left(\sum_k a_k^2\right)^q \cdot 1\cdot 3\cdot 5\cdot \ldots (2q-1) \tag{6.15}$$

для любого положительного целого числа $q$.

Достаточно доказать справедливость выражения (6.15) для $n$, принадлежащих конечному множеству $1, 2, \ldots, N$. Пусть $r_k(t)$, $k = 1, 2, \ldots$, обозначают функции Радемахера, определяемые на интервале $[0, 1]$ следующим образом:

$$r_k(t) = \begin{cases} 1, & t \in \left[\dfrac{2i}{2^k}, \dfrac{2i+1}{2^k}\right], \quad i = 0, 1, \ldots, 2^{k-1}-1, \\ -1 & \text{в остальных случаях.} \end{cases}$$

Согласно неравенству (6.11) имеем

$$E\varphi(r_k) \geqslant E\varphi(Y_k), \qquad k = 1, 2, \ldots$$

Так как неравенство (6.15) справедливо в случае функций Радемахера (см. работы Хинчина [1923] и Пэли и Зигмунда [1930]), то результат (6.15) следует из (6.14).

Пример 6.8 (взят из работы Циглера [1966]). Представляет интерес следующая задача. При каких условиях на последовательности $\{a_i\}$, $\{b_i\}$, $i = 0, 1, \ldots, n$, можно построить функцию $\varphi \in C(u_0, u_1, \ldots, u_n)$, определенную на $[a, b]$ и удовлетворяющую граничным условиям

$$\varphi^{(i-1)}(a) = a_{i-1}, \qquad \varphi^{(i-1)}(b) = b_{i-1}, \qquad i = 1, 2, \ldots, n+1. \tag{6.16}$$

Кроме того, какие необходимы условия для того, чтобы обеспечить однозначное определение функции $\varphi$, обладающей указанными свойствами?

Заметим прежде всего, что задание начальных значений в (6.16) эквивалентно заданию следующих значений:

$$\frac{\varphi(a)}{w_0(a)} = a_0^*, \quad \varphi(b) = \bar{b}_0 = b_0, \tag{6.17}$$

$$\frac{(D_i \ldots D_0\varphi)(a)}{w_{i+1}(a)} = a_{i+1}^*, \quad (D_i \ldots D_0\varphi)(b) = \bar{b}_{i+1}, \qquad i = 0, 1, \ldots, n-1.$$

Один подход к решению этой задачи состоит в следующем.

Пусть множества $\{a_i^*\}$, $\{\bar{b}_i\}$, $i=0, 1, \dots, n$, заданы. Определим

$$c_i^* = \bar{b}_{n-i} - \sum_{j=n-i}^{n} a_j^* D_{n-i-1} \dots D_1 D_0 u_j(b), \qquad i = 0, 1, \dots, n,$$

где $D_{-1}$ — тождественный оператор. Тогда возможны четыре случая.

1) Если $c_0^* = 0$, то решение существует тогда и только тогда, когда $c_i^* = 0$, $i = 1, 2, \dots, n$, и оно единственно.

Если $c_0^* \neq 0$, имеются три случая.

2) Если числа $(c_i^*/c_0^*) w_n(b)$, $i = 1, \dots, n$, заключены внутри выпуклого конуса, описываемого кривой

$$\begin{gathered} f_1(x) = D_{n-2} \dots D_1 D_0 \varphi_n(b; x), \\ f_2(x) = D_{n-3} \dots D_1 D_0 \varphi_n(b; x), \dots, f_{n-1}(x) = D_0 \varphi_n(b; x), \\ f_n(x) = \varphi_n(b; x), \qquad a \leq x \leq b, \end{gathered} \tag{6.18}$$

то существует бесчисленное множество решений (определение функции $\varphi_n(b; x)$ см. в (2.5)).

3) Если числа $(c_i^*/c_0^*) w_n(b)$, $i = 1, \dots, n$, находятся на границе этого конуса, то существует единственное решение.

4) Если эти числа лежат за пределами конуса, то решения нет.

Согласно следствию 2.1 каждая функция $\varphi(t)$ из конуса $C(u_0, u_1, \dots, u_n)$, имеющая требуемые производные при $t = a$, может быть представлена в виде

$$\varphi(t) = \int_a^b \varphi_n(t; x)\, d\rho(x) + \sum_{j=0}^{n} \frac{D_{j-1} \dots D_0 \varphi(a)}{w_j(a)} u_j(t).$$

Применяя оператор $D_{n-i-1} \dots D_1 D_0$, получим

$$\bar{b}_{n-i} = D_{n-i-1} \dots D_1 D_0 \varphi(b) = \int_a^b D_{n-i-1} \dots D_0 \varphi_n(b; x)\, d\rho(x) + \\ + \sum_{j=n-i}^{n} a_j^* D_{n-i-1} \dots D_0 u_j(b), \qquad i = 1, 2, \dots, n,$$

так что

$$c_i^* = \int_a^b D_{n-i-1} \dots D_0 \varphi_n(b; x)\, d\rho(x), \qquad i = 1, 2, \dots, n. \tag{6.19}$$

Далее

$$c_0^* = \bar{b}_n - a_n^* w_n(b) = \\ = \int_a^b D_{n-1} \dots D_1 D_0 \varphi_n(b; x)\, d\rho(x) = w_n(b) \int_a^b d\rho(x). \tag{6.20}$$

Следовательно, если $c_0^* = 0$, то $d\rho(t) \equiv 0$, так как это неотрицательная мера. Таким образом, $\varphi(t)$ должна удовлетворять следующему условию:

$$d\left[\frac{D_{n-1}\dots D_0\varphi(t)}{w_n(t)}\right] \equiv 0,$$

и в этом случае очевидно, что решение существует тогда и только тогда, когда $c_i^* = 0$, $i = 0, 1, \dots, n$. Линейная комбинация $\varphi(t) = \sum_{j=0}^{n} a_j^* u_j(t)$, которая принадлежит конусу $C(u_0, \dots, u_n)$, является тогда единственным решением задачи.

В другом случае, т. е. когда $c_0^* \neq 0$, существование неотрицательной меры $d\sigma(t)$, такой, что

$$\int_a^b d\sigma = 1,$$

$$\int_a^b D_{n-i-1}\dots D_0\varphi_n(b; x)\, d\sigma(x) = \frac{c_i^*}{c_0^*} w_n(b), \qquad i = 1, \dots, n, \tag{6.21}$$

является необходимым и достаточным условием существования решения, в чем можно убедиться, обратившись к равенствам (6.19) и (6.20). Условия (6.21) предполагают, что в нашем распоряжении имеются все три варианта, упомянутые в пунктах 2), 3) и 4).

## § 7. Дискретные примеры

Цель этого параграфа состоит в описании некоторых дискретных аналогов и примеров теории, изложенной в предыдущих параграфах.

В начале рассмотрим конус конечных последовательностей $\{a_0, a_1, \dots, a_n\}$ таких, что

$$\sum_0^n a_k\varphi(k) \geqslant 0 \tag{7.1}$$

для любой выпуклой последовательности $\varphi(k)$, $k = 0, 1, \dots, n$.

Легко видеть, что к этому дискретному аналогу тоже можно отнести предыдущие рассуждения, и что он соответствует подклассу чисто атомных обобщенных мер $d\mu$, которые сосредоточены в целых числах $k = 0, 1, \dots, n$ с соответствующими массами $a_0, \dots, a_n$. Условия (5.10) и (5.11) теоремы 5.1 сводятся к условиям

$$\sum_{0}^{n} a_k = 0, \tag{7.2}$$

$$\sum_{0}^{n} k a_k = 0, \tag{7.3}$$

$$\sum_{j=0}^{k} \sum_{i=0}^{j} a_i \geqslant 0, \qquad k = 0, 1, \ldots, n, \tag{7.4}$$

которые являются необходимыми и достаточными для выполнения неравенства (7.1) для всех выпуклых последовательностей $\{\varphi(k)\}$. Следует отметить, что нет необходимости ограничиваться только конечными рядами; интерпретация ряда $\sum\limits_{0}^{\infty} a_k \varphi(k) \geqslant 0$ для выпуклых функций имеет обычный смысл, если ряд сходится, и значение этого выражения стремится к бесконечности, если ряд расходящийся.

Легко доказать следующие леммы (ср. с теоремами 5.5 и 5.4).

Лемма 7.1. *Если условия* (7.2) *и* (7.3) *выполняются, то последовательность* $\{a_0, \ldots, a_n\}$ *содержит по крайней мере две перемены знака.*

Лемма 7.2. *Если условия* (7.2) *и* (7.3) *выполняются и перемены знака в последовательности* $\{a_0, \ldots, a_n\}$ *происходят следующим образом:*

$$a_k \geqslant 0 \quad \text{для} \quad 0 \leq k \leq k_1,$$

$$a_k \leq 0 \quad \text{для} \quad k_1 + 1 \leq k \leq k_2,$$

$$a_k \geqslant 0 \quad \text{для} \quad k_2 + 1 \leq k \leq n,$$

*причем строгое неравенство имеет место по крайней мере один раз в каждой из указанных трех областей, то неравенство* (7.1) *выполняется для всех выпуклых последовательностей.*

Пример 7.1. Пусть $0 \leq p_i \leq 1$, $i = 1, \ldots, n$, $p_i + q_i = 1$, и определим $\{b_k\}$ посредством равенства $\sum\limits_{0}^{n} b_k x^k = \prod\limits_{i=1}^{n} (q_i + p_i x)$. Далее пусть

$$d_k = \binom{n}{k} p^k (1-p)^{n-k}, \quad \text{т. е.} \quad \sum_{0}^{n} d_k x^k = (q + px)^n,$$

где $p = \frac{1}{n}(p_1 + \ldots + p_n)$ и $p + q = 1$. Тогда

$$\sum_{k=0}^{n} b_k \varphi(k) \leqslant \sum_{k=0}^{n} d_k \varphi(k) \tag{7.5}$$

для любой выпуклой функции $\varphi(k)$. Этот результат можно интерпретировать как неравенство вида

$$E_Y(\varphi) \leq E_X(\varphi) \tag{7.6}$$

для выпуклой функции $\varphi$.($E_X(\varphi)$ обозначает математическое ожидание случайной величины $\varphi(X)$). Здесь $Y$ и $X$ представляют собой случайные величины, обозначающие число благоприятных исходов в $n$ независимых биномиальных испытаниях с соответствующими вероятностями $p_i$ ($i = 1, \ldots, n$) и в $n$ независимых биномиальных испытаниях с вероятностью $p$ в каждом соответственно.

Для доказательства неравенства (7.5) необходимо только показать, что $a_k = d_k - b_k$ удовлетворяют условиям (7.2) — (7.4). В следующем примере устанавливается несколько более сильный результат.

Пример 7.2. Пусть $p_i$ $(0 \leq p_i \leq 1,\ \ i = 1, \ldots, n)$ определены как в примере 7.1. Пусть $\|\pi_{ij}\|_{i,j=1}^{n}$ — бистохастическая матрица $\left(\text{т. е. } \pi_{ij} \geq 0,\ \sum_{i=1}^{n} \pi_{ij} = \sum_{j=1}^{n} \pi_{ij} = 1\right)$ и $p'_i = \sum_{j=1}^{n} \pi_{ij} p_j,\ \ i = 1, \ldots, n$. Пусть далее $b_k$ определены, как в примере 7.1, и $b'_k$ определены подобным же образом посредством $\sum_{0}^{n} b'_k x^k = \prod_{i=1}^{n} (q'_i + p'_i x)$. Если $Y'$ представляет собой случайную величину, обозначающую число успешных исходов в $n$ независимых биномиальных испытаниях с соответствующими вероятностями $p'_i$ $(i = 1, \ldots, n)$, то $E_Y(\varphi) \leq$ $\leq E_{Y'}(\varphi)$ для любой выпуклой функции $\varphi$, или, что эквивалентно,

$$\sum_{k=0}^{n} \varphi(k)(b'_k - b_k) \geq 0 \tag{7.7}$$

для любой выпуклой функции $\varphi$. Результат примера 7.1 получается путем конкретизации матрицы $\pi_{ij}$. Мы утверждаем, что достаточно установить это в случае:

$$p'_1 = tp_1 + (1-t)p_2,$$
$$p'_2 = (1-t)p_1 + tp_2, \qquad 0 < t < 1,$$
$$p'_i = p_i, \qquad i = 3, 4, \ldots, n,$$

где $t$ — указанный параметр.

Если матрица этого преобразования заменяется матрицей, в которой первые две строки и два столбца меняются местами с произвольными строками и столбцами, то требуемый результат все равно имеет место, так как меняется только порядок $\{p_i\}$ и $\{p'_i\}$.

Однако, как показано у Харди, Литтлвуда и Полиа [1952] (гл. I), произвольная бистохастическая матрица представима как произведение матриц такого же типа.

Простые вычисления показывают, что

$$b'_k - b_k = (p'_1 p'_2 - p_1 p_2)\,\Delta^2 R_k,$$

где

$$\Delta^2 R_k = R_{k-2} - 2R_{k-1} + R_k, \quad R_{-1} = R_{-2} = 0,$$

$$\sum_0^{\infty} R_k x^k = \prod_{i=3}^{n} q_i + p_i x).$$

Заметим, что $R_{n-1} = R_n = \ldots = 0$. Отсюда следует, что

$$\sum_{j=0}^{k} \sum_{i=0}^{j} (b'_i - b_i) = (p'_1\, p'_2 - p_1 p_2)\, R_k,$$

что является неотрицательной величиной, так как $R_k$, очевидно, удовлетворяет условию $R_k \geqslant 0$, в то время как

$$p'_1\, p'_2 - p_1 p_2 = t\,(1-t)\,(p_1^2 + p_2^2 - 2p_1 p_2) \geqslant 0.$$

Пример 7.3. Рассмотрим бесконечную последовательность независимых биномиальных испытаний с вероятностями благоприятных исходов $p_1, p_2, p_3, \ldots,\ 0 < p_i < 1$. Предположим, $\sum_{i=1}^{\infty} p_i = \lambda < \infty$. Вероятность $k$ благоприятных исходов $a_k$ можно вычислить с помощью производящей функции $\sum_{k=0}^{\infty} a_k x^k = \prod_{i=1}^{\infty} (q_i + p_i x),\quad q_i + p_i = 1$. Мы утверждаем, что

$$\sum_{k=0}^{\infty} a_k \varphi(k) \leq \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \varphi(k) \tag{7.8}$$

для любой выпуклой последовательности $\varphi(k)$, которая возрастает при $k \to \infty$ не быстрее, чем экспонента. Учитывая, что $\log(k!\, a_k)$ является строго вогнутой функцией (см. Харди, Литтлвуд и Полиа [1952], стр. 52), и что нулевой и первый моменты от $b_k = (\lambda^k/k!)\, e^{-\lambda} - a_k$ обращаются в нуль, заключаем, что последовательность $\{b_k\}_0^{\infty}$ меняет знак ровно два раза. Кроме того, $a_0 = \prod_{i=1}^{\infty} (1 - p_i) < e^{-\Sigma p_i} = e^{-\lambda}$, откуда видно, что $b_0 > 0$. Применение леммы 7.2 приводит к неравенству (7.8). Тот же результат можно получить путем соответствующего предельного перехода из результата примера 7.1.

Результат данного примера может быть записан в виде $E_X(\varphi) \leq \mathbf{E}_Y(\varphi)$, где $X$ — число благоприятных исходов в бесконечной последовательности независимых биномиальных испытаний, а $Y$ — соответствующая пуассоновская случайная величина.

Пример 7.4. Определим $a_k$ с помощью производящей функции:

$$\sum_{k=0}^{\infty} a_k x^k = \prod_{i=1}^{n} \frac{q_i}{1-p_i x},$$

где $p_1, p_2, \ldots, p_n$ заданы так, что $0 < p_i < 1$, $p_i + q_i = 1$. Если $p_1 = p_2 = \ldots = p_n$, то для $a_k$ известны явные выражения.

В частности, имеем

$$\sum_{k=0}^{\infty} \binom{n+k-1}{n-1} q^n p^k x^k = \left(\frac{q}{1-px}\right)^n.$$

Как показано у Харди, Литтлвуда и Полиа [1952] (стр. 164), функция

$$\log\left[\frac{a_k}{\binom{n+k-1}{n-1} q^n p^k}\right]$$

является строго выпуклой, как последовательность по $k$, если не выполняется $p_1 = p_2 = \ldots = p_n$.

В частности, если $p$ определяется равенством

$$\frac{n}{p} = \frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_n}$$

и $p + q = 1$, то условия леммы 7.2 выполняются, и можно сделать вывод, что

$$\sum_{k=0}^{\infty} a_k \varphi(k) \geqslant \sum_{k=0}^{\infty} \binom{n+k-1}{n-1} p^k q^n \varphi(k) \tag{7.9}$$

для любой выпуклой последовательности $\varphi(k)$ (следует обратить внимание на изменение смысла неравенства на противоположный по сравнению с неравенством (7.8)).

Этот результат можно обобщить следующим образом. Если

$$\frac{p_i'}{q_i'} = \sum_{j=1}^{n} \pi_{ij} \frac{p_j}{q_j}, \qquad 0 < p_i' < 1, \quad p_i' + q_i' = 1 \qquad (i = 1, \ldots, n)$$

и $\| \pi_{ij} \|_{i,j=1}^{n}$, как и прежде, — бистохастическая матрица, то

$$\sum_{k=0}^{\infty} \varphi(k) a_k \geqslant \sum_{k=0}^{\infty} \varphi(k) a_k', \tag{7.10}$$

где $\varphi$ — выпуклая функция (не исключается возможность, что обе части неравенства обращаются в бесконечность).

Пример 7.5. Следует указать на связь условий (5.10) и (5.11) для $n = 1$ с известным критерием Карамата (см. Харди, Литтлвуд и Полиа [1952], гл. I): неравенство

$$\sum_{i=1}^{n} \varphi(a_i) \leqslant \sum_{i=1}^{n} \varphi(a'_i) \tag{7.11}$$

имеет место, когда $\varphi$ — выпуклая функция. У Харди, Литтлвуда и Полиа показано, что если $\{a_i\}$ и $\{a'_i\}$ упорядочены так, что $0 \leq a_1 \leq a_2 \leq \ldots \leq a_n$, $0 \leq a'_1 \leq a'_2 \leq \ldots \leq a'_n$, то неравенство (7.11) справедливо тогда и только тогда, когда

$$\sum_{i=1}^{r} a_i \leq \sum_{i=1}^{r} a'_i, \qquad r = 1, 2, \ldots, n-1, \tag{7.12}$$

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} a'_i. \tag{7.13}$$

Пусть $d\mu_1$ ($d\mu_2$) обозначает меру, приписывающую единичную массу каждой из точек $\{a_i\}_{i=1}^{n}$ ($\{a'_i\}$) (если среди значений $a_i$ встречаются равные, то мы приписываем общей точке массу, равную ее кратности).

Введем следующие обозначения. Пусть $\nu_1(x) = \int\limits_0^x d\mu_1$ — число значений $a_i$, которые меньше чем $x$, $\nu_2(x) = \int\limits_0^x d\mu_2$ — число значений $a'_i$, которые меньше чем $x$, для всех $x > 0$. Соотношения (5.10) и (5.11) перепишутся в следующем виде:

$$\int d\mu_1(x) = \int d\mu_2(x) = n, \tag{7.14}$$

$$\int x d\mu_1(\mu) = \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} a'_i = \int x d\mu_2(x), \tag{7.15}$$

$$\int\limits_0^x \nu_1(y)\, dy \geqslant \int\limits_0^x \nu_2(y)\, dy. \tag{7.16}$$

Нетрудно проверить непосредственно эквивалентность неравенств (7.16) и (7.12), что мы и проделаем ниже. (Ввиду этого результат, представленный соотношениями (7.14)—(7.16), может рассматриваться как обобщение теоремы Карамата, относящейся к неравенствам (7.12).) Заметим, что по определению $\nu_1(y) = k$ в интервале $(a_k, a_{k+1}]$ и $\nu_2(y) = k$ в промежутке $(a'_k, a'_{k+1}]$.

Заметим, что

$$\int\limits_0^{a_k} \nu_1(y)\, dy = (a_2 - a_1) + 2(a_3 - a_2) + \ldots + (k-1)(a_k - a_{k-1}) =$$
$$= -a_1 - a_2 - \ldots - a_k + k a_k.$$

С другой стороны, если $l = l(k)$ выбрано так, что $a'_l \leq a_k < a'_{l+1}$, то

$$\int\limits_0^{a_k} \nu_2(y)\,dy = -a'_1 - a'_2 - \ldots - a'_l + la_k.$$

Если теперь предположить, что условие (7.16) выполняется для всех $x$, то оно справедливо, в частности, для $x = a_k$, и мы имеем

$$-\sum_1^k a_i + ka_k \geqslant -\sum_1^l a'_i + la_k, \tag{7.17}$$

или

$$\sum_1^k a_i + la_k \leq \sum_1^l a'_i + ka_k.$$

Из неравенства (7.17) теперь можно вывести неравенство (7.12).

Если $l \leq k$, то $\sum\limits_1^l a'_i + ka_k \leq \sum\limits_1^k a'_i + la_k$, так как $a'_{l+1} \geqslant a_k$, а если $l > k$, то $\sum\limits_1^l a'_i \leq \sum\limits_1^k a'_i + (l-k)\,a_k$. В обоих случаях выполнение (7.17) влечет выполнение (7.12).

С другой стороны, если неравенство (7.17) имеет место для $k = 1, \ldots, n$, то неравенство (7.16) будет справедливо для $x = a_1$, $a_2, \ldots, a_n$. Так как $\int\limits_0^x \nu_1(y)\,dy$ является линейной функцией в интервалах между этими точками, а $\int\limits_0^x \nu_2(y)\,dy$ есть выпуклая функция, то неравенство (7.16) выполняется между этими точками, т. е. оно выполняется для всех $x$. Таким образом, для того чтобы показать, что из неравенства (7.12) следует неравенство (7.16), достаточно показать, что из (7.12) вытекает (7.17), что и завершает доказательство эквивалентности. Предположим поэтому, что $\sum\limits_1^k a_i \leq \sum\limits_1^k a'_i$ для $k = 1, \ldots, n$; если $l \leq k$, то $\sum\limits_1^k a_i \leq \sum\limits_1^l a'_i + (k-l)\,a_k$, а если $l > k$, то имеем $\sum\limits_1^k a_i + (l-k)\,a_k \leq \sum\limits_1^l a'_i$. Таким образом, $\sum\limits_1^k a_i \leq \sum\limits_1^l a'_i + (k-l)\,a_k$ в любом случае, что соответствует неравенству (7.17).

Пример 7.6. Дискретный аналог конуса $C(1)$ представляет собой конус возрастающих последовательностей. Соответствующие необходимые и достаточные условия того, чтобы $\sum_1^n a_k d_k \geqslant 0$ для всех возрастающих последовательностей $\{a_k\}$, состоят в том, что $\sum_1^n d_k = 0$, $\sum_{j=1}^k d_j \leqq 0$, $k = 1, \ldots, n$. Использование этого факта приводит к элементарной теореме Чебышева о перестановках. Пусть $a_1 \leqq a_2 \leqq \ldots \leqq a_n$, и пусть $\sigma$ обозначает перестановку $\{1, 2, \ldots, n\}$. Тогда

$$\sum_1^n a_k d_{\sigma(k)} \leqq \sum_1^n a_k d_{\sigma_0(k)},$$

где $\sigma_0$ — такая перестановка, что

$$d_{\sigma_0(1)} \leqq d_{\sigma_0(2)} \leqq \ldots \leqq d_{\sigma_0(n)}.$$

Пример 7.7. Пусть

$$u_0(x) \equiv 1, \quad u_1(x) = \begin{cases} 0, & x = 0, 1, \ldots, r, \\ 1, & x = r+1, \; r+2, \ldots, n \end{cases} \qquad (r < n).$$

Конус $C(u_0, u_1)$ в этом случае состоит из таких последовательностей $\{\varphi(i)\}$ длины $n$, которые не возрастают на подынтервале $[0, r]$ и не убывают на $[r+1, n]$. Заметим, что $u_0(x) \equiv 1$ и

$$\begin{vmatrix} u_0(x_1) & u_0(x_2) \\ u_1(x_1) & u_1(x_2) \end{vmatrix} \geqslant 0,$$

причем строгое неравенство имеет место, если $x_1 \leqq r$ и $x_2 > r$.

Сформулируем следующий результат (доказательство непосредственно получается при использовании методов аппроксимации и здесь опускается). Пусть $a_0, a_1, a_2, \ldots, a_n$ определяется следующим образом:

$$\sum_{j=0}^n a_j x^j = \prod_{i=1}^n (q_i + p_i x), \quad p_i + q_i = 1, \qquad 0 \leqq p_i \leqq 1.$$

Определим $\tilde{p}$ с помощью условия на моменты

$$\sum_{j=r+1}^n a_j = \sum_{j=r+1}^n \binom{n}{j} \tilde{p}^j (1 - \tilde{p})^{n-j}.$$

Если

$$b_j = a_j - \binom{n}{j} \tilde{p}^j (1 - \tilde{p})^{n-j}, \qquad j = 0, 1, \ldots, n,$$

то

$$\sum_{j=0}^{n} b_j \varphi(j) \leqslant 0$$

для всех последовательностей $\varphi$ из конуса $C(u_0, u_1)$, т. е. для последовательностей $\varphi$, не возрастающих в подынтервале $[0, r]$ и не убывающих в подынтервале $[r+1, n]$.

Пример 7.8. Этот пример, относящийся к свойствам монотонности сумм Римана, рассматривался в работе Сеге и Турана [1961]. Здесь используются методы, отличные от методов, применявшихся указанными авторами.

Пусть $0 = x_0 < x_1 < \ldots < x_n < x_{n+1} = 1$, $0 = y_0 < y_1 < \ldots < y_{n+1} < 1$, и пусть эти числа удовлетворяют условию

$$x_\nu < y_{\nu+1} < x_{\nu+1}, \qquad \nu = 0, \ldots, n. \tag{7.18}$$

Введем обозначения $y_{\nu+1} - x_\nu = p_\nu$, $x_{\nu+1} - y_{\nu+1} = q_\nu$, $\nu = 0, 1, \ldots, n$, $p_{n+1} = 0$, и образуем суммы Римана, соответствующие этим разбиениям,

$$S_n^{(l)}(f) = \sum_{\nu=0}^{n} (x_{\nu+1} - x_\nu) f(x_\nu),$$

$$S_{n+1}^{(l)}(f) = \sum_{\nu=0}^{n+1} (y_{\nu+1} - y_\nu) f(y_\nu) \qquad (y_{n+2} = 1),$$

$$S_{n+1}^{(r)}(f) = \sum_{\nu=0}^{n+1} (y_{\nu+1} - y_\nu) f(y_{\nu+1}),$$

$$S_n^{(r)}(f) = \sum_{\nu=0}^{n} (x_{\nu+1} - x_\nu) f(x_{\nu+1}).$$

Необходимые и достаточные условия того, чтобы соотношения

$$S_n^{(r)}(f) \leq S_{n+1}^{(r)}(f); \quad S_n^{(l)}(f) \geqslant S_{n+1}^{(l)}(f) \tag{7.19}$$

выполнялись для любой убывающей и выпуклой функции $f(x)$, заключаются в том, что должны иметь место неравенства

$$\sum_{\nu=0}^{\mu} q_\nu (p_{\nu+1} - p_\nu) \leq 0, \qquad \mu = 0, 1, \ldots, n. \tag{7.20}$$

Докажем этот результат для $S_n^{(l)}$, при небольшой модификации это доказательство можно применить и для $S_n^{(r)}$. Имеем

$$S_n^{(l)}(f) - S_{n+1}^{(l)}(f) = \sum_{\nu=1}^{2n+2} a_\nu f(z_\nu), \tag{7.21}$$

где $\{z_\nu\}$ — числа $\{x_\nu\}$, $\{y_\nu\}$, упорядоченные по величине. Коэффициенты $a_\nu$ определяются следующим образом:

$$a_1 = q_0, \quad a_{2k} = -(p_k + q_{k-1}), \qquad k = 1, \ldots, n+1,$$
$$a_{2k+1} = q_k + p_k, \qquad k = 1, \ldots, n.$$

Теперь видно, что теорема 5.3 применима при $w_0(t) \equiv w_1(t) \equiv 1$, $a = 0$, $b = 1$. В данной задаче рассматриваются последовательности, а $d\mu(t)$ — дискретная мера со скачками $a_\nu$ при $t = z_\nu$, где $\nu = 1, \ldots, 2n+2$. Необходимые и достаточные условия того, чтобы выражение (7.21) было неотрицательно, сводятся к следующему:

$$\int_0^1 d\mu = 0, \tag{7.22}$$

$$\int_0^1 \tilde{\varphi}_1(t; x)\, d\mu(t) \geqslant 0, \tag{7.23}$$

где $\tilde{\varphi}_1(t; x) = \begin{cases} w_0(t) \int_t^x w_1(\xi)\, d\xi = x - t, & t < x, \\ 0, & t \geqslant x. \end{cases}$

Условие (7.22) немедленно следует из равенств

$$\int_0^1 d\mu = \sum_{i=1}^{2n+2} a_i = q_0 + \sum_{k=1}^{n} [(q_k + p_k) - (p_k + q_{k-1})] - q_n = 0.$$

Функция $\int_0^x d\mu$ принимает значения

$$\int_0^x d\mu = \begin{cases} q_i, & x_i \leq x < y_{i+1}, \quad i = 0, 1, \ldots, n, \\ -p_i, & y_i \leq x < x_i, \quad i = 1, 2, \ldots, n, \\ 0, & y_{n+1} \leq x \leq 1. \end{cases}$$

Несложные алгебраические преобразования показывают, что

$$\int_0^1 \tilde{\varphi}_1(t; x)\, d\mu(t) = \int_0^x (x - t)\, d\mu(t) \geqslant 0, \qquad 0 \leq x \leq 1;$$

тогда и только тогда, когда имеют место неравенства (7.20).

Пример 7.9. Пусть $X(t)$ — неубывающая и непрерывная функция при $\alpha \leq t \leq \beta$, где $[\alpha, \beta)$ — конечный интервал. Если $X(\beta) < \infty$, то продолжим $X(t)$ на интервал $[\beta, \infty)$ таким образом, чтобы $X(t)$ оставалась непрерывной и неубывающей, и чтобы $\lim\limits_{t\to\infty} X(t) = \infty$.

Аналогично, если $X(\alpha) > -\infty$, то продолжим $X(t)$ на $(-\infty, \alpha]$, сохранив свойства непрерывности и неубывания функции $X(t)$, и таким образом, чтобы $\lim\limits_{t\to\infty} X(t) = -\infty$.

Пусть $G(t)$ — функция ограниченной вариации на интервале $[\alpha, \beta]$, для которой $G(\alpha) = 0$. Кроме того, предположим, что $X(t)$ интегрируема относительно меры $|dG(t)|$. Дадим теперь необходимые и достаточные условия того, чтобы неравенство

$$\int_{\alpha}^{\beta} f(X(t))\, dG(t) \geqslant f\left(\int_{\alpha}^{\beta} X(t)\, dG(t)\right) \tag{7.24}$$

имело место для любой выпуклой функции $f$. Допустим, что функция $f$ определена на наименьшем интервале, включающем область изменения функции $X(t)$ $(\alpha \leqslant t \leqslant \beta)$ и значение $m = \int\limits_{\alpha}^{\beta} X(t)\, dG(t)$.

Пусть $\gamma_0$ определена таким образом, что

$$\int_{\alpha}^{\beta} X(t)\, dG(t) = X(\gamma_0).$$

Это возможно благодаря указанным в определении свойствам функции $X(t)$.

Требуемые условия для того, чтобы неравенство (7.24) выполнялось, непосредственно следуют из теоремы 5.1. Таким образом, неравенство (7.24) выполняется, как показано, тогда и только тогда, когда

$$\begin{cases} G(\beta) = 1, & \\ \int\limits_{t}^{\beta} [1 - G(u)]\, dX(u) \geqslant 0, & \beta \geqslant t \geqslant \gamma_0, \\ \int\limits_{\alpha}^{t} G(u)\, dX(u) \geqslant 0, & \gamma_0 \geqslant t \geqslant \alpha. \end{cases} \tag{7.25}$$

Когда $\gamma_0 > \beta$ $(\gamma_0 < \alpha)$, первая (вторая) группа интегральных неравенств отпадает. Доказательство очень простое. Пусть $H$ представляет собой вырожденную меру, сконцентрированную в $\gamma_0$. Тогда неравенство (7.24) можно записать в следующем виде:

$$\int_{\alpha}^{\beta} f(X(t))\, [dG(t) - dH(t)] \geqslant 0.$$

Условия теоремы 5.1 для рассматриваемого случая сводятся к соотношениям (7.25) после соответствующего интегрирования по частям.

Несложное доказательство показывает, что предположение о непрерывности $X(t)$ можно отбросить при сохранении условия, что $X(t)$ — неубывающая функция.

Приведенные рассуждения были предложены Брунком [1964]. Однако его изложение результатов не полностью совпадает с вышеприведенным. Некоторые многомерные аналоги неравенства (7.24) приводятся у Брунка [1964] и Сизельски [1957].

1) Достаточно условие, которое, очевидно, обеспечивает выполнение соотношений (7.25), состоит в том, что $0 \leq G(t) \leq 1$ на отрезке $\alpha \leq t \leq \beta$ и $G(\beta) = 1$.

Элементарное применение условия (1) заключается в следующем.

2) Пусть $f(x)$ — выпуклая функция на интервале $[0, b]$, т. е. $\alpha = 0$ и $\beta = b$, и рассмотрим последовательность $\{a_i\}_1^{2m-1}$, удовлетворяющую условию $b > a_1 \geqslant a_2 \geqslant \ldots \geqslant a_{2m-1} \geqslant 0$. Тогда

$$\sum_{j=1}^{2m-1} (-1)^{j-1} f(a_j) \geqslant f\left(\sum_{j=1}^{2m-1} (-1)^{j-1} a_j\right) \tag{7.26}$$

(см. Сеге [1950]).

В самом деле, пусть $G(t)$ — функция ограниченной вариации, которая имеет скачок $(-1)^{j-1}$ при $a_j$. Очевидно, что $0 \leq G(t) \leq 1$ и $G(b) = 1$. Положим $X(t) = t$; тогда неравенство (7.24) переходит в неравенство (7.26).

Если $m$ — четное, можно доказать, что

$$\sum_{j=1}^{2m} (-1)^{j-1} f(a_j) \geqslant f\left(\sum_{j=1}^{2m} (-1)^{j-1} a_j\right)$$

(см. Беллман [1953]) при условии $f(0) \leq 0$. Действительно, пусть $G(t)$—функция ограниченной вариации, которая имеет скачок $(-1)^{j-1}$ при $a_j$ и 1 в точке 0. В этом случае еравенство (7.24) сводится к следующему;

$$f(0) + \sum_{j=1}^{2m} (-1)^{j-1} f(a_j) \geqslant f\left(\sum_{j=1}^{2m} (-1)^{j-1} a_j\right),$$

и отсюда следует справедливость неравенства (7.27), так как $f(0) \leq 0$. Если функция $G(t)$ имеет скачки $(-1)^{j-1} w_j$ при $a_j$ $(j = 1, 2, \ldots, r)$, где $1 \geqslant w_1 \geqslant w_2 \geqslant \ldots \geqslant w_r \geqslant 0$, и последовательность $\{a_i\}_1^r$ удовлетворяет условию $b \geqslant a_1 \geqslant a_2 \geqslant \ldots \geqslant a_r \geqslant 0$, то аналогичным способом можно получить неравенство

$$\left[1 - \sum_{1}^{r} (-1)^{j-1} w_j\right] f(0) + \sum_{1}^{r} (-1)^{j-1} w_j f(a_j) \geqslant f\left(\sum_{1}^{r} (-1)^{j-1} w_j a_j\right)$$

(см. Олкин [1959]).

Пример 7.10. Другое простое приложение теоремы 5.4: функция $\left(\sum_{i=1}^{n} a_i^r \Big/ \sum_{i=1}^{n} b_i^r\right)^{1/r}$ — монотонная возрастающая относительно $r$ при условии, что монотонно возрастающей является последовательность $a_i/b_i$, причем $a_1 \geqslant a_2 \geqslant \ldots \geqslant a_n > 0$ и $b_1 \geqslant b_2 \geqslant \ldots \geqslant b_n > 0$. Существует также непрерывный аналог этого результата. Этот пример был предоставлен нам Олкином, который предложил другое доказательство.

## § 8. Обобщенные абсолютно монотонные функции

Пусть функция $\varphi$ определена на интервале $(0, b)$, и при этом $\varphi^{(k)}(t) \geqslant 0$ для $t \in (0, b)$ и всех $k = 0, 1, \ldots$, т. е. пусть функция $\varphi$ является абсолютно монотонной на интервале $(0, b)$. Известно, что функцию, удовлетворяющую указанным условиям, можно разложить в сходящийся степенной ряд

$$\varphi(t) = \sum_{k=0}^{\infty} \varphi^{(k)}(0+) \frac{t^k}{k!}, \qquad t \in [0, b). \tag{8.1}$$

Далее, если $u_i(t) = t^i/i!$, $i = 0, 1, \ldots$, то класс абсолютно монотонных функций на интервале $(0, b)$ совпадает с конусом $C^+ \cap \left[\bigcap_{k=0}^{\infty} C(u_0, u_1, \ldots, u_k)\right]$, где $C^+ = C^+(0, b)$ обозначает конус неотрицательных функций на $[0, b]$. Действительно, если $u_i(t) = t^i/i!$, то $w_i(t) \equiv 1$, $i = 0, 1, \ldots$ и $D_{k-1} \ldots D_0\varphi(t) = \varphi^{(k)}(t)$, $k = 1, 2, \ldots$ Из теоремы 3.2 следует, что $\varphi \in C^+ \cap \left[\bigcap_{k=0}^{\infty} C(u_0, \ldots, u_k)\right]$ тогда и только тогда, когда $\varphi(t) \geqslant 0$, $t \in (0, b)$ и $D_{k-1} \ldots D_0\varphi(t) = \varphi^{(k)}(t) \geqslant 0$, $k = 1, 2, \ldots$, $t \in (0, b)$.

В задачу этого параграфа входит обобщение представления (8.1) на функции $\varphi$, принадлежащие $C^+ \cap \left[\bigcap_{k=0}^{\infty} C(u_0, u_1, \ldots, u_k)\right]$, где $\{u_i\}_0^k$ $(k = 0, 1, 2, \ldots)$ — *ECT*-система на $[a, b]$, обладающая представлением (2.1).

Теорема 8.1. *Пусть $\{u_i\}_0^\infty$ представляет собой систему функций на интервале $[a, b]$ такую, что $\{u_i\}_0^k$ — ECT-система вида (2.1) для любого целого $k$. Пусть $m_i$ и $M_i$ определяются следующим образом:*

$$0 < m_i = \min_{a \leqslant t \leqslant b} w_i(t) \leqslant \max_{a \leqslant t \leqslant b} w_i(t) = M_i, \quad i = 0, 1, \ldots \tag{8.2}$$

*Если существует* $\varepsilon > 0$ *такое, что*

$$\lim_{n\to\infty}\left(\prod_{i=0}^{n}\frac{M_i}{m_i}\right)\varepsilon^n = 0, \tag{8.3}$$

*то любая функция* $\varphi$, *принадлежащая* $C^+ \cap \left[\bigcap_{k=0}^{\infty} C(u_0, \ldots, u_k)\right]$, *представима в виде следующего сходящегося ряда*:

$$\varphi(t) = \sum_{k=0}^{\infty} \rho_k(a+)\, u_k(t), \qquad t \in (a, b), \tag{8.4}$$

где $\rho_0(t) = \dfrac{\varphi(t)}{w_0(t)}$ и $\rho_k(t) = \dfrac{D_{k-1}\ldots D_1 D_0\varphi(t)}{w_k(t)}$, $k = 1, 2, \ldots$

Замечание 8.1. Условие (8.3) обязательно выполняется, если $w_i(t)$ равномерно ограничены сверху и снизу.

Замечание 8.2. Сходимость ряда (8.4) — равномерная на любом компактном подынтервале интервала $(a, b)$. Это непосредственно вытекает из теоремы Дини, так как все слагаемые, включая предельную функцию, неотрицательны и непрерывны. Кроме того, если функция $\varphi(t)$ является непрерывной в граничной точке $b$, то представление (8.4) справедливо и в точке $b$. Чтобы доказать это, заметим, что так как $\varphi(t)/w_0(t)$ непрерывна в точке $b$ и $\varphi(t)/w_0(t)$ — неубывающая функция, то для заданного $\varepsilon > 0$ существует достаточно малое $\eta(\varepsilon)$, такое, что

$$\frac{\varphi(b)}{w_0(b)} \le \frac{\varphi(b-\eta)}{w_0(b-\eta)} + \varepsilon.$$

Далее сходимость ряда (8.4) в точке $b-\eta$ означает, что для достаточно большого $n$

$$\frac{\varphi(b-\eta)}{w_0(b-\eta)} \le \sum_{k=0}^{n} \rho_k(a+)\frac{u_k(b-\eta)}{w_0(b-\eta)} + \varepsilon \le \sum_{k=0}^{n} \rho_k(a+)\frac{u_k(b)}{w_0(b)} + \varepsilon \le$$

$$\le \sum_{k=0}^{\infty} \rho_k(a+)\frac{u_k(b)}{w_0(b)} + \varepsilon,$$

второе неравенство следует из того, что $u_k(t)/w_0(t)$—неубывающая функция. Если $\varepsilon \to 0$, то

$$\frac{\varphi(b)}{w_0(b)} \le \sum_{k=0}^{\infty} \rho_k(a+)\frac{u_k(b)}{w_0(b)}.$$

С другой стороны, $\varphi(t)/w_0(t)$ — неубывающая функция, так что

$$\frac{\varphi(b)}{w_0(b)} \geqslant \frac{\varphi(b-\varepsilon)}{w_0(b-\varepsilon)} \geqslant \sum_{k=0}^{n} \rho_k\,(a+)\,\frac{u_k(b-\varepsilon)}{w_0(b-\varepsilon)}.$$

Устремляя $\varepsilon$ к нулю, получаем, что

$$\frac{\varphi(b)}{w_0(b)} \geqslant \sum_{k=0}^{n} \rho_k\,(a+)\,\frac{u_k(b)}{w_0(b)},$$

и это неравенство справедливо для всех $n$.

Следовательно,

$$\varphi(b) = \sum_{k=0}^{\infty} \rho_k\,(a+)\,u_k(b).$$

Для $t=a$ ситуация еще проще, так как $u_k(a)=0,\ k=1,2,\ldots$

Если функция $\varphi$ непрерывна при $t=a$ и $t=b$, то сходимость ряда (8.4) на интервале $[a, b]$ равномерная.

Доказательство теоремы 8.1. Пусть функция $\varphi$ принадлежит конусу $\overset{+}{C} \cap \left[\bigcap\limits_{k=0}^{\infty} C(u_0, \ldots, u_k)\right]$. Так как функция $\varphi$, в частности, принадлежит конусам $C(u_0, \ldots, u_{n-1})$ и $C(u_0, \ldots, u_n)$, то из теоремы 8.2 следует, что

$$\rho_n(t) = \frac{D_{n-1}\ldots D_0\varphi(t)}{w_n(t)}$$

—неотрицательная и неубывающая функция. В этом случае $\rho_n(a+) \geqslant 0$, так что, согласно следствию 2.1, функция $\varphi(t)$ представлена в виде

$$\varphi(t) = \int_a^b \varphi_n(t;x)\,d\rho_n(x) + \sum_{k=0}^{n} \rho_k(a+)\,u_k(t). \tag{8.5}$$

Так как $\sum\limits_{k=0}^{n} \rho_k\,(a+)\,u_k(t)$ — невозрастающая функция от $n$ при фиксированном $t$, то из равенства (8.5) заключаем, что

$$g_n(t) = \int_a^b \varphi_n(t;x)\,d\rho_n(x) \tag{8.6}$$

является невозрастающей последовательностью по $n$. Следовательно, из того факта, что $g_n(t) \geqslant 0$, делаем вывод о существовании предела

$$\lim_{n\to\infty} g_n(t) = g(t) \geqslant 0. \tag{8.7}$$

Докажем, что $g(t)=0$ для $t\in[a, b)$. С этой целью сначала покажем, что для любой функции $\varphi$, принадлежащей конусу $C^+\cap \cap\left[\bigcap_{k=0}^{\infty} C(u_0, \ldots, u_k)\right]$, и любого $c\in[a, b)$ и $t$, принадлежащего некоторому невырожденному интервалу $[c, c+b)$,

$$\lim_{n\to\infty}\int_c^b \varphi_n(t; x)\,d\rho_n(x; \varphi)=0, \tag{8.8}$$

где $\rho_n(x; \varphi)=\dfrac{D_{n-1}\ldots D_1D_0\varphi(x)}{w_n(x)}$.

Допустим, что соотношение (8.8) доказано. Тогда известно, что $g(t)=0$ на некотором интервале $[a, a+\delta)$. Пусть $[a, c^*)$ представляет собой максимальный интервал, левым концом которого является $a$, на котором $g(t)=0$. Согласно следствию 3.2

$$\varphi_n(t; x)\in C^+\cap\left[\bigcap_{k=0}^{m} C(u_0, \ldots, u_k)\right],\qquad n\geqslant m,$$

и, следовательно,

$$g(t)=\lim_{n\to\infty}\int_a^b \varphi_n(t; x)\,d\rho_n(x)\in C^+\cap\left[\bigcap_{k=0}^{m} C(u_0, \ldots, u_k)\right].$$

Но это справедливо для всех $m$, и поэтому

$$g(t)\in C^+\cap\left[\bigcap_{k=0}^{\infty} C(u_0, \ldots, u_k)\right]. \tag{8.9}$$

Так как $g(t)=0$ для $t\in[a, c^*)$, представление (8.5) для $g(t)$ сводится к $g(t)=\int_{c^*}^{b}\varphi_n(t; x)\,d\rho_n(x; g)$. Применяя соотношение (8.8), получаем

$$g(t)=\lim_{n\to\infty}\int_{c^*}^{b}\varphi_n(t; x)\,d\rho_n(x; g)=0$$

для $t\in[c^*, c+\delta)$ и некоторого $\delta>0$. Это противоречит тому, что интервал $[a, c^*)$ является максимальным, если $c^*\neq b$. Таким образом, $g(t)\equiv 0$ для $t\in[a, b)$, и выражение (8.5) преобразуется в следующее:

$$\varphi(t)=\sum_{k=0}^{\infty}\rho_k(a+)\,u_k(t),\qquad t\in(a, b).$$

Обратимся теперь к доказательству утверждения, сформулированного в (8.8). Рассматривая выражение (8.5), убеждаемся, что для

$$a \leq x \leq t < \beta \leq b,$$

$$\varphi(\beta) \geqslant \int_x^\beta \varphi_n(\beta; \xi)\, d\rho_n(\xi) = \int_x^\beta \varphi_n(\beta; \xi)\, w_{n+1}(\xi)\, \rho_{n+1}(\xi)\, d\xi \geqslant$$

$$\geqslant \rho_{n+1}(x) \int_x^\beta \varphi_n(\beta; \xi)\, w_{n+1}(\xi)\, d\xi \geqslant \rho_{n+1}(x) \int_t^\beta \varphi_n(\beta; \xi)\, w_{n+1}(\xi)\, d\xi.$$

Кроме того, из определения $\varphi_n(\beta; \xi)$ следует, что

$$\int_t^\beta \varphi_n(\beta; \xi)\, w_{n+1}(\xi)\, d\xi \geqslant \left(\prod_{i=0}^{n+1} m_i(t; \beta)\right) \frac{(\beta - t)^{n+1}}{(n+1)!}, \tag{8.10}$$

где $m_i(t; \beta) = \min\limits_{t \leqslant z \leqslant \beta} w_i(z)$, $i = 0, 1, \ldots, n+1$. Следовательно, если $a \leq x \leq t < \beta \leq b$, то

$$\rho_{n+1}(x) \leq \frac{\varphi(\beta) \cdot (n+1)!}{(b - t)^{n+1} \prod\limits_{i=0}^{n+1} m_i(t; \beta)}. \tag{8.11}$$

Учитывая неравенство (8.11), приходим к выводу, что если $c < t < \beta$, то

$$\int_c^b \varphi_n(t; x)\, d\rho_n(x) = \int_c^t \varphi_n(t; x)\, w_{n+1}(x)\, \rho_{n+1}(x)\, dx \leq$$

$$\leq \frac{\varphi(\beta)\,(n+1)!}{(\beta - t)^{n+1} \prod\limits_{i=0}^{n+1} m_i(t; \beta)} \int_c^t \varphi_n(t; x)\, w_{n+1}(x)\, dx \leq$$

$$\leq \varphi(\beta) \left(\frac{t - c}{\beta - t}\right)^{n+1} \prod_{i=0}^{n+1} \frac{M_i(c; t)}{m_i(t; \beta)},$$

где $M_i(c; t) = \max\limits_{c \leqslant z \leqslant t} w_i(z)$, $i = 0, 1, \ldots, n+1$.

Учитывая соотношения (8.2) и (8.3), заключаем, что

$$\lim_{n \to \infty} \int_c^\beta \varphi_n(t; x)\, d\rho_n(x) = 0$$

для $t$ в некотором невырожденном интервале $[c, c + \delta]$.

Замечание 8.3. Отметим, что, когда $w_i(t)$ — неубывающая функция по $t$, оценка в выражении (8.12) сводится к следующему:

$$\int_c^b \varphi_n(t; x)\, d\rho_n(x) \leq \varphi(\beta)\left(\frac{t-c}{\beta-t}\right)^{n+1},$$

так как в этом случае $m_i(t; \beta) = M_i(c; t)$.

С помощью теоремы 8.1 можно теперь описать конус, дуальный к конусу $C^+ \cap \left[\bigcap_{k=0}^{\infty} C(u_0, \ldots, u_k)\right]$.

Теорема 8.2. *При условиях* (8.2) *и* (8.3) *обобщенная мера* $d\mu$ *принадлежит конусу, дуальному к конусу* $C^+ \cap \left[\bigcap_{k=0}^{\infty} C(u_0, \ldots, u_k)\right]$ *тогда и только тогда, когда*

$$\int_a^b u_i d\mu \geqslant 0, \qquad i = 0, 1, 2, \ldots$$

Доказательство. Из следствия 3.2 получаем

$$u_i \in C^+ \cap \left[\bigcap_{k=0}^{\infty} C(u_0, \ldots, u_k)\right],$$

так что условие (8.13) является необходимым. Кроме того, с помощью представления (8.4), находим, что условие (8.13) является также достаточным.

## § 9. Интерполяция функций обобщенными сплайн-многочленами

Последнее время наблюдается повышенный интерес к теории аппроксимации функций обобщенными сплайн-многочленами. Для ознакомления с историей вопроса и полной библиографией отсылаем читателя к Сарду [1963] и Шенбергу [1964a]. Большинство работ в этой области ограничивается рассмотрением частного случая сплайн-многочленов, т. е., многочленов, которые образуются с помощью обычных многочленов. Мы будем излагать теорию по отношению к обобщенным сплайн-многочленам, индуцированным обобщенной полной $T$-системой. Другой подход к этому общему случаю используется Гревилем [1964 a], [1964 b]; см. также Алберг, Нильсон и Уолш [1964].

Пусть $\{u_i\}_0^n$ означает $ECT$-систему на интервале $[a, b]$ вида (1.5), и определим $\varphi_n(t; x)$ посредством (2.5).

Определение 9.1. Функция $p(t)$, определенная на интервале $[a, b]$, называется *обобщенной сплайн-функцией порядка* $n+1$ *с узлами* $x_1, x_2, \ldots, x_k$ $(a < x_1 < x_2 < \ldots < x_k < b)$, если:

1) на каждом из интервалов $[a, x_1), [x_1, x_2), \ldots, [x_k, b]$ $p(t)$ совпадает с некоторым $u$-многочленом;

2) функция $p(t)$ принадлежит классу $C^{n-1}[a, b]$.

Следующая лемма указывает каноническое представление обобщенного сплайн-многочлена с заданными узлами.

Лемма 9.1. *Сплайн-многочлен с узлами $\{x_i\}_{i=1}^{k}$ представим в следующем виде*:

$$p(t) = \sum_{i=1}^{k} a_i \varphi_n(t; x_i) + \sum_{i=0}^{n} b_i u_i(t). \tag{9.1}$$

Доказательство. Функция вида (9.1), очевидно, удовлетворяет условиям 1) и 2), указанным выше.

Обратно, пусть $p$ является сплайн-многочленом, удовлетворяющим условиям определения 9.1. Таким образом, $p$ совпадает с многочленом на каждом из интервалов $[a, x_1], [x_1, x_2], \ldots, [x_k, b]$. Разумеется, функция $(1/w_n(t))(D_{n-1}^R D_{n-2} \ldots D_0 p)(t)$ определена на каждом из этих интервалов. Пусть $a_i$, $i = 1, 2, \ldots, k$, являются скачками функции $(1/w_n(t))(D_{n-1}^R D_{n-2} \ldots D_0 p)(t)$ в точках $x_i$, $i = 1, 2, \ldots, k$. Рассмотрим обобщенный сплайн-многочлен

$$p^*(t) = \sum_{i=1}^{k} a_i \varphi_n(t; x_i) + \sum_{i=0}^{n} b_i u_i(t),$$

где $\sum_{i=0}^{n} b_i u_i(t) = p(t)$ для $t \in [a, x_1)$. Функция $(D_n D_{n-1} \ldots D_0)(p - p^*)$ тождественно равна нулю. Следовательно, по лемме 2.3 $p - p^*$ является многочленом на всем интервале $[a, b]$. Но $p(t) - p^*(t) = 0$ для $t \in [a, x_1)$ и, таким образом, $p(t) = p^*(t)$ на всем интервале $[a, b]$.

Для дальнейших рассуждений потребуется следующая лемма, представляющая и самостоятельный интерес.

Лемма 9.2. *Пусть $\{u_i\}_0^n$ представляет собой ECT-систему вида* (1.5) *на интервале* $[a, b]$, *пусть функция $\varphi_n(t; x)$ определяется соотношением* (2.5). *Если $\{t_i\}_1^k$ и $\{x_i\}_1^k$ таковы, что $a \leq x_1 < x_2 < \ldots < x_k \leq b$ и $a \leq t_1 < t_2 < \ldots < t_k \leq b$, то*

$$\det \| \varphi_n(t_i; x_j) \|_{i,j=1}^{k} \geq 0, \tag{9.2}$$

*причем строгое неравенство будет иметь место тогда и только тогда, когда*

$$t_{i-n-1} < x_i < t_i, \qquad i = 1, 2, \ldots, k, \tag{9.3}$$

*где случаю $i \leq n+1$ соответствует только правое неравенство. Для $n = 0$ допускается равенство в правой части* (9.3).

Доказательство. Доказательство проводится индукцией по $n$. Для $n = 0$ имеем

$$\varphi_0(t; x) = \begin{cases} w_0(t), & t \geq x, \\ 0, & t < x. \end{cases}$$

Более удобно $\varphi_0(t; x)$ представить следующим образом: $\varphi_0(t; x) = w_0(t) K(t; x)$, где

$$K(t; x) = \begin{cases} 1, & t \geqslant x, \\ 0, & t < x. \end{cases}$$

В примере 9 § 3 гл. I показано, что для $a \leq t_1 < \ldots < t_k \leq b$ и $a \leq x_1 < \ldots < x_k \leq b$

$$K\begin{pmatrix} t_1, \ldots, t_k \\ x_1, \ldots, x_k \end{pmatrix} = \det \| K(t_i; x_j) \|_{i,j=1}^{k} \geqslant 0, \tag{9.4}$$

причем строгое неравенство имеет место тогда и только тогда, когда

$$t_{i-1} < x_i \leq t_i, \qquad i = 1, 2, \ldots, k. \tag{9.5}$$

Отсюда следует, что

$$\det \| \varphi_0(t_i; x_j) \|_{i,j=1}^{k} = \left( \prod_{i=1}^{k} w_0(t_i) \right) K\begin{pmatrix} t_1, \ldots, t_k \\ x_1, \ldots, x_k \end{pmatrix} \geqslant 0,$$

и строгое неравенство имеет место тогда и только тогда, когда выполняется условие (9.5). Таким образом, утверждение теоремы доказано для случая $n = 0$.

Теперь допустим, что теорема доказана для всех $ECT$-систем вида (1.5), включающих, самое большее, $n$ функций. Можно записать

$$\varphi_n(t; x) = w_0(t) \int_a^b K(t; \xi) \tilde{\varphi}_{n-1}(\xi; x) \, d\xi,$$

где

$$\tilde{\varphi}_{n-1}(\xi; x) = \begin{cases} w_1(\xi) \int\limits_x^{\xi} w_2(\xi_2) \int\limits_x^{\xi_2} w_3(\xi_3) \ldots \int\limits_x^{\xi_{n-1}} w_n(\xi_n) d\xi_n \ldots d\xi_2 & \text{для } \xi \geqslant x, \\ 0 & \text{для } \xi < x. \end{cases}$$

Использование композиционной формулы (1.3.12) дает

$$\det \| \varphi_n(t_i; x_j) \|_{i,j=1}^{k} = \prod_{i=1}^{k} w_0(t_i) \int \ldots \int_{a \leqslant \xi_1 < \ldots < \xi_k \leqslant b} K\begin{pmatrix} t_1, \ldots, t_k \\ \xi_1, \ldots, \xi_k \end{pmatrix} \times$$
$$\times \det \| \tilde{\varphi}_{n-1}(\xi_i; x_j) \|_{i,j=1}^{k} d\xi_1 \ldots d\xi_k. \tag{9.6}$$

Из предположения индукции следует, что

$$\det \| \tilde{\varphi}_{n-1}(\xi_i; x_j) \|_{i,j=1}^{k} \geqslant 0,$$

причем строгое неравенство имеет место тогда и только тогда, когда

$$\xi_{i-n} < x_i < \xi_i, \qquad i = 1, 2, \ldots, k, \tag{9.7}$$

и $a \leqslant t_1 < t_2 < \ldots < t_k \leqslant b$, $a \leqslant \xi_1 < \xi_2 < \ldots < \xi_k \leqslant b$. Так как определитель $K\begin{pmatrix} t_1, \ldots, t_k \\ \xi_1, \ldots, \xi_k \end{pmatrix}$ всегда неотрицателен и положителен, только если

$$t_{i-1} < \xi_i \leqslant t_i, \qquad i = 1, 2, \ldots, k, \tag{9.8}$$

то отсюда следует, что определитель (9.6) неотрицателен и строго положителен тогда и только тогда, когда неравенства (9.7) и (9.8) описывают множество положительных $k$-мерных мер Лебега. Это, как легко показать, имеет место при условии, что выполняется соотношение $t_{i-n-1} < x_i < t_i$ и только в этом случае. Теорема доказана.

Теперь можно сформулировать и проанализировать некоторые результаты, относящиеся к интерполяции функций сплайн-многочленами. Система функций $\{u_i\}_0^n$ вида (1.5) определяется задан ем функций $w_0, w_1, \ldots, w_n$, обладающих требуемыми свойствами непрерывности. Необходимо конкретизировать выбор $\{w_i\}$ следующим образом. С этого момента будем считать, что $\{u_i\}_0^N$, где $N = 2n + 1$, задаются соотношением (1.5), где $w_{n+1} \equiv 1$ и $w_{n+i-1} = w_{n-i+1}$ $(w_i \in C^N[a, b])$, $i = 0, 1, 2, \ldots, n$. Другими словами, система функций $\{u_i\}_0^N$ определяется посредством (1.5) с помощью множества положительных функций

$$w_0, w_1, \ldots, w_n, 1, w_n, w_{n-1}, \ldots, w_2, w_1.$$

Удобно ввести в рассмотрение дифференциальные операторы

$$(D_i^* f)(t) = \frac{1}{w_i(t)} \frac{df(t)}{dt}, \qquad i = 0, 1, \ldots, n\,^{1)},$$

и положить

$$L = D_n D_{n-1} \ldots D_0,$$

$$L^* = D_0^* D_1^* \ldots D_n^*.$$

Для системы функций $\{u_i\}_0^N$ обобщенная сплайн-функция $p$ с узлами $x_1, \ldots, x_k$ удовлетворяет условию

$$(L^* L p)(t) = 0, \qquad x_i < t < x_{i+1}, \qquad i = 0, 1, \ldots, k$$

$$(x_0 = a, \; x_{k+1} = b) \tag{9.9}$$

и функция $p$ принадлежит классу $C^{N-1}$.

Важным подклассом сплайн-функций служит класс натуральных сплайнов, определенных следующим образом.

---

[1]) Звездочка означает, что $D_i$ и $D_i^*$ представляют собой сопряженные дифференциальные операторы.

Определение 9.2. По отношению к системе функций $\{u_i\}_0^N$ ($N = 2n + 1$) *натуральный сплайн-многочлен* $p$ *с узлами* $x_1, x_2, \ldots$ $\ldots, x_k$ представляет собой сбобщенный сплайн-многочлен, удовлетворяющий дополнительному условию

$$(Lp)(t) = 0 \tag{9.10}$$

для $t > x_k$ и $t < x_1$.

Напомним, что обобщенная сплайн-функция для системы функций $\{u_i\}_0^N$ имеет вид (9.1), где $n$ заменено на $N$. Так как очевидно, что $\sum_{i=1}^{k} a_i \varphi_N(t; x_i) = 0$, когда $t < x_1$, то условие (9.10) означает, что $L\left(\sum_{i=0}^{N} b_i u_i(t)\right) = 0$ для $t < x_1$. (Мы предостерегаем читателя, чтобы он не путал этот факт со стандартным свойством $L^*L\left(\sum_{i=0}^{N} b_i u_i(t)\right) = 0$, которое выполняется для любого многочлена). С учетом леммы 2.3 можно заключить, что $b_i = 0$, $i = n+1, \ldots$ $\ldots, N$. Следовательно, любой натуральный сплайн-многочлен по отношению к $ECT$-системе $\{u_i\}_{i=0}^N$, имеющий узлы $x_1, \ldots, x_k$, допускает представление вида

$$p(t) = \sum_{i=1}^{k} a_i \varphi_N(t; x_i) + \sum_{i=0}^{n} b_i u_i(t) \tag{9.11}$$

(таким образом, $N$ понижается до $n$ во второй сумме), где на $a_i$ наложено дополнительное ограничение

$$\sum_{i=1}^{k} a_i L\varphi_N(t; x_i) = 0, \qquad t > x_k.$$

Теперь можно сформулировать основную теорему интерполяции для натуральных сплайн-многочленов.

Теорема 9.1. *Пусть заданы числа* $\{x_i\}_1^k$ ($k \geqslant n+1$), *удовлетворяющие условию* $a < x_1 < \ldots < x_k < b$. *Тогда для любого множества вещественных чисел* $y_1, y_2, \ldots, y_k$ *существует единственный натуральный сплайн-многочлен* $p$ *с узлами в точках* $x_1, x_2, \ldots$ $\ldots, x_k$ *такой, что* $p(x_i) = y_i$, $i = 1, 2, \ldots, k$.

Замечание 9.1. Для случая, когда $u_i(t) = t^i$, $i = 0, 1, 2, \ldots, N$, эта теорема впервые рассматривалась де Буром [1963]. Доказательство в этой работе основывается на теореме Шенберга и Уитни [1953], которая является вариантом леммы 9.2 для случая степенных функций.

Доказательство. Справедливость утверждения теоремы будет доказана в том случае, если можно получить коэффициенты

$\{a_i\}_1^k$ и $\{b_i\}_0^n$ со свойством

$$\sum_{i=1}^{k} a_i \varphi_N (x_j; x_i) + \sum_{i=0}^{n} b_i u_i (x_j) = y_j, \qquad j = 1, 2, \ldots, k, \quad (9.12)$$

и удовлетворяющие также условию

$$\sum_{i=1}^{k} a_i L\varphi_N (t; x_i) = 0, \qquad t > x_k. \quad (9.13)$$

Выберем любые $n+1$ точек $\{x_{k+j}\}_{j=1}^{n+1}$, удовлетворяющих условию $x_{k+1} < x_{k+2} < \ldots < x_{k+n+1} < b$, и заменим (9.13) требованием

$$\sum_{i=1}^{k} a_i L\varphi_N (x_{k+j}; x_i) = 0, \qquad j = 1, 2, \ldots, n+1. \quad (9.14)$$

Ниже будет показано, что из (9.14) следует (9.13).

Соотношения (9.12) и (9.14) образуют систему $k+n+1$ линейных уравнений с $k+n+1$ неизвестными $\{a_i\}_1^k$ и $\{b_i\}_0^n$. Докажем теперь, что матрица этой системы уравнений невырожденная. Предположим противное, т. е. что соответствующий определитель равен нулю. Тогда существуют константы $\{e_i\}_1^k$ и $\{f_i\}_0^n$, не все равные нулю, такие, что

$$\sum_{i=1}^{k} e_i \varphi_N (t; x_i) + \sum_{i=0}^{n} f_i u_i (t) = 0 \quad (9.15)$$

при $t = x_j$, $j = 1, 2, \ldots, k$, и

$$\sum_{i=1}^{k} e_i L\varphi_N (x_{k+r}; x_i) = 0, \qquad r = 1, 2, \ldots, n+1. \quad (9.16)$$

Применяя оператор $D_0$ к функции, стоящей в левой части равенства (9.15), и используя теорему Ролля, заключаем, что результирующее выражение обращается в нуль на множестве точек $\{\xi\}_1^{k-1}$, удовлетворяющих условию $x_1 < \xi_1 < x_2 < \xi_2 < \ldots < \xi_{k-1} < x_k$. Таким же образом, последовательно применяя операторы $D_1, D_2, \ldots, D_n$, в результате получаем следующие равенства:

$$\sum_{i=1}^{k} e_i L\varphi_N (x_j^*; x_i) = 0, \qquad j = 1, 2, \ldots, k-n-1, \quad (9.17)$$

где $\{x_j^*\}$ удовлетворяет условию

$$x_j < x_j^* < x_{j+n+1}, \qquad j = 1, 2, \ldots, k-n-1. \quad (9.18)$$

Объединяя (9.16) и (9.17), получаем

$$\sum_{i=1}^{k} e_i L\varphi_N(t; x_i) = 0 \tag{9.19}$$

для $t = x_1^*, \ldots, x_{k-n-1}^*, x_{k+1}, \ldots, x_{k+n+1}$. Заметим, что

$$L\varphi_N(t; x) = \int_x^t w_n(\xi_n) \int_x^{\xi_n} w_{n-1}(\xi_{n-1}) \ldots \int_x^{\xi_2} w_1(\xi_1)\, d\xi_1 \ldots d\xi_{n-1} d\xi_n$$

для $t \geqslant x$ и $L\varphi_N(t; x) = 0$ при $t < x$. Другими словами, $L\varphi_N(t; x) = \tilde{\varphi}_n(t; x)$ представляет собой ядро вида (2.5), базирующееся на функциях $1, w_n(t), w_{n-1}(t), \ldots, w_1(t)$.

Можно применить лемму 9.2, согласно которой определитель системы уравнений (9.19) положителен, если только $\{x_i\}_{i=1}^k$ и $\{x_i^*\}_{i=1}^k$ ($x_i^* = x_{i+n+1}$ для $i = k-n, \ldots, k$) удовлетворяют следующим неравенствам:

$$\begin{aligned} &x_{i-n-1}^* < x_i < x_i^*, \qquad && i = 1, 2, \ldots, k-n-1, \\ &x_{i-n-1}^* < x_i < x_{n+1+i}, \qquad && i = k-n, \ldots, k. \end{aligned} \tag{9.20}$$

В соответствии с выбором $x_{k+1}, \ldots, x_{k+n+1}$ и с учетом условия (9.18) видно, что требования (9.20) выполняются. На основании леммы 9.2 заключаем, что $e_i = 0$, $i = 1, 2, \ldots, k$. В этом случае из равенств (9.15) следует, что $\sum_{i=0}^{n} f_i u_i(t) = 0$ для $t = x_1, \ldots, x_k$. Так как $k \geqslant n+1$ и $\{u_i\}_0^n$ образуют $ECT$-систему, то отсюда следует, что $f_i = 0$, $i = 0, 1, \ldots, n$. Это противоречие означает, что система уравнений (9.12) и (9.14) имеет единственное решение.

Доказательство существования завершится, как только мы покажем, что из (9.14) следует (9.13). Очевидно, что

$$L^*\left(\sum_{i=1}^{k} a_i L\varphi_N(t; x_i)\right) = 0, \qquad t > x_k,$$

так что по лемме 2.8 $\sum_{i=1}^{k} a_i L\varphi_N(t; x_i)$ совпадает на $[x_k, b]$ с многочленом по $T$-системе, образованной функциями $1, w_n, w_{n-1}, \ldots, w_1$. В этом случае выполнение соотношений (9.14) означает и справедливость условия (9.13), так как $k \geqslant n+1$. Доказательство теоремы закончено.

Пользуясь теоремой интерполяции, мы далее рассмотрим некоторые оптимальные свойства натуральных сплайн-многочленов.

Свойства наилучшей среднеквадратичной аппроксимации натуральными сплайн-многочленами. Для каждой пары функций $g$ и $h$ из класса $C^{n+1}[a, b]$ определим квазискалярное произведение

$$(g, h)_n = \int_a^b (Lg)(Lh)\,dt,$$

и пусть $\|g\|_n^2 = (g, g)_n$.

Характерное экстремальное свойство натуральных сплайн-многочленов составляет содержание следующей теоремы.

Теорема 9.2. *Пусть* $x_1, x_2, \ldots, x_k$ $(k \geqslant n+1)$ *удовлетворяют условию* $a < x_1 < x_2 < \ldots < x_k < b$, *и пусть* $y_1, y_2, \ldots, y_k$ — *произвольные вещественные числа. Пусть далее* $p(t)$ *обозначает единственную натуральную сплайн-функцию с узлами в точках* $x_1, x_2, \ldots, x_k$, *интерполирующую* $y_i$ *в точках* $x_i$, *т. е.* $p(x_i) = y_i$, $i = 1, 2, \ldots, k$ (*см. теорему* 9.1). *Тогда среди всех функций* $g$, *принадлежащих классу* $C^{n+1}[a, b]$, *для которых* $g(x_i) = y_i$, $i = 1, 2, \ldots, k$, *натуральный спайн однозначно минимизирует* $\|g\|_n^2$.

Доказательство теоремы 9.2 основывается на следующей лемме.

Лемма 9.3. *Если функции* $g$ *и* $p$ *удовлетворяют условиям теоремы* 9.2, *то*

$$\|g\|_n^2 - \|p\|_n^2 = \|g - p\|_n^2.$$

Доказательство. Рассмотрим функцию $h = g - p$. Тогда $\|g\|_n^2 = \|h\|_n^2 + \|p\|_n^2 + 2(h, p)_n$. Поэтому достаточно показать, что $(h, p)_n = 0$. Интегрирование по частям приводит к формуле

$$(h, p)_n = \int_a^b (D_n D_{n-1} \ldots D_0 h)(D_n D_{n-1} \ldots D_0 p)\,dt =$$

$$= \frac{1}{w_n} (D_{n-1} \ldots D_0 h)(D_n \ldots D_0 p)\Big|_a^b - \int_a^b (D_{n-1} \ldots D_0 h)(D_n^* D_n \ldots D_0 p)\,dt.$$

Но $p$ — натуральный сплайн, так что

$$(L_p)(t) = 0, \tag{9.21}$$

если $t > x_k$ и $t < x_1$, и, следовательно, $(h, p)_n = -\int_a^b (D_{n-1} \ldots$ $\ldots D_0 h)(D_n^* D_n \ldots D_0 p)\,dt$. Повторяя этот процесс еще $n - 1$ раз, получаем равенство

$$(h, p)_n = (-1)^n \int_a^b (D_0 h)(D_1^* D_2^* \ldots D_n^* D_n \ldots D_1 D_0 p)\,dt. \tag{9.22}$$

Так как $L_n^* L_n p = 0$ на любом интервале $[x_i, x_{i+1}]$, приходим к выводу, что $D_1^*D_2^* \dots D_n^*D_n \dots D_1D_0p$ равно некоторой константе $c_i$ на любом из интервалов $[x_i, x_{i+1}), i = 1, \dots, k-1$, так что

$$(h, p)_n = (-1)^n \sum_{i=1}^{k-1} c_i \int_{x_i}^{x_{i+1}} (D_0h)\, dt =$$

$$= (-1)^n \sum_{i=1}^{k-1} c_i \left[ \frac{h(x_{i+1})}{w_0(x_{i+1})} - \frac{h(x_i)}{w_0(x_i)} \right].$$

Так как $g(x_i) = y_i = p(x_i)$ и, таким образом, $h(x_i) = g(x_i) - p(x_i) = 0$, $i = 1, \dots, k$, заключаем, что $(h, p)_n = 0$. Итак, лемма 9.3 доказана.

Доказательство теоремы 9.2. Теорема 9.2 непосредственно следует из леммы 9.3. Действительно, если функции $g$ и $p$ удовлетворяют условиям теоремы, то $\|g\|_n^2 - \|p\|_n^2 = \|g-p\|_n^2 \geqslant 0$. Кроме того, строгое неравенство имеет место, если $\|g-p\|^2 \neq 0$ или $D_n \dots$ $\dots D_0(g-p)(t) = 0$ почти всюду. Тогда по лемме 2.3 $g - p$ есть многочлен вида $\sum_{i=0}^{n} b_i u_i(t)$ на $[a, b]$. Функция $g - p$ обращается в нуль в $k$ точках $x_1, \dots, x_k$, причем $k \geqslant n+1$, и, следовательно, $g(t) - p(t) \equiv 0$.

Теорема 9.3. *Пусть $N$ представляет собой класс всех натуральных сплайн-функций с узлами в точках $x_1, \dots, x_k$ $(a < x_1 < \dots$ $\dots < x_k < b)$, и пусть $p^*$ — натуральный сплайн, интерполирующий данную функцию $f$, принадлежащую классу $C^{n+1}[a, b]$, в точках $x_1, \dots, x_k$. Если $p \in \mathcal{N}$ и разность $p^* - p$ не представляет собой многочлен, то*

$$\|f - p^*\|_n < \|f - p\|_n.$$

Доказательство. Пусть $p$ — произвольный элемент из $\mathcal{N}$. Тогда единственная натуральная сплайн-функция, существование которой гарантируется теоремой 9.1 и которая интерполирует $f - p$ в точках $x_1, \dots, x_k$, есть функция $p^* - p$. По лемме 9.3 имеем

$$\|f - p^*\|_n^2 = \|f - p - (p^* - p)\|_n^2 = \|f - p\|_n^2 - \|p^* - p\|_n^2.$$

Следовательно, $\|f - p^*\|_n^2 < \|f - p\|_n^2$, если $\|p^* - p\|_n^2 \neq 0$, т. е. если разность $p^* - p$ не равна многочлену $\sum_{i=0}^{n} b_i u_i(t)$.

Методы доказательства теорем 9.2 и 9.3 аналогичны методам, впервые использованным де Буром для случая обычных сплайн-многочленов (см. де Бур [1963]).

## § 10. Наилучшие квадратурные формулы, использующие натуральные сплайн-многочлены

В этом параграфе мы представляем фундаментальное описание натуральных сплайн-многочленов, введенных в § 9. Эти результаты обобщают некоторые результаты из работы Шенберга [1964a], в которой рассматривалась задача нахождения наилучших квадратурных формул, приближающих некоторые линейные операторы для случая обычных многочленов. Естественная формулировка более общих результатов, представленных ниже, была недавно дана де Буром и Линчем [1966] на основе теории проективных операторов в конечномерных гильбертовых пространствах с воспроизводящими ядрами. Некоторые экстремальные свойства натуральных сплайн-многочленов рассматривались также с точки зрения гильбертова пространства в более ранней работе Голомба и Вейнбергера [1959] (см. также Вейнбергер [1961]). Последняя работа, кроме всего прочего, содержит элегантную геометрическую интерпретацию сплайн-многочленов. Методы, использованные вышеупомянутыми авторами, относятся к явным, при этом используется удобное выражение для воспроизводящего ядра. В общем случае представить воспроизводящее ядро в удобной форме оказывается сложной задачей. Перейдем к непосредственному анализу некоторых соотношений двойственности, имеющих значение в теории $ET$-систем. Некоторые из сопутствующих лемм могут представлять также самостоятельный интерес.

Начнем с постановки задачи. Пусть $\mathscr{L}$ обозначает линейный функционал, определенный для функций $f$, принадлежащих классу $C^{n+1}[a, b]$, и пусть $\{u_i\}_0^n$ представляет собой $ECT$-систему на интервале $[a, b]$ вида (1.5), образованную $n+1$ положительными функциями $1, w_1, \ldots, w_n$ $(w_i \in C^{2n+1}[a, b])$. Примерами $\mathscr{L}$ служат функционалы типа

$$\mathscr{L}f = \sum_{\nu,\mu=1}^{n+1} a_{\mu\nu} \int_a^b f^{(\mu)}(x)\, d\sigma_\nu(x),$$

где $\sigma_\nu(x)$ обозначают обобщенные меры ограниченной вариации.

Для фиксированного множества из $k$ точек $(k \geqslant n+1)$ $\{x_i\}_1^k$ $(a < x_1 < x_2 < \ldots < x_k < b)$ запишем

$$\mathscr{L}f = \sum_{\nu=1}^{k} B_\nu f(x_\nu) + \mathscr{R}f. \tag{10.1}$$

Константы $B_\nu$ должны удовлетворять следующим условиям:

$$\mathscr{L}u_i = \sum_{\nu=1}^{k} B_\nu u_i(x_\nu), \qquad i = 0, 1, \ldots, n, \tag{10.2}$$

т. е. $\mathcal{R}u_i = 0$, $i = 0, 1, \ldots, n$. В случае, когда $k = n + 1$, условия (10.2) однозначно определяют костанты $\{B_\nu\}_1^k$. Однако при $k > n + 1$ имеется еще $k - n - 1$ свободных параметров. Выражение $\sum_{\nu=1}^{k} B_\nu f(x_\nu)$ представляет собой квадратурную формулу, аппроксимирующую $\mathcal{L}f$, которая точна для функций $f$, являющихся решением уравнения $D_n \ldots D_1 D_0 f = 0$. Член $\mathcal{R}f$ является остатком квадратурной формулы.

Далее для любой функции $f$, принадлежащей классу $C^{n+1}[a, b]$, повторное интегрирование по частям (см. доказательство леммы 2.2) приводит к формуле

$$f(t) = \int_a^b \varphi_n(t; x)\,[(D^n f)(x)]\,dx + \sum_{i=0}^{n} a_i u_i(t) \tag{10.3}$$

$$(D^n = D_n D_{n-1} \ldots D_0),$$

поэтому

$$\mathcal{R}f = \int_a^b \mathcal{R}_t \varphi_n(t; x)\, D^n f(x)\, dx\,^{1)}.$$

Рассматриваемая здесь задача состоит в определении постоянных $B_1, \ldots, B_k$, удовлетворяющих условию (10.2), при которых интеграл

$$\int_a^b (\mathcal{R}_t \varphi_n(t; x))^2\, dx \tag{10.4}$$

достигает минимального значения. Величина интеграла (10.2) есть норма линейного функционала $\mathcal{R}$ в гильбертовом пространстве всех таких функций $f$, для которых

$$\int_a^b [D^n f(x)]^2\, dx < \infty.$$

Заметим, что формулировка задачи минимизации включает только $ECT$-систему $u_0, u_1, \ldots, u_n$. Однако замечательно то, что решение ее требует введения натуральных сплайн-многочленов для соответствующей $ECT$-системы $2n + 2$ функций. В этом случае результат приводимой ниже теоремы 10.1 раскрывает неожиданный внутренний смысл понятия сплайн-многочленов. Чтобы сформулировать основной результат, положим, что $\{u_i\}_0^N$ ($N = 2n + 1$) представляет собой $ECT$-систему, образованную, как в (1.5), функциями

$$1, w_1, \ldots, w_n, 1, w_n, \ldots, w_1. \tag{10.5}$$

---

[1]) Индекс $t$ указывает на то, что операция осуществляется по переменной $t$, в то время как $x$ остается фиксированным.

Пусть $L_\nu(t)$ обозначают фундаментальные натуральные сплайн-многочлены Лагранжа, соответствующие точкам $x_1, \ldots, x_n$, т. е. $L_\nu(t)$ есть единственный натуральный сплайн-многочлен по отношению к $ECT$-системе, образованной функциями (10.5), который имеет узлы $x_1, \ldots, x_k$, и $L_\nu(x_\mu) = \delta_{\nu\mu}$, $\nu, \mu = 1, 2, \ldots, k$ (ср. определение 9.2 и теорему 9.1).

Следующая теорема является основной в данном параграфе.

Теорема 10.1. *Квадратурная формула для оператора* $\mathcal{L}$ *вида* (10.1), *которая минимизирует норму* (10.4) *при условиях* (10.2), *однозначно определяется заданием коэффициентов* $B_\nu$ *в виде*

$$B_\nu^* = \mathcal{L}L_\nu, \qquad \nu = 1, \ldots, k. \tag{10.6}$$

Замечание 10.1. Кроме решения важной вариационной задачи, имеющей значение в теории интерполяции, ценность и изящество теоремы 10.1 повышаются благодаря тому замечательному свойству, что независимо от специфического вида функционала $\mathcal{L}$ оптимальный выбор констант состоит в простом применении данного функционала $\mathcal{L}$ к специальным натуральным сплайн-многочленам $L_\nu$.

Прежде чем приступить к доказательству, введем дополнительные обозначения и докажем две вспомогательные леммы, представляющие самостоятельный интерес. Вспомним, что с данным множеством положительных фукций $1, w_1, \ldots, w_n$ ($w_i \in C^N[a,b]$) связана функция $\varphi_n(t; x)$, определяемая формулой (2.5). Функции $\varphi_n^*(t; x)$ и $\Omega_n(t; x)$ будут необходимы для дальнейшего анализа. Функция $\varphi_n^*(t; x)$ определяется формулой (2.5) с использованием функций $1, w_n, \ldots, w_1$, т. е. простым изменением на обратный порядка образующих функций $w_1, \ldots, w_n$. Функция $\Omega_n(t; x)$ представляет собой неусеченный вариант функции $\varphi_n(t; x)$, определенной для всех $t$ и $x$ в $[a, b]$ посредством

$$\Omega_n(t; x) = \int_x^t w_1(\xi_1) \int_x^{\xi_1} w_2(\xi_2) \ldots \int_x^{\xi_{n-1}} w_n(\xi_n)\, d\xi_n d\xi_{n-1} \ldots d\xi_1. \tag{10.7}$$

Функция $\Omega_N(t; x)$ для $N = 2n + 1$ описывается выражением (10.7) с использованием усеченной системы (10.5).

При указанных выше определениях имеют место следующие тождества.

Лемма 10.1.

a) $\varphi_n(t; x) + (-1)^n \varphi_n^*(x; t) = \Omega_n(t; x)$,

$\varphi_N(t; x) - \varphi_N(x; t) = \Omega_N(t; x)$;

b) $\displaystyle\int_a^b \varphi_n(t; x)\, \varphi_n^*(x; s)\, dx = \varphi_N(t; s)$,

c) $\displaystyle\Omega_N(t; x) = \sum_{k=0}^{N} (-1)^{k+1} u_{N-k}(x)\, u_k(t)$.

Доказательство. Чтобы доказать первое приведенное в пункте а) утверждение, заметим, что для $t > x$ $\varphi_n^*(x; t) \equiv 0$, и тогда, конечно, $\varphi_n(t; x) + (-1)^n \varphi_n^*(x, t) = \varphi_n(t; x) = \Omega_n(t; x)$. Для $t < x$ имеем $\varphi_n(t; x) = 0$, и перестановка интегралов показывает, что $(-1)^n \varphi_n^*(x; t) = \Omega_n(t; x)$. Вторая формула из (а) есть специальный случай; заметим при этом, что $\varphi_N^*(x; t) = \varphi_N(x; t)$.

При многократном выполнении интегрирования по частям с использованием тождества (2.11), согласно которому

$$\frac{d}{dx}\varphi_n(t; x) = -w_n(x)\varphi_{n-1}(t; x),$$

получаем, что $\int\limits_a^b \varphi_n(t; x)\varphi_n^*(x; s)\,dx = \varphi_N(t; s)$.

Тождество в с) доказывается следующим образом. Непосредственное дифференцирование $\Omega_N(t; x)$ показывает, что для фиксированного $x \in [a, b]$ $d[D_1^* \ldots D_n^* D_n \ldots D_0 \Omega_N(t; x)] = 0$. Следовательно, по лемме 2.3 функция $\Omega_N(t; x)$ представляет собой многочлен в системе функций $\{u_i\}_0^N$, т. е.

$$\Omega_N(t; x) = \sum_{k=0}^{N} c_k(x)\,u_k(t).$$

Коэффициенты $c_k(x)$ вычисляются по известным значениям функции $\Omega_N(t; x)$ и ее производных при $t = a$. Например, используя тождества пункта а), получим

$$c_0(x) = \sum_{k=0}^{N} c_k(x)\,u_k(a) = \Omega_N(a; x) = -\varphi_N(x; a) = -u_N(x) = -u_{2n+1}(x),$$

$$c_1(x) = \left.\frac{D_0\Omega_N(t; x)}{w_1(t)}\right|_{t=a} = \left.\frac{-D_0\varphi_N(x; t)}{w_1(t)}\right|_{t=a} = u_{2n}(x).$$

Остальные коэффициенты $c_k(x)$ определяются подобным же образом.

Лемма 10.2. *Пусть $A_\mu$, $\mu = 1, \ldots, k$, представляют собой константы, удовлетворяющие условию*

$$\sum_\mu A_\mu u_i(x_\mu) = 0, \qquad i = 0, 1, \ldots, n.$$

*Тогда*

$$g(t) = \sum_\mu A_\mu\Big[\varphi_N(x_\mu; t) - \sum_\nu L_\nu(t)\,\varphi_N(x_\mu; x_\nu)\Big] = 0, \qquad t \in [a, b].$$

Доказательство. Функция $g(t)$, очевидно, обращается в нуль при $t = x_1, \ldots, x_k$. Согласно теореме 9.1 достаточно доказать, что $g(t)$ — натуральный сплайн. Так как $L_\nu(t)$, $\nu = 1, \ldots, k$, —

натуральные сплайны, то второй член в выражении для $g(t)$ тоже является натуральным сплайном. Кроме того, $\varphi_N(x_\mu; t) = 0$ для $t > x_k$. Следовательно, достаточно показать, что $(D_n D_{n-1} \ldots \ldots D_0) \sum_\mu A_\mu \varphi_N(x_\mu; t) = 0$ для $t < x_1$. Теперь по лемме 10.1 a)

$$\sum_\mu A_\mu \varphi_N(x_\mu; t) = \sum_\mu A_\mu [\varphi_N(t; x_\mu) + \Omega_N(x_\mu; t)] =$$

$$= \sum_\mu A_\mu \Omega_N(x_\mu; t), \qquad t < x_1.$$

Из леммы 10.1 c) следует, что

$$\sum_\mu A_\mu \Omega_N(x_\mu; t) = \sum_\mu A_\mu \sum_{i=0}^{\mu} (-1)^{i+1} u_{N-i}(t) u_i(x_\mu) =$$

$$= \sum_{i=0}^{N} (-1)^{i+1} u_{N-i}(t) \sum_\mu A_\mu u_i(x_\mu) =$$

$$= \sum_{i=n+1}^{N} (-1)^{i+1} \left[\sum_\mu A_\mu u_i(x_\mu)\right] u_{N-i}(t),$$

так как $\sum_\mu A_\mu u_i(x_\mu)$ обращается в нуль для $i = 0, 1, 2, \ldots, n$ по условию леммы. Последнее выражение, очевидно, представляет собой многочлен по $u_0, \ldots, u_n$ и, следовательно,

$$D_n D_{n-1} \ldots D_0 \left(\sum_\mu A_\mu \varphi_N(x_\mu; t)\right) = 0, \qquad t < x_1.$$

Следствие 10.1. *Существуют константы* $\lambda'_i$, $i = 0, 1, \ldots, n$, *такие, что*

$$\mathcal{L}\left(\varphi_N(x_\mu; t) - \sum_{\nu=1}^{k} L_\nu(t) \varphi_N(x_\mu; x_\nu)\right) = \sum_{i=0}^{n} \lambda_i u_i(x_\mu), \qquad \mu = 1, 2, \ldots, k.$$

Доказательство. Рассмотрим матрицу $\mathbf{M}$ размером $(n+1) \times k$ с вектор-строками

$$(u_i(x_1), \ldots, u_i(x_k)), \qquad i = 0, 1, \ldots, n.$$

Обращаясь к лемме 10.2, видим, что вектор с компонентами

$$\mathcal{L}(\varphi_N(x_\mu; t) - \sum_\nu L_\nu(t) \varphi_N(x_\mu x_\nu), \qquad \mu = 1, \ldots, k,$$

принадлежит к $\mathcal{N}(\mathbf{M})^{\perp}$ — ортогональному дополнению нуль-пространства матрицы $\mathbf{M}$. Однако для любой матрицы $\mathcal{N}(\mathbf{M})^{\perp} = \mathcal{R}(\mathbf{M}')$, где $\mathbf{M}'$ — транспонированная матрица. откуда и следует требуемый результат.

Доказательство теоремы 10.1. Так как

$$\mathcal{R}_t\varphi_n(t; x) = \mathcal{L}_t\varphi_n(t; x) - \sum_{\nu} B_\nu\varphi_n(x_\nu; x),$$

то надо показать, что $B_\nu^* = \mathcal{L}L_\nu$, $\nu = 1, \ldots, k$, однозначно минимизирует интеграл

$$\int_a^b \left[\mathcal{L}_t\varphi_n(t; x) - \sum_{\nu} B_\nu\varphi_n(x_\nu; x)\right]^2 dx \tag{10.8}$$

при ограничениях (10.2). Используя метод множителей Лагранжа, продифференцируем выражение

$$\int_a^b \left[\mathcal{L}_t\varphi_n(t; x) - \sum_{\nu} B_\nu\varphi_n(x_\nu; x)\right]^2 dx + \sum_i \lambda_i \left[\sum_{\nu} B_\nu u_i(x_\nu) - \mathcal{L}u_i\right]$$

относительно неизвестных $B_\mu$ и $\lambda_i$, мы получаем следующие соотношения:

$$-2\int_a^b \left[\mathcal{L}_t\varphi_n(t; x) - \sum_{\nu} B_\nu\varphi_n(x_\nu; x)\right]\varphi_n(x_\mu; x)\,dx +$$

$$+ \sum_i \lambda_i u_i(x_\mu) = 0, \qquad \mu = 1, \ldots, k, \tag{10.9}$$

$$\sum_{\nu} B_\nu u_i(x_\nu) = \mathcal{L}u_i, \qquad i = 0, 1, \ldots, n. \tag{10.10}$$

Доказательство теперь делится на две части, а именно, надо доказать, что: a) выбор постоянных $B_\nu^* = \mathcal{L}L_\nu$, $\nu = 1, \ldots, k$, удовлетворяет соотношениям (10.9) и (10.10) при соответствующих $\{\lambda_i^*\}_0^n$ и b) система из $n+1+k$ уравнений (10.9) и (10.10) относительно неизвестных $\{B_\nu\}_1^k$ и $\{\lambda_i\}_0^n$ имеет единственное решение.

Докажем пункт a). Используя лемму 10.1 a), запишем

$$\varphi_n(t; x)\,\varphi_n(x_\mu; x) = \varphi_n(x_\mu; x)\left[(-1)^{n+1}\varphi_n^*(x; t) + \Omega_n(t; x)\right] =$$

$$= \varphi_n(x_\mu; x)\left[(-1)^{n+1}\varphi^*(x; t) + \sum_{i=0}^{n} c_i(x)\,u_i(t)\right].$$

Интеграл от этого выражения равен

$$\int_a^b \varphi_n(t;x)\,\varphi_n(x_\mu;x)\,dx =$$

$$= (-1)^{n+1}\int_a^b \varphi_n(x_\mu;x)\,\varphi_n^*(x;t)\,dx + \sum_{i=0}^n b_i(x_\mu)\,u_i(t),$$

где $b_i(x_\mu) = \int_a^b c_i(x)\,\varphi_n(x_\mu;x)\,dx.$

Следовательно, по лемме 10.1 b) получаем

$$\int_a^b \varphi_n(t;x)\,\varphi_n(x_\mu;x)\,dx =$$

$$= (-1)^{n+1}\varphi_N(x_\mu;t) + \sum_{i=0}^n b_i(x_\mu)\,u_i(t), \qquad (10.11)$$

откуда

$$\int_a^b \mathscr{L}_t\varphi_n(t;x)\,\varphi_n(x_\mu;x)\,dx =$$

$$= (-1)^{n+1}\mathscr{L}_t\varphi_N(x_\mu;t) + \sum_i b_i(x_\mu)\,\mathscr{L}u_i. \qquad (10.12)$$

Подставив $t = x_\nu$ в (10.11), далее получим

$$\sum_\nu B_\nu \int_a^b \varphi_n(x_\nu;x)\,\varphi_n(x_\mu;x)\,dx =$$

$$= (-1)^{n+1}\sum_\nu B_\nu\varphi_N(x_\mu;x_\nu) + \sum_i b_i(x_\mu)\sum_\nu B_\nu u_i(x_\nu). \qquad (10.13)$$

Используя (10.12) и (10.13), уравнения (10.9) можно переписать в виде

$$-2\Big[(-1)^{n+1}\Big(\mathscr{L}_t\varphi_n(x_\mu;t) - \sum_\nu B_\nu\varphi_N(x_\mu;x_\nu)\Big) +$$

$$+ \sum_i b_i(x_\mu)\Big(\mathscr{L}u_i - \sum_\nu B_\nu u_i(x_\nu)\Big)\Big] + \sum_i \lambda_i u_i(x_\mu) = 0, \qquad (10.14)$$

$$\mu = 1, 2, \ldots, k.$$

Теперь при специальном выборе $B_\nu = B_\nu^* = \mathscr{L}L_\nu$ имеем

$$\sum_i b_i(x_\mu)\Big[\mathscr{L}u_i - \sum_\nu (\mathscr{L}L_\nu)\,u_i(x_\nu)\Big] =$$

$$= \sum_i b_i(x_\mu)\Big[\mathscr{L}\Big(u_i(t) - \sum_\nu L_\nu(t)\,u_i(x_\nu)\Big)\Big].$$

Так как $L_\nu(x_\mu) = \delta_{\mu\nu}$, то видно, что $u_i(t) - \sum L_\nu(t) u_i(x_\nu) = R_i(t)$ обращается в нуль при $t = x_1, \dots, x_k$. Но эта функция есть натуральный сплайн, и из теоремы 9.1 следует, что $R_i(t)$ тождественно равна нулю. Тогда (10.14) сводится к следующему:

$$2(-1)^{n+2}\left[\mathcal{L}_t\left[\varphi_N(x_\mu; t) - \sum_\nu L_\nu(t)\varphi_N(x_\mu; x_\nu)\right]\right] + \\ + \sum_i \lambda_i u_i(x_\mu) = 0, \qquad \mu = 1, 2, \dots, k. \tag{10.15}$$

Теперь, используя следствие 10.1, можно записать:

$$\mathcal{L}\left(\varphi_N(x_\mu; t) - \sum_\nu L_\nu(t)\varphi_N(x_\mu; x_\nu)\right) = \sum_{i=0}^{n} \lambda_i' u_i(x_\mu), \qquad \mu = 1, \dots, k.$$

Очевидно, что условия (10.15) удовлетворяются при

$$\lambda_i^* = \lambda_i = 2(-1)^{n+1} \cdot \lambda_i'.$$

Мы показали, что уравнения (10.9) и (10.10) имеют решение $B_\nu^* = \mathcal{L}L_\nu$ при соответствующим образом выбранных $\lambda_i$, $i = 0, \dots, n$; таким образом, первая часть доказательства завершена.

Теперь докажем, что уравнения (10.9) и (10.10) имеют единственное решение. Ясно, что для этого достаточно показать, что система уравнений

$$\sum_\nu B_\nu \int_a^b \varphi_n(x_\nu; x)\varphi_n(x_\mu; x)\,dx + \\ + \sum_i \lambda_i u_i(x_\mu) = 0, \qquad \mu = 1, \dots, k, \tag{10.16}$$

$$\sum_\nu B_\nu u_i(x_\nu) = 0, \qquad i = 1, \dots, n, \tag{10.17}$$

имеет только тривиальное решение $B_\nu = 0$, $\nu = 1, \dots, k$; $\lambda_i = 0$, $i = 0, \dots, n$.

Сначала заметим, что, согласно (10.16), функция

$$g(t) = \sum_\nu B_\nu \int_a^b \varphi_n(x_\nu; x)\varphi_n(t; x)\,dx + \sum_{i=0}^{n} \lambda_i u_i(t) \tag{10.18}$$

обращается в нуль при $t = x_1, x_2, \dots, x_k$. Применим операторы $D_0, D_1, \dots, D_n$ к функции $g(t)$ и воспользуемся теоремой Ролля. Так как $D_n D_{n-1} \dots D_0 u_i = 0$, $i = 0, 1, \dots, n$, и

$$D_n D_{n-1} \dots D_0 \left(\int_a^b \varphi_n(x_\nu; x)\varphi_n(t; x)\,dx\right) = \varphi_n(x_\nu; t), \tag{10.19}$$

то можно сделать вывод о существовании различных значений $\{x_j^*\}_{j=1}^{k-n-1}$, таких, что

$$\sum_\nu B_\nu \varphi_n(x_\nu; t) = 0 \qquad (10.20)$$

для $t = x_j$, $j = 1, \ldots, k-n-1$, где $x_j < x_j^* < x_{j+n+1}$, $j = 1, \ldots$ $\ldots, k-n-1$.

Далее мы покажем, что

$$\sum_\nu B_\nu \varphi_n(x_\nu; t) = 0$$

для всех $t < x_1$. Для каждого фиксированного $t_0 < x_1$ из леммы 2.3 следует, что $\varphi_n(t; t_0)$ совпадает с многочленом по функциям $u_0, \ldots, u_n$ для $t > t_0$; таким образом,

$$\varphi_n(t; t_0) = \sum_{j=0}^{n} b_j(t_0)\, u_j(t).$$

Используя равенство (10.17), можно сделать вывод, что для произвольного $t_0 < x_1$

$$\sum_\nu B_\nu \varphi_n(x_\nu; t_0) = \sum_j b_j(t_0) \sum_\nu B_\nu u_j(x_\nu) = 0.$$

Пусть $\bar{x}_1, \ldots, \bar{x}_{n+1}$ — любые фиксированные точки, удовлетворяющие условию $\bar{x}_1 < \ldots < \bar{x}_{n+1} < x_1$. Тогда

$$\sum_{\nu=1}^{k} B_\nu \varphi_n(x_\nu; t) = 0$$

при $k$ значениях $\bar{x}_1 < \ldots < \bar{x}_{n+1} < x_1^* < \ldots < x_{k-n-1}^*$. По лемме 9.2 определитель этой системы уравнений с $k$ неизвестными $B_1, \ldots, B_k$ положителен, и, следовательно, $B_1 = \ldots = B_k = 0$.

Уравнения (10.16) теперь сводятся к $\sum_{i=0}^{n} \lambda_i u_i(x_\mu) = 0$, $\mu = 1, \ldots, k$. Но $k \geq n+1$, а $\{u_i\}_0^n$ — $T$-система, из чего можно заключить, что $\lambda_0 = \lambda_1 = \ldots = \lambda_n = 0$. Этим завершается доказательство.

## § 11. Свойства дифференцируемости выпуклых в обобщенном смысле функций

Этот параграф посвящен рассмотрению доказательств свойств гладкости для выпуклых в обобщенном смысле функций, описанных в § 2. Для простоты изложения имеет смысл рассмотреть вспомогательный аппарат в виде ряда лемм.

Все рассматриваемые нами функции определены на фиксированном открытом интервале $(a, b)$. Говорят, что функция является ограниченной, если она ограничена на любом замкнутом подынтервале. Будем пользоваться сокращениями:

$$D^{-1} \text{ — тождественный оператор,}$$

$$D^k = D_k D_{k-1} \ldots D_1 D_0 \qquad (k = 0, 1, \ldots),$$

$$(D_i f)(x) = \frac{d}{dx} \frac{f(x)}{w_i(x)},$$

и пусть функция $\varphi_i(y; x)$ определена соотношением (2.5).

Следующие леммы понадобятся в дальнейшем, но могут представлять и самостоятельный интерес.

Лемма 11.1. *Предположим, что функция $g(x) = [D^{k-1}\varphi(x)/w_k(x)]$ абсолютно непрерывна на $(a, b)$ и имеет ограниченную производную почти всюду. Допустим также, что*

$$\lim_{y \downarrow x} \frac{\varphi(y) - \sum_{i=0}^{k} \varphi_i(y; x) \dfrac{D^{i-1}\varphi(x)}{w_i(x)}}{\int_x^y \varphi_k(y; \xi)\, d\xi} = L_k(x; \varphi)$$

*существует*[1]) *на $(a, b)$. Если $D^k\varphi$ существует при $x_0$, то $D^k\varphi(x_0) = L_k(x_0; \varphi)$.*

Обозначения. Величина

$$\frac{\varphi(y) - \sum_{i=0}^{k} \varphi_i(y; x) \dfrac{D^{i-1}\varphi(x)}{w_i(x)}}{\int_x^y \varphi_k(y; \xi)\, d\xi} = L_k(y, x; \varphi), \qquad y > x, \tag{11.1}$$

встречается часто, и поэтому удобно ввести для нее обозначение. Числитель в выражении (11.1) иногда встречается отдельно, поэтому обозначим

$$\varphi(y) - \sum_{i=0}^{k} \varphi_i(y; x) \frac{D^{i-1}\varphi(x)}{w_i(x)} = M_k(y, x; \varphi). \tag{11.2}$$

Доказательство. Так как функция $g(x)$ является абсолютно непрерывной, то $D^k\varphi(x)$ существует почти всюду. Поэтому можно осуществить неоднократное интегрирование по частям, которое

[1]) $\lim\limits_{y \downarrow x}$ означает, что предел берется при приближении $y$ к $x$ справа.

приводит к тождеству

$$\int_x^y \varphi_k(y, \xi) D^k\varphi(\xi)\, d\xi = M_k(y, x; \varphi).$$

(ср. с (2.10)). В действительности первое интегрирование по частям приводит к равенству

$$\int_x^y \varphi_k(y; \xi) D^k\varphi(\xi)\, d\xi =$$
$$= \int_0^{y-x} \varphi_{k-1}(y; x+\eta)\, w_k(x+\eta)[g(x+\eta) - g(x)]\, d\eta.$$

Справедливо следующее тождество:

$$\int_0^{y-x} \varphi_{k-1}(y; x+\eta)\, w_k(x+\eta)[g(x+\eta) - g(x)]\, d\eta = M_k(y, x; \varphi). \tag{11.3}$$

Кроме того, имеем

$$\int_0^{y-x} \eta\varphi_{k-1}(y; x+\eta)\, w_k(x+\eta)\, d\eta = \int_x^y \varphi_k(y; \xi)\, d\xi.$$

Теперь пусть $x_0$ является точкой, в которой существует предел $\lim\limits_{\eta\downarrow 0} [g(x_0+\eta) - g(x_0)]/\eta = D^k\varphi(x_0)$. С помощью тождества (11.3) можно получить соотношение

$$\frac{\displaystyle\int_0^{y-x_0} \eta\varphi_{k-1}(y; x_0+\eta)\, w_k(x_0+\eta)\left\{\frac{g(x_0+\eta) - g(x)}{\eta} - L_k(x_0, \varphi)\right\} d\eta}{\displaystyle\int_0^{y-x_0} \eta\varphi_{k-1}(y; x_0+\eta)\, w_k(x_0+\eta)\, d\eta} =$$
$$= L_k(y, x_0; \varphi) - L_k(x_0; \varphi). \tag{11.4}$$

Выражение в правой части стремится к нулю по предположению и

$$\frac{g(x_0+\eta) - g(x_0)}{\eta} - L_k(x_0; \varphi) \to D^k\varphi(x_0) - L_k(x_0; \varphi) \text{ при } \eta \downarrow 0$$

за счет выбора $x_0$.

Ядро интегрального преобразования в формуле (11.4) является неотрицательным, и, используя свойства суммируемости, можно заключить, что $D^k\varphi(x_0) = L_k(x_0; \varphi)$, что и требовалось доказать.

Лемма 11.2. *Пусть условия леммы* 11.1 *имеют место, и, кроме того, допустим, что функция* $L_k(x; \varphi)$ *непрерывна. Тогда* $D^k\varphi(x)$ *всюду существует и совпадает с* $L_k(x; \varphi)$.

Доказательство. Так как $g(x)$ абсолютно непрерывна, то для всех $x$ и $\eta > 0$ имеем

$$\frac{1}{\eta}\int_{x}^{x+\eta} D^k\varphi(\xi)\,d\xi = \frac{1}{\eta}[g(x+\eta) - g(x)].$$

Но $D^k\varphi(\xi) = L_k(\xi; \varphi)$ почти всюду, так что

$$\frac{1}{\eta}\int_{x}^{x+\eta} L_k(\xi; \varphi)\,d\xi = \frac{1}{\eta}[g(x+\eta) - g(x)].$$

Так как $L_k(\xi; f)$ непрерывна, то из известной теоремы численного анализа следует, что выражение в левой части стремится (при $\eta \downarrow 0$) к $L_k(x; \varphi)$ для всех $x$.

Аналогичные рассуждения позволяют доказать следующее следствие.

Следствие 11.2. *Если предположить, что $L_k(x, \varphi)$ непрерывна справа, то $D_R^k\varphi(x)$ (производная справа от $g(x)$) существует всюду и непрерывна справа.*

Лемма 11.3. *Допустим, что $g(x) = [D^{k-1}\varphi(x)/w_k(x)]$ непрерывна на интервале $(a, b)$ и $M_k(y; x; \varphi) \geqslant 0$ для всех $a < x < y < b$. Тогда $g(x)$ — неубывающая функция.*

*Если же $M_k(y, x; \varphi) \leqslant 0$ для всех $x < y$, то $g(x)$ — невозрастающая функция.*

Доказательство. Докажем только первое утверждение. Тождество (11.3) в сочетании с предположением леммы означает, что

$$\int_{0}^{y-x} R(y; x+\eta)[g(x+\eta) - g(x)]\,d\eta \geqslant 0 \tag{11.5}$$

для всех $x < y$, где $R(y; x+\eta) = \varphi_{k-1}(y; x+\eta)\,w_k(x+\eta) > 0$ при $0 \leqslant \eta < y - x$. Предположим противное, т. е., что $g(x)$ не является неубывающей функцией. Тогда существуют $\alpha < \beta$ ($a < \alpha < \beta < b$) такие, что $g(\alpha) > g(\beta)$. Пусть $\gamma$ будет максимальной точкой из $[\alpha, \beta]$, где функция $g$ достигает своего максимума на $[\alpha, \beta]$. Очевидно, что $\gamma < \beta$. Выбрав $x = \gamma$ и $y = \beta$, имеем $g(\gamma+\eta) - g(\gamma) < 0$ для всех $0 < \eta < \beta - \gamma$, что, очевидно, вступает в противоречие с неравенством (11.5). Это противоречие и доказывает желаемый результат.

Лемма 11.4. *Пусть функция $g(x) = [D^{k-1}\varphi(x)/w_k(x)]$ является ограниченной и удовлетворяет свойству квазинепрерывности: $\overline{\lim\limits_{\xi \downarrow x}}\, g(\xi) = g(x)$ для всех $x$. Тогда из того, что $M_k(y, x; \varphi) \geqslant 0$ для всех $a < x < y < b$, следует, что $g(x)$ — неубывающая функция.*

Доказательство. Опять предположим противное, т. е. что $g(x)$ не является неубывающей функцией. Это означает, что существуют $\alpha < \beta$ такие, что $g(\alpha) > g(\beta)$. Пусть $m = \sup\limits_{x \in [\alpha, \beta]} g(x)$. Если $m$ достигается в некоторой точке, то мы приходим к противоречию с предыдущей леммой. Предположим, что $m$ нигде не достигается. Пусть $\gamma_0$ определена так, что $\sup\limits_{x \in [\gamma_0, \beta]} g(x) < m$, но $\sup\limits_{x \in [\gamma_0 - \delta, \beta]} g(x) = m$ для всех достаточно малых положительных $\delta$ (это возможно, так как $g$ квазинепрерывна справа). Ясно, что $\alpha < \gamma_0 \leq \beta$, так как мы предполагаем, что $m$ не достигается. Существует последовательность $\gamma_n \uparrow \gamma_0$ $(\alpha < \gamma_n \leq \beta)$ такая, что $g(\gamma_n) \uparrow m$ и $g(\xi) \leq m - \varepsilon$ для всех $\xi \in [\gamma_0, \beta]$ и некоторого $\varepsilon > 0$. Поскольку $g(\beta) < g(\alpha)$, то можно выбрать $\varepsilon$ таким образом, чтобы $g(\alpha) - \varepsilon > g(\beta)$. Из-за того, что функция $g(x)$ квазинепрерывна справа, существует малый интервал $[\beta, \beta + \delta]$ такой, что $g(\xi) \leq m - \varepsilon$ на $[\beta, \beta + \delta)$. Следовательно, $g(\xi) \leq m - \varepsilon$ на $[\gamma_0, \beta + \delta)$. Теперь зададим $x = \gamma_n$ для больших $n$ и положим $y = \beta + \delta$. Очевидно, для больших $n$ $g(\gamma_n) > g(\xi) + \varepsilon/2$ для всех $\xi \in [\gamma_0, \beta + \delta]$. Хотя $g(\xi) \leq m$ на $[\gamma_n, \gamma_0)$, можно взять $n$ таким большим и, следовательно, $\gamma_0 - \gamma_n$ таким малым, что вклад интеграла

$$\int\limits_0^{\gamma_0 - \gamma_n} R(y; \gamma_n + \eta)\,[g(\gamma_n + \eta) - g(\gamma_n)]\,d\eta$$

будет сколь угодно малым. С другой стороны,

$$\int\limits_{\gamma_0 - \gamma_n}^{\beta + \delta - \gamma_n} R(y; \gamma_n + \eta)\,[g(\gamma_n + \eta) - g(\gamma_n)]\,d\eta$$

является отрицательной фиксированной величиной. Это находится в противоречии с неотрицательностью выражения (11.5).

Подобным же образом можно установить следующую лемму.

Лемма 11.5. *Предположим, что функция $g(x) = (D^{k-1}\varphi(x))/w(x)$ ограничена и удовлетворяет свойству квазинепрерывности: $\overline{\lim\limits_{\xi \downarrow x}} g(\xi) = g(x)$ для всех $x$ в интервале $(a, b)$. Пусть для всех $a < x < y < b$ выполняется неравенство $M_k(y, x; \varphi) \leq 0$. Тогда функция $g(x)$ — невозрастающая.*

Теперь мы можем доказать следующую важную лемму.

Лемма 11.6. *Пусть существуют $\varphi$, $(D_0\varphi/w_1), \ldots, (D^{k-1}\varphi/w_k)$, и предположим, что функция $g(x) = [D^{k-1}\varphi(x)/w_k(x)]$ удовлетворяет условиям квазинепрерывности, введенным в леммах 11.4 и 11.5 (два условия квазинепрерывности вместе означают, что функция $g(x)$ непрерывна справа). Предположим, что $L_k(x, y; \varphi)$ ограничена при $x < y$ (т. е. ограничена, когда $x$ и $y$ $(x < y)$ принадлежат любому замкнутому подынтервалу интервала $(a, b)$). Тогда*

*$g(x)$ имеет ограниченные разделенные разности. В частности, $g(x)$ абсолютно непрерывна на $(a, b)$.*

Доказательство. Заменим $\varphi$ на $\tilde{\varphi} = \varphi + C\theta_k(x, a)$, где $\theta_k(x, a)$ определяется как решение дифференциального уравнения $D^k\theta_k(x, a) \equiv 1$ при начальных условиях $\theta_k(a, a) = 0$ и $D^r\theta_k(a, a) = 0$, $r = 0, 1, \ldots, k-1$. Из определения $L_k(y, x; \varphi)$, данного в (11.1), найдем, что

$$L_k(y, x; \tilde{\varphi}) = L_k(y, x; \varphi) + C.$$

При соответствующем выборе $C$ можно гарантировать, что $L_k(y, x; \tilde{\varphi}) \geqslant 0$ для всех $a + \delta \leqslant x < y \leqslant b - \delta$, где $\delta$ — положительная и произвольно малая величина, а $C$, конечно и зависит от $\delta$. Применив лемму 11.4, заключаем, что $\tilde{g}(x) = [D^{k-1}\tilde{\varphi}(x)/w_k(x)]$ — неубывающая функция, что, в свою очередь, означает

$$g(y) - g(x) \geqslant -C(y - x). \tag{11.6}$$

Аналогично при рассмотрении $\varphi^* = \varphi - C_1\theta_k(x, a)$ и соответствующем $C_1$ находим, что $L_k(y, x; \varphi^*) \leqslant 0$ для всех $x$ и $y$ $(x < y)$, определенных на подынтервале интервала $(a, b)$. Применив лемму 11.5, приходим к выводу, что $g^* = (D^{k-1}\varphi^*)/w_k$ — убывающая функция, или

$$g(y) - g(x) \leqslant C_1(y - x). \tag{11.7}$$

Соотношения (11.6) и (11.7) показывают, что функция

$$g(x) = D^{k-1}\varphi(x)/w_k(x)$$

имеет ограниченные разделенные разности на любом подынтервале. Доказательство леммы завершено.

Следствие 11.6. *Пусть функция $g(x) = [D^{k-1}\varphi(x)/w_k(x)]$ удовлетворяет условиям леммы 11.6, за исключением того, что $L_k(y; x; \varphi)$ предполагается ограниченной только снизу на любом замкнутом подынтервале интервала $(a, b)$. Тогда $g(x)$ дифференцируема почти всюду.*

Доказательство. Анализ доказательства предыдущей леммы показывает, что $g(x) + cx$ — неубывающая функция при соответствующем $c$, и, следовательно, $g$ — дифференцируемая почти всюду функция.

Теперь мы располагаем необходимыми результатами, чтобы доказать упомянутые свойства регулярности. В дальнейшем будем считать, что $\varphi \in C(u_0, u_1, u_2, \ldots, u_n)$. Предположим, что уже доказано следующее:

$$D^r\varphi \in C(a, b), \quad r < n - 2, \quad n \geqslant 2. \tag{11.8}$$

(Для удобства определим $D^{-2}\varphi \equiv 0$ и вспомним, согласно нашим предыдущим обозначениям, что $D^{-1}\varphi=\varphi$, $D^0\varphi=D_0\varphi$ и т. д.) Наша следующая задача состоит в том, чтобы доказать, что функция $D^{r+1}\varphi$ существует и непрерывна. Для $r<n-1$ выберем

$$x<y<z<x_{r+5}<x_{r+6}<\ldots<x_{n+2},$$

где $x_{r+5}<\ldots<x_{n+2}$ выбираются вблизи $b$, а $x, y, z$ можно рассматривать как три произвольные точки внутри $(a, b)$, расположенные в возрастающем порядке; если $r+2=n$, то переменные, соответствующие $x_{r+5}, x_{r+6}, \ldots$, отсутствуют.

Теперь составим определитель

$$\begin{vmatrix} u_0(x), & D^0u_0(x), & \ldots, & D^ru_0(x), & u_0(y), & u_0(z), & u_0(x_{r+5}), & \ldots, & u_0(x_{n+2}) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ u_n(x), & D^0u_n(x), & \ldots, & D^ru_n(x), & u_n(y), & u_n(z), & u_n(x_{r+5}), & \ldots, & u_n(x_{n+2}) \\ \varphi(x), & D^0\varphi(x), & \ldots, & D^r\varphi(x), & \varphi(y), & \varphi(z), & \varphi(x_{r+5}), & \ldots, & \varphi(x_{n+2}) \end{vmatrix} \tag{11.9}$$

который неотрицателен, так как $\varphi\in C(u_0, u_1, \ldots, u_n)$. Вспомним, что

$$L_{r+1}(y, x; \theta)=\frac{\theta(y)-\sum\limits_{i=0}^{r+1}\varphi_i(y; x)\dfrac{D^{i-1}\theta(x)}{w_i(x)}}{\int\limits_x^y \varphi_{r+1}(y;\xi)\,d\xi},$$

откуда получаем неравенство

$$\begin{vmatrix} u_0(x), & \ldots, & D^ru_0(x), & L_{r+1}(y, x; u_0), & L_{r+1}(z, x; u_0), & u_0(x_{r+5}), & \ldots, & u_0(x_{n+2}) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ u_n(x), & \ldots, & D^ru_n(x), & L_{r+1}(y, x; u_n), & L_{r+1}(z, x; u_n), & u_n(x_{r+5}), & \ldots, & u_n(x_{n+2}) \\ \varphi(x), & \ldots, & D^r\varphi(x), & L_{r+1}(y, x; \varphi), & L_{r+1}(z, x; \varphi), & \varphi(x_{r+5}), & \ldots, & \varphi(x_{n+2}) \end{vmatrix} \geqslant$$

$$\geqslant 0. \tag{11.10}$$

Раскрывая теперь определитель по элементам последней строки, приходим к неравенству

$$(-1)^{n+r+1}L_{r+1}(y, x; \varphi)\,D_r(x, z, x_{r+5}, \ldots, x_{n+2})+$$
$$+(-1)^{n+r}L_{r+1}(z, x; \varphi)\,D_r(x, y, x_{r+5}, \ldots, x_{n+2})+$$
$$+E(x, y, z, x_{r+5}, \ldots)\geqslant 0, \tag{11.11}$$

где $D_r(x, z, x_{r+5}, \ldots)$ — минор элемента $L_{r+1}(y, x; \varphi)$.

Заметим, что, когда $y \downarrow x$, то

$$D_r(x, y, x_{r+5}, \ldots) \to$$

$$\to \begin{vmatrix} u_0(x), \ldots D^r u_0(x), & D^{r+1}u_0(x), & u_0(x_{r+5}), \ldots, u_0(x_{n+2}) \\ u_1(x), \ldots D^r u_1(x), & D^{r+1}u_1(x), & u_1(x_{r+5}), \ldots, u_1(x_{n+2}) \\ \cdots & \cdots & \cdots \\ u_n(x), \ldots D^r u_2(x), & D^{r+1}u_n(x), & u_n(x_{r+5}), \ldots, u_n(x_{n+2}) \end{vmatrix} = \widetilde{D}(x). \qquad (11.12)$$

Правая часть в выражении (11.12) строго положительна, так как $\{u_i\}_0^n$ — обобщенная полная $T$-система. Кроме того,

$$\lim_{z \downarrow x} D_r(x, z, x_{r+5}, \ldots) = \widetilde{D}(x) > 0.$$

Наконец, следует указать, что

$$E(x, y, z, x_{r+5}, \ldots) \to 0,$$

когда $y \downarrow x$ и $z \downarrow x$, так как после перехода к пределам $E$ превращается в сумму определителей, каждый из которых содержит два равных столбца.

В случае, когда $r < n-2$, можно также составить определитель

$$\begin{vmatrix} u_0(x_1) & u_0(x) & D^0 u_0(x) & \ldots D^r u_0(x) & u_0(y) & u_0(z) \ldots & u_0(x_{n+2}) \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ u_n(x_1) & u_n(x) & D^0 u_n(x) & \ldots D^r u_n(x) & u_n(y) & u_n(z) \ldots & u_n(x_{n+2}) \\ \varphi(x_1) & \varphi(x) & D^0\varphi(x) & \ldots D^r\varphi(x) & \varphi(y) & \varphi(z) \ldots & \varphi(x_{n+2}) \end{vmatrix} \geqslant 0, \qquad (11.13)$$

где $x_1 < x < y < z < x_{r+6} < \ldots < x_{n+2}$. Величина $x_1$ задается вблизи $a$, а $x_{r+6}, \ldots, x_{n+2}$ выбираются вблизи $b$, и для наших целей $x < y < z$ можно считать произвольными внутри интервала $(a, b)$. Основное различие между (11.9) и (11.13) заключается в том, что в последнем выражении столбцы с переменной $x$ начинаются со второго.

Выполнив очевидные преобразования определителя (11.13), приходим к следующему неравенству:

$$\begin{vmatrix} u_0(x_1) & u_0(x), \ldots, D^r u_0(x), & L_{r+1}(x, y; u_0), & L_{r+1}(x, z; u_0), \ldots, & u_0(x_{n+2}) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ u_n(x_1) & u_n(x), \ldots, D^r u_n(x), & L_{r+1}(x, y; u_n), & L_{r+1}(x, z; u_n), \ldots, & u_n(x_{n+2}) \\ \varphi(x_1) & \varphi(x), \ldots, D^r\varphi(x), & L_{r+1}(x, y; \varphi), & L_{r+1}(x, z; \varphi), \ldots, & \varphi(x_{n+2}) \end{vmatrix} \geqslant 0. \qquad (11.14)$$

Разложив теперь определитель (11.14) по элементам последней строки, приходим к неравенству

$$(-1)^{n+r} L_{r+1}(y, x; \varphi) D_r^*(x_1, x, z, x_{r+6}, \ldots) +$$
$$+ (-1)^{n+r+1} L_{r+1}(z, x; \varphi) D_r^*(x_1, x, y, \ldots) + E^*(x_1, x, y, z \ldots) \geqslant 0. \qquad (11.15)$$

Снова заметим, что, когда $y \downarrow x$, то

$$D_r^*(x_1, x, y, x_{r+6}, \ldots) \to$$

$$\to \begin{vmatrix} u_0(x_1), & u_0(x), & \ldots, & D^r u_0(x), & D^{r+1} u_0(x), & u_0(x_{r+6}), & \ldots, & u_0(x_{n+2}) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ u_n(x_1), & u_n(x), & \ldots, & D^r u_n(x), & D^{r+1} u_n(x), & u_n(x_{r+6}), & \ldots, & u_n(x_{n+2}) \end{vmatrix}, \tag{11.16}$$

и предельная величина является строго положительной вследствие обобщенного чебышевского свойства системы функций $\{u_0, u_1, \ldots \ldots, u_n\}$. Кроме того, $D_r^*(x_1, x, z, x_{r+6}, \ldots)$ имеет тот же самый предел при $z \downarrow x$, а $E^*(\ldots)$ стремится к нулю, когда $y \downarrow x$ и $z \downarrow x$.

Сравнивая (11.11) и (11.15) (существенное различие между ними состоит в знаке этих двух неравенств), легко получаем, что существует $C$, при котором

$$L_{r+1}(y, x, \varphi) \leqslant C, \qquad x < y, x, y \in (a, b). \tag{11.17}$$

Пусть $y \downarrow x$ и $z \downarrow x$ тогда из неравенства (11.11) с учетом того, что $D_r(x)$ в (11.11) строго положительна, и $E \to 0$, заключаем, что

$$\varlimsup_{y \downarrow x} (-1)^{n+r} L_{r+1}(y, x; \varphi) \leqslant \varliminf_{z \downarrow x} (-1)^{n+r} L_{r+1}(z, x, \varphi). \tag{11.18}$$

Это означает, что

$$\lim_{y \downarrow x} L_{r+1}(y, x; \varphi) = L_{r+1}(x; \varphi)$$

существует для всех $x$.

Далее возьмем $z > x_0$. Пусть $y \downarrow x$, $x \to x_0$ и $z \downarrow x_0$. Тогда получим, что

$$\varlimsup_{x \to x_0} L_{r+1}(x, \varphi) \leqslant L_{r+1}(x_0, \varphi), \tag{11.19}$$

если $n + r + 1$ нечетное, и

$$\varliminf_{x \to x_0} L_{r+1}(x, \varphi) \geqslant L_{r+1}(x_0, \varphi),$$

если $n + r + 1$ четное. Применяя те же рассуждения, основанные на неравенстве (11.15), приходим к следующему результату:

$$\varliminf_{x \to x_0} L_{r+1}(x; \varphi) \geqslant L_{r+1}(x_0; \varphi), \tag{11.20}$$

если $n + r + 1$ нечетное и

$$\varlimsup_{x \to x_0} L_{r+1}(x; \varphi) \leqslant L_{r+1}(x_0; \varphi),$$

если $n + r + 1$ четное.

Сравнение (11.19) и (11.20) показывает, что функция $L_{r+1}(x; \varphi)$ непрерывна по $x$.

Учитывая это, а также (11.17) и леммы 11.2 — 11.4, получим утверждение части а) следующей теоремы.

Теорема 11.1. *Пусть* $\varphi \in C(u_0, u_1, \ldots, u_n)$, *где* $\{u_i\}_0^n$ — *ECT-система на* $(a, b)$.

(a) *Предположим, что функции* $\varphi, D^0\varphi, D^1\varphi, \ldots, D^r\varphi$ *существуют и непрерывны, где* $r < n-2$ $(r = -1, 0, 1, 2, \ldots)$. (*Условие для* $r = -1$ *соответствует требованию непрерывности функции* $\varphi$.) *Тогда* $D^{r+1}\varphi(x)$ *существует и непрерывна.*

b) *Пусть* $\varphi, D^0\varphi, \ldots, D^{n-2}\varphi$ *непрерывны на* $(a, b)$, *тогда* $D^{n-2}\varphi(x)/w_{n-1}(x)$ *имеет производную справа* $D_R^{n-1}\varphi$, *которая непрерывна справа и* $D_R^{n-1}\varphi(x)/w_n(x)$ *возрастает на* $(a, b)$.

Доказательство части b). Воспользуемся неравенством

$$\begin{vmatrix} u_0(x), & D^0u_0(x), & \ldots, & D^{n-2}u_0(x), & u_0(y), & u_0(z) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ u_n(x), & D^0u_n(x), & \ldots, & D^{n-2}u_n(x), & u_n(y), & u_n(z) \\ \varphi(x), & D^0\varphi(x), & \ldots, & D^{n-2}\varphi(x), & \varphi(y), & \varphi(z) \end{vmatrix} \geqslant 0 \qquad (11.21)$$

для $x < y < z$, откуда получаем

$$\begin{vmatrix} u_0(x), & 0, & 0, & \ldots, & 0, & 0, & 0 \\ u_1(x), & w_1(x), & 0, & \ldots, & 0, & 0, & 0 \\ u_2(x), & D^0u_2(x), & w_2(x), & \ldots, & 0, & 0, & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & w_{n-1}(x), & 0, & 0 \\ u_n(x), & D^0u_n(x), & & \ldots, & D^{n-2}u_n(x), & X_n(y, x), & X_n(z, x) \\ \varphi(x), & D^0\varphi(x), & D^1\varphi(x), & \ldots, & D^{n-2}\varphi(x), & L_{n-1}(y, x; \varphi), & L_{n-1}(z, x; \varphi) \end{vmatrix} \geqslant 0,$$

где

$$X_n(y, x) = \frac{\int_x^y \varphi_{n-1}(y; \xi) w_n(\xi)\, d\xi}{\int_x^y \varphi_{n-1}(y; \xi)\, d\xi} = L_{n-1}(y, x; u_n).$$

Мы используем здесь тот факт, что

$$L_{n-1}(y, x; \theta) = \int_x^y \varphi_{n-1}(y; \xi) D^{n-1}\theta(\xi)\, d\xi \Big/ \int_x^y \varphi_{n-1}(y; \xi)\, d\xi \qquad (11.22)$$

для $\theta \in C^{n-1}$. Это выражение равно нулю, когда $\theta = u_0, u_1, \ldots, u_{n-1}$, так как $D^{n-1}u_k \equiv 0$ $(k = 0, 1, \ldots, n-1)$.

Раскрывая определитель по элементам последнего столбца, получим

$$L_{n-1}(z, x; \varphi) X_n(y; x) \geqslant L_{n-1}(y, x; \varphi) X_n(z, x). \qquad (11.23)$$

Очевидно, $X_n(y, x) \to w_n(x)$, когда $y \downarrow x$.

Аналогично тому, как было установлено неравенство (11.18), в данном случае можно установить, что

$$\lim_{y \downarrow x} L_{n-1}(y, x; \varphi) = L_{n-1}(x; \varphi). \tag{11.24}$$

Специальное рассуждение показывает, что функция $L_{n-1}(y, x; \varphi)$ равномерно ограничена, когда $x$ и $y\,(x < y)$ лежат внутри интервала $(a, b)$. Действительно, из (11.23) следует, что $L_{n-1}(y, x; \varphi)$ ограничена сверху. Раскрыв определитель в неравенстве

$$\begin{vmatrix} u_0(x_1), & u_0(x), & \ldots, & D^{n-2}u_0(x), & L_{n-1}(y, x; u_0) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ u_n(x_1), & u_n(x), & \ldots, & D^{n-2}u_n(x), & L_{n-1}(y, x; u_n) \\ \varphi(x_1), & \varphi(x), & \ldots, & D^{n-2}\varphi(x), & L_{n-1}(y, x; \varphi) \end{vmatrix} \geqslant 0 \tag{11.25}$$

$(x_1 < x < y)$, получим нижнюю границу для $L_{n-1}(y, x; \varphi)$. Отсюда, согласно лемме 11.6, следует, что функция $D^{n-2}\varphi(x)/w_{n-1}(x)$ абсолютно непрерывна и имеет ограниченные разделенные разности.

Теперь в неравенстве (11.23) будем считать, что $z > x_0$, пусть $y \downarrow x$, $x \to x_0$ и $z \downarrow x_0$.

Имеем

$$\overline{\lim_{x \to x_0}} L_{n-1}(x; \varphi) \leqslant L_{n-1}(x_0; \varphi) \tag{11.26}$$

для всех $x_0 \in (a, b)$. Из леммы 11.1 следует, что в любой точке, где производная

$$g(x) = \frac{d}{dx}\left(\frac{D^{n-2}\varphi(x)}{w_{n-1}(x)}\right)$$

существует, выполняется равенство $g(x) = L_{n-1}(x; \varphi)$. Доказательство этого утверждения основывается на тождестве

$$\frac{\int\limits_{x_0}^{y} \varphi_{n-1}(y; \xi)\, g(\xi)\, d\xi}{\int\limits_{x_0}^{y} \varphi_{n-1}(y; \xi)\, d\xi} = L_{n-1}(y, x; \varphi). \tag{11.27}$$

Из этого представления легко следует также, что

$$\overline{\lim_{x \downarrow x_0}} L_{n-1}(x, \varphi) \geqslant L_{n-1}(x_0; \varphi).$$

Действительно, из выражения (11.27) видно, что существует по крайней мере одно значение $x_1\,(x_0 < x_1 < y)$, при котором $g(x_1) = L_{n-1}(x_1; \varphi) \geqslant L_{n-1}(y, x_0; \varphi) - \varepsilon$, где $\varepsilon$ — фиксированное произвольное положительное число. Далее рассмотрим формулу (11.27), где $y$ заменено на $(x_0 + x_1)/2$. Снова из представления (11.27) заключаем, что существует такое $x_2\,(x_0 < x_2 < x_1)$, что

$$L_{n-1}(x_2; \varphi) \geqslant L_{n-1}(x_1, x_0; \varphi) - \varepsilon.$$

Продолжая, получим последовательность $x_1 \geqslant x_2 \geqslant x_3 \geqslant \ldots \geqslant x_0$, удовлетворяющую условию

$$L_{n-1}(x_k;\varphi) \geqslant L_{n-1}(x_{k-1}, x_0;\varphi) - \varepsilon, \qquad k = 2, 3, \ldots$$

Отсюда следует, что

$$\overline{\lim_{x \downarrow x_0}} L_{n-1}(x;\varphi) \geqslant L_{n-1}(x_0;\varphi) - \varepsilon.$$

Но $\varepsilon > 0$ — произвольное, и поэтому

$$\overline{\lim_{x \downarrow x_0}} L_{n-1}(x;\varphi) \geqslant L_{n-1}(x_0;\varphi),$$

что в сочетании с неравенством (11.26) приводит к следующему:

$$\overline{\lim_{x \downarrow x_0}} L_{n-1}(x;\varphi) = L_{n-1}(x_0;\varphi). \tag{11.28}$$

Наконец, справедливо неравенство

$$\begin{vmatrix} u_0(x), & 0, & \ldots, & 0, & 0, & u_0(y) \\ u_1(x), & w_1(x), & \ldots, & 0, & 0, & u_1(y) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ u_n(x), & D^0 u_n(x), & \ldots, & D^{n-2}u_n(x), & w_n(x), & u_n(y) \\ \varphi(x), & D^0\varphi(x), & \ldots, & D^{n-2}\varphi(x), & L_{n-1}(x;\varphi), & \varphi(y) \end{vmatrix} \geqslant 0 \tag{11.29}$$

для всех $x < y$, так как $\varphi \in C(u_0, u_1, \ldots, u_n)$.

Пусть

$$K_n(y, x;\theta) = \frac{\theta(y) - \sum_{i=0}^{n-1} \varphi_i(y;x)\,\frac{D^{i-1}\theta(x)}{w_i(x)} - \frac{\varphi_n(y;x)}{w_n(x)}\,L_{n-1}(x;\theta)}{\int\limits_x^y \varphi_n(y;\xi)d\xi} \tag{11.30}$$

Заметим, что $L_{n-1}(x;\theta) = (D^{n-1}\theta(x))$, когда $\theta$ принадлежит классу $C^n$. Теперь, сформировав функцию $K$ в последнем столбце определителя (11.29), получим

$$\begin{vmatrix} u_0(x), & 0, & \ldots, & 0, & 0 \\ u_1(x) & w_1(x), & \ldots, & 0, & 0 \\ u_n(x), & D^0u_n(x), & \ldots, & w_n(x), & 0 \\ \varphi(x), & D^0\varphi(x), & \ldots, & L_{n-1}(x;\varphi), & K_n(x, y;\varphi) \end{vmatrix} \geqslant 0,$$

или, эквивалентно,

$$K_n(x, y;\varphi) \geqslant 0 \tag{11.31}$$

для всех $a < x < y < b$.

Используя леммы 11.5 и 11.6, заключаем, что функция $L_{n-1}(x;\varphi)/w_n(x)$ монотонно возрастающая. Отсюда и из равенства (11.28) следует, что функция $L_{n-1}(x;\varphi) = D_R^{n-1}\varphi(x)$ непрерывна справа и неубывающая. Теорема 11.1 доказана.

Глава XII

# ОБОБЩЕННЫЕ НЕРАВЕНСТВА ЧЕБЫШЕВА

## § 1. Введение

В этой главе и в заключительных двух главах книги мы начинаем систематическое исследование нескольких родов обобщенных неравенств Чебышева. В простейших формах неравенства чебышевского типа определяют верхнюю или нижнюю границу для интеграла от некоторой функции или функционала над некоторым классом мер.

В гл. III нам встретилось специфически чебышевское неравенство в теореме Маркова—Крейна. Мы воспроизведем этот результат с целью сравнения. Пусть $\{u_i\}_0^n$ есть $T$-система на $[a, b]$ и $\Omega$ — заданная непрерывная функция. Требуется определить точные верхнюю и нижнюю границы для интеграла

$$\int_a^\xi \Omega(t)\, d\sigma(t), \tag{1.1}$$

где $\sigma$ пробегает множество неотрицательных конечных мер, удовлетворяющих моментным условиям

$$c_i^0 = \int_a^b u_i(t)\, d\sigma(t), \qquad i = 0, 1, \ldots, n. \tag{1.2}$$

При некоторых дополнительных ограничениях на функции $u_0, \ldots, u_n$ и $\Omega$ были установлены точные границы для (1.1) и было определено, что соответствующие экстремальные меры есть канонические меры. В решении этой проблемы существенно использовались предположения о том, что системы $\{u_i\}_0^k$ и $\{u_0, u_1, \ldots, u_k, \Omega\}$ для каждого $k = 0, 1, \ldots, n$ образуют $T$-системы и что $\Omega(t) > 0$, $t \in [a, b]$.

В этой главе мы исследуем неравенства чебышевского типа при значительно более общих условиях, накладываемых на функции $u_0, u_1, \ldots, u_k$ и $\Omega$. Стоящая перед нами задача — определить общие границы для $\int_a^b \Omega(t)\,d\sigma(t)$, где $\sigma$ подчинено моментным условиям вида (1.2) и другим естественным ограничениям гладкости.

Наиболее раннее из неравенств чебышевского типа приписывают Чебышеву (см. Шохат и Тамаркин [1943]). Результат Чебышева теперь рассматривают как специальный случай неравенств Маркова — Крейна. Вклад в эту область внесли Гаусс, Валле-Пуссен и др. В нашем столетии выросла многотомная литература, посвященная неравенствам чебышевского типа. Обзорный очерк результатов, полученных вплоть до 1955 года, содержится в работе Гудвина [1955]. Другие более специальные работы собраны Севиджем [1961]. Мы также обращаем внимание на монографию Гудвина [1964], которая частично совпадает с материалом следующих трех глав. Как обзорные статьи, так и монография Гудвина содержат важнейшую библиографию, к которой мы отсылаем заинтересованного читателя. Неравенства чебышевского типа представляют интерес в теории вероятностей, статистическом анализе, теории моментных неравенств, дифференциальных уравнениях, теории интерполяции и других областях.

Формулировку, описание и анализ обобщенных чебышевских неравенств удобно проводить на языке случайных величин и математических ожиданий. Пусть $(\mathcal{X}, \mathcal{B}, P)$ — вероятностное пространство, т. е. $\mathcal{X}$ — абстрактное пространство, $\mathcal{B}$ — $\sigma$-поле подмножеств $\mathcal{X}$, и $P$ — вероятностная мера, определенная на $\mathcal{B}$; т. е. мера всего пространства есть единица. Случайная величина $X$ определяется как вещественная измеримая функция, заданная на $\mathcal{X}$. Для большинства задач достаточно рассматривать соответствующую функцию распределения

$$F(t) = \Pr(X \leqslant t) = P(\{x \in \mathcal{X} \mid X(x) \leqslant t\}). \tag{1.3}$$

Функция $F(t) = F_X(t)$ непрерывна справа, не убывает и удовлетворяет условиям $\lim\limits_{t\to\infty} F(t) = F(\infty) = 1$, $\lim\limits_{t\to-\infty} F(t) = F(-\infty) = 0$. Среднее и дисперсия случайной величины $X$ определяются соответственно как $\mu = EX = \int_{-\infty}^{\infty} t\,dF(t)$ и $\sigma^2 = E(X-\mu)^2 = \int_{-\infty}^{\infty} (t-\mu)^2 dF(t)$, где интегралы понимаются в смысле Лебега — Стилтьеса.

Наиболее классическое из всех чебышевских неравенств есть специальное неравенство Маркова, которое утверждает, что $\mu/\lambda$ есть наименьшая верхняя граница для $\Pr(X \geqslant \lambda)$ по отношению к классу всех неотрицательных случайных величин $X$, которые имеют среднее $\mu$ $(\mu < \lambda)$, и что эта верхняя граница достигается, только если $X$ удовлетворяет равенствам $\Pr(X = 0) = 1 - \Pr(X = \lambda) = 1 - \mu/\lambda$. Этот простой результат доказывается следующим образом.

Так как

$$\mu = \int_0^\infty t\,dF(t) \geqslant \int_\lambda^\infty t\,dF(t) \geqslant \lambda \Pr(X \geqslant \lambda), \tag{1.4}$$

то немедленно получаем

$$\Pr(X \geqslant \lambda) \leqslant \mu/\lambda. \tag{1.5}$$

Более того, равенство всюду в (1.4) появляется тогда и только тогда, когда $X$ принимает только значения 0 и $\lambda$ и $\Pr(x = \lambda) = \mu/\lambda$.

Неравенство Маркова есть специальный случай общей теоремы Маркова—Крейна (см. теорему 5.2 гл. V). Отметим, что в обозначениях гл. V предельное распределение есть каноническая мера индекса 3/2, которая сосредоточивает массы в точках 0 и $\lambda$.

Из неравенства Маркова вытекают несколько стандартных неравенств. Например, если $X$ — случайная величина со средним $\mu$ и дисперсией $\sigma^2$, то неравенство Бьеньеме—Чебышева утверждает, что

$$\Pr(|X - \mu| \geqslant \lambda\sigma) \leqslant 1/\lambda^2. \tag{1.6}$$

Этот результат получается заменой в (1.5) $X$ на $(X-\mu)^2/\sigma^2$. Другое классическое неравенство, которое может быть получено немедленно из (1.5), есть неравенство Пирсона, утверждающее, что если $X$ есть случайная величина со средним $\mu$ и $E|X-\mu|^r = \beta_r$ $(r > 0)$, то

$$\Pr(|X - \mu|^r \geqslant \beta_r\lambda^r) \leqslant 1/\lambda^r. \tag{1.7}$$

Многие неравенства чебышевского типа следуют из более фундаментальных посредством замены переменных и тем не менее приводятся в различных работах как самостоятельные результаты.

Неравенство Маркова является прототипом большой коллекции неравенств, каждое из которых утверждает, что некоторое число $A$ есть верхняя граница для $\Pr(\mathbf{X} \in E)$ по отношению к классу векторнозначных случайных величин $\mathbf{X} = (X_1, \ldots, X_k)$, удовлетворяющих некоторым моментным условиям вида

$$\int_T \cdots \int t_1^{r_1} t_2^{r_2} \ldots t_k^{r_k}\, dF_k(t_1, t_2, \ldots, t_k) = c_{r_1, r_2, \ldots, r_k}. \tag{1.8}$$

Множество $E$ есть некоторое подмножество заданной области $T$, которое может и не содержаться целиком в $k$-мерном евклидовом пространстве.

В общем случае верхняя граница $A$ зависит по крайней мере от некоторых моментных констант $c_{r_1, r_2, \ldots, r_k}$. Довольно часто множество $E$ предполагается симметричным относительно заданного вектора средних как в случае неравенства Бьеньеме — Чебышева.

Другой род неравенств представляется классическим неравенством Гаусса (см. ниже § 4), утверждающим, в частности, что

$4\sigma^2/9\lambda^2$ есть верхняя грань $\Pr\{|X| \geqslant \lambda\}$, взятая по классу вещественных случайных величин $X$, которые имеют среднее 0 и дисперсию $\sigma^2$ (при $\lambda \geqslant (2/\sqrt{3})\,\sigma$) и унимодальны с модой в точке 0 (т. е. производная от $F_X(t)$ существует и неотрицательна для $t < 0$ и неположительна для $t > 0$). Мы определяем здесь класс случайных величин в терминах условий «гладкости» и моментных условий. Обобщения понятия унимодальности будут даны в § 5.

Упомянем еще два типа неравенств. Первый тип требует границ некоторого вида для случайной величины: либо явно в виде требования, чтобы она менялась в пределах некоторого ограниченного множества, либо неявно, посредством задания бесконечного числа моментных неравенств вида $\int\limits_T |t|^i\, dF_X(t) \leqslant c_i$, $i = 1, 2, 3, \ldots$ Второй вид неравенств устанавливается после наложения моментных условий, которые включают условные математические ожидания.

Не существует единого метода вывода неравенств чебышевского типа. В § 2 мы опишем общий принцип (теорема 2.1), который дает возможность систематически выводить чебышевские неравенства. То, что этот принцип имеет широкую сферу приложения, будет многократно продемонстрировано в оставшейся части книги. Остальные параграфы этой главы содержат ряд примеров, в которых множество $T$ из (1.8) есть подмножество вещественной прямой. В каждом случае мы будем применять метод теоремы 2.1. Большинство этих случаев являются классическими и первоначально анализировались *прямыми* методами.

В гл. XIII мы применяем метод теоремы 2.1 к изучению многомерных чебышевских неравенств. В гл. XIV трактуются некоторые аспекты чебышевских неравенств, связанные с суммами случайных величин и некоторыми нелинейными задачами.

## § 2. Общая теорема

Пусть $u_0, u_1, \ldots, u_n$ — вещественные, измеримые по Борелю функции, определенные для $\mathbf{t} \in T$, где $T$ есть подмножество $k$-мерного евклидова пространства, и пусть $\Sigma$ обозначает совокупность всех конечных регулярных мер, определенных на $T$ и удовлетворяющих условиям интегрируемости $\int\limits_T |u_i(\mathbf{t})|\, d\sigma(\mathbf{t}) < \infty$, $i = 0, 1, \ldots, n$. Следует отметить, что в случае, когда $T$ есть подмножество вещественной прямой, функции $u_0, u_1, \ldots, u_n$ — общего вида и не должны образовывать чебышевскую систему.

Пусть для любого заданного $\mathbf{c}^0 = (c_0^0, c_1^0, \ldots, c_n^0)$

$$V(\mathbf{c}^0) = \left\{\sigma \mid \sigma \in \Sigma,\ \int\limits_T u_i(\mathbf{t})\, d\sigma(\mathbf{t}) = c_i^0, \qquad i = 0, 1, \ldots, n\right\}. \quad (2.1)$$

Задача состоит в определении точной верхней и нижней границ интеграла $\int\limits_T \Omega(\mathbf{t})\,d\sigma(\mathbf{t})$ при $\sigma \in V(\mathbf{c}^0)$, т. е. мы хотим оценить

$$\begin{aligned} I_{\max} &= \sup_{\sigma \in V(\mathbf{c}^0)} \int\limits_T \Omega(\mathbf{t})\,d\sigma(\mathbf{t}), \\ I_{\min} &= \inf_{\sigma \in V(\mathbf{c}^0)} \int\limits_T \Omega(\mathbf{t})\,d\sigma(\mathbf{t}), \end{aligned} \tag{2.2}$$

Если $V(\mathbf{c}^0)$ пусто, то условимся считать, что $I_{\max} = +\infty$, $I_{\min} = -\infty$. Даже если $V(\mathbf{c}^0)$ не пусто, экстремальные значения могут все же быть бесконечными. В большинстве примеров мы исследуем также вопрос о том, могут ли sup и inf в (2.2) быть заменены на max и min соответственно. Это обстоятельство второстепенно, так как почти во всех приложениях имеет значение именно неравенство.

Далее мы накладываем существенное минимальное требование линейной независимости функций $u_0, u_1, \ldots, u_n$ над $T$, т. е. если $\sum\limits_{i=0}^{n} a_i u_i(t) = 0$ для всех $t \in T$, то $a_i = 0$ $(i = 0, 1, \ldots, n)$.

Геометрическая интерпретация этого допущения такова. Рассмотрим пространство моментов

$$\mathcal{M}_{n+1} = \left\{ \mathbf{c} = (c_0, \ldots, c_n) \mid c_i = \int\limits_T u_i(\mathbf{t})\,d\sigma(\mathbf{t}),\ i = 0, 1, \ldots, n,\ \sigma \in \Sigma \right\}. \tag{2.3}$$

Если функции $u_0, u_1, \ldots, u_n$ линейно независимы, то конус $\mathcal{M}_{n+1}$ имеет размерность $n+1$, т. е. не существует гиперплоскости, содержащей $\mathcal{M}_{n+1}$.

В самом деле, допущение о том, что $\mathcal{M}_{n+1}$ содержится в гиперплоскости, влечет существование вещественных констант $\{a_i\}_0^n$, которые не все равны нулю и таковы, что

$$\sum_{i=0}^{n} a_i c_i + \gamma = 0 \text{ для всех } \mathbf{c} \in \mathcal{M}_{n+1}. \tag{2.4}$$

Из того, что $\mathcal{M}_{n+1}$ — конус, следует, что $\gamma = 0$. Таким образом, (2.4) принимает вид

$$\sum_{i=0}^{n} a_i u_i(\mathbf{t}) = 0 \text{ для всех } \mathbf{t} \in T$$

в противоречии с линейной независимостью. Наоборот, если $\{u_i\}_0^n$ линейно зависимы, мы получаем немедленно, что $\mathcal{M}_{n+1}$ заключено в $n$-мерном подпространстве.

Пространство моментов $\mathcal{M}_{n+1}$, образованное $\{u_i\}_0^n$ для произвольных $u_i$, не обладает богатой структурой в противоположность случаю, когда $u_0, u_1, \ldots, u_n$ образуют $T$-систему. Тем не менее формулировка и доказательство следующей теоремы основано на геометрической точке зрения, присущей изучению моментных пространств, порожденных $T$-системами функций $\{u_i\}_0^n$.

Для заданного $\Omega(\mathbf{t})$, определенного на $T$, мы вводим множество $\mathcal{P}_+$ всех многочленов, которые ограничивают $\Omega(\mathbf{t})$ сверху, т. е.

$$\mathcal{P}_+ = \left\{ u(\mathbf{t}) = \sum_{i=0}^{n} a_i u_i(\mathbf{t}) \,|\, u(\mathbf{t}) \geqslant \Omega(\mathbf{t}),\ \mathbf{t} \in T \right\}.$$

И аналогично

$$\mathcal{P}_- = \left\{ u(\mathbf{t}) = \sum_{i=0}^{n} a_i u_i(\mathbf{t}) \,|\, u(\mathbf{t}) \leqslant \Omega(\mathbf{t}),\ \mathbf{t} \in T \right\}.$$

Теперь мы можем сформулировать главную теорему настоящего параграфа.

Теорема 2.1. *Пусть* $\mathbf{c}^0$ *есть внутренняя точка* $\mathcal{M}_{n+1}$ *и пусть* $\Omega(\mathbf{t})$ *таково, что* $\mathcal{P}_+$ *и* $\mathcal{P}_-$ *непустые множества.*

*Тогда*

$$I_{\max} = \inf \sum_{i=0}^{n} a_i c_i^0, \tag{2.5}$$

$$I_{\min} = \sup \sum_{i=0}^{n} a_i c_i^0, \tag{2.6}$$

*где нижняя и верхняя грани берутся по всем многочленам* $u = \sum_{i=0}^{n} a_i u_i$, *содержащимся в* $\mathcal{P}_+$ *и* $\mathcal{P}_-$ *соответственно.*

*Более того, равенства в* (2.5) *и* (2.6) *достигаются.*

Замечание 2.1. Прежде чем приступить к доказательству, удобно в данном месте обсудить значение теоремы. Метод вывода чебышевских неравенств, заключенный в этой теореме, используется давно. Он, вероятно, восходит к Маркову, Чебышеву или, возможно, Поссе. Основной вклад и акцент теоремы 2.1 заключается в том, что неравенство, получаемое посредством данного метода, является точным. Ранее приходилось исследовать точность в каждом случае отдельно.

Общий принцип теоремы 2.1, по-видимому, был впервые сформулирован Исии [1960]. Независимо и одновременно, Карлин в 1960 г. (записи лекций в Станфорде) установил тот же самый результат.

Доказательство, основанное на геометрии пространства моментов, дается ниже. Специальные случаи теоремы 2.1 можно найти у Маршалла и Олкина [1960 a], [1960 b] (см. также ссылки в этих работах). Обзорная дискуссия по содержанию теоремы 2.1 дана у Кингмана [1963] и Исии [1960], [1963].

Один важный вопрос, не укладывающийся в общий контекст, касается единственности последовательности $\{a_i\}^n_{i=0}$, на которой реализуются экстремальные значения в (2.5) и (2.6). Информация такого рода полезна для облегчения подсчета $I_{\max}$ и $I_{\min}$.

Доказательство. Мы доказываем только (2.5), так как (2.6) либо доказывается подобными рассуждениями, либо получается из (2.5) заменой $\Omega$ на $-\Omega$, $u_i$ на $-u_i$ и т. д. Сначала отметим, что $I_{\max}$ конечно, так как $\mathcal{P}_+$ и $\mathcal{P}_-$ не являются пустыми множествами. Введем пространство моментов

$$\mathcal{M}=\left\{(c_0,\ldots,c_{n+1})\,|\,c_i=\int\limits_T u_i d\sigma,\ i=0,1,\ldots,n,\ c_{n+1}=\int\limits_T \Omega\,d\sigma,\ \sigma\in\Sigma\right\},$$

и пусть $\overline{\mathcal{M}}$ обозначает его замыкание. Ясно, что $\overline{\mathcal{M}}$ есть замкнутый выпуклый конус. Рассмотрим точку $\gamma^0=(c_0^0, c_1^0,\ldots,c_n^0, I_{\max})$, которая, очевидно, есть граничная точка $\overline{\mathcal{M}}$.

Построим опорную плоскость к конусу $\overline{\mathcal{M}}$ в точке $\gamma^0$. Существуют вещественные константы $\{b_i\}_0^{n+1}$, не все равные нулю и такие, что

$$\sum_{i=0}^{n+1} b_i c_i \geqslant 0 \tag{2.7}$$

для всех $\mathbf{c}\in\overline{\mathcal{M}}$ и

$$\sum_{i=0}^{n} b_i c_i^0 + b_{n+1} I_{\max} = 0.$$

В частности,

$$\sum_{i=0}^{n} b_i u_i(\mathbf{t}) + b_{n+1}\Omega(\mathbf{t}) \geqslant 0 \tag{2.8}$$

для всех $\mathbf{t}\in T$.

Докажем теперь, что $b_{n+1}<0$. В самом деле, точка $\mathbf{q}_\lambda=(c_0^0,\ldots$ $\ldots, c_n^0, I_{\max}+\lambda)$ для всех $\lambda>0$, очевидно, лежит в полупространстве, дополнительном к (2.7). Следовательно,

$$\sum_{i=0}^{n} b_i c_i^0 + b_{n+1} I_{\max} + \lambda b_{n+1} < 0$$

для всех $\lambda>0$. Отсюда, очевидно, следует соотношение $b_{n+1}\leqslant 0$. Если теперь $b_{n+1}=0$, то получаем противоречие с тем фактом, что $\{c_0^0,\ldots,c_n^0\}$ содержится в $\operatorname{Int}\mathcal{M}_{n+1}$. Отсюда следует, что $b_{n+1}$ отрицательно.

Разделив на $-b_{n+1}$, мы можем переписать (2.8) в виде

$$\sum_{i=0}^{n} a_i^0 u_i(\mathbf{t}) \geqslant \Omega(\mathbf{t}) \tag{2.9}$$

для всех $\mathbf{t}\in T$, где $a_i^0=b_i/(-b_{n+1})$ $(i=0,1,\ldots,n)$ и (2.7) принимает вид

$$\sum_{i=0}^{n} a_i^0 c_i^0 = I_{\max}. \tag{2.10}$$

Этим завершается доказательство (2.5).

Сформулируем несколько замечаний, формальное доказательство которых просто.

Замечание 2.2. Если $\mathbf{c}^0 = (c_0^0, \ldots, c_n^0) \in \mathcal{DM}_{n+1}$, $T$ есть подмножество вещественной прямой, а $\{u_i\}_0^n$ есть $T$-система, то $V(\mathbf{c}^0)$ содержит единственную меру $\sigma^*$ и, таким образом, $I_{\max} = I_{\min} = \int_T \Omega(\mathbf{t})\, d\sigma^*(\mathbf{t})$.

Замечание 2.3. В общем случае мы всегда имеем $\inf_{\mathbf{t}\in T}\left[\sum_{i=0}^{n} a_i^0 u_i(\mathbf{t}) - \Omega(\mathbf{t})\right] = 0$, где $\sum_{i=0}^{n} a_i^0 u_i(\mathbf{t})$ — многочлен, построенный в теореме. Более того, если существует экстремальная мера $\sigma^*$ такая, что $I_{\max} = \int_T \Omega\, d\sigma^*$, то спектр $\sigma^*$ содержится в множестве $S \subset T$, на котором в (2.9) достигается равенство, т. е.

$$S = \left\{\mathbf{t} \,\middle|\, \sum_{i=0}^{n} a_i^0 u_i(\mathbf{t}) = \Omega(\mathbf{t})\right\}. \tag{2.11}$$

В большинстве приложений, особенно когда $T$ — линейная область, множество $S$ содержит только конечное число точек и $n+2$ есть обычно подходящая верхняя граница этого числа. Кроме того, если $\{u_i\}_0^n$ — $T$-система, то, считая, что $\mathbf{c}^0$ — внутренняя точка, мы получаем из теоремы 2.1 гл. II, что множество $S$ содержит как минимум $[(n+1)/2]+1$ точек.

Замечание 2.4. Другой метод, часто используемый для вывода чебышевских неравенств, может быть описан следующим образом. Используя идеи замечания 2.3, мы можем иногда описать спектр экстремального распределения. Тогда для произвольного распределения на этих точках условия для моментов явно выписываются. Таким образом получаем уравнения, которые могут быть иногда легко решены при заданном спектре экстремального распределения. Отсюда получаем $a_i$ из теоремы 2.1, доставляющие минимум вместе с распределением, на котором достигается равенство.

## § 3. Примеры

В оставшейся части этой главы мы применяем метод теоремы 2.1 для получения ряда классических одномерных неравенств и нескольких обобщений. Первый пример излагается подробно, для того чтобы проиллюстрировать ход мыслей и приемы, которые помогают применять теорему 2.1. Искушенный читатель может посетовать на чрезмерную подробность изложения и предложить более краткие доказательства. Тем не менее нам кажется, что тонкие дополнительные рассуждения полезны, для того чтобы проиллюстрировать технику общего метода. В дальнейших примерах большинство стандартных алгебраических преобразований будут опущены; однако существенные различия в каждом случае будут ясно указаны.

Пример 3.1. Неравенство Селберга[1]). Мы хотим установить точную верхнюю границу вероятности

$$\Pr(X \in (-\infty, \mu-\alpha] \cup [\mu+\beta, \infty)), \qquad -\alpha < \beta, \tag{3.1}$$

для случайных величин $X$ со средним $\mu$ и дисперсией $\sigma^2 > 0$. Мы установим ниже, что максимум (3.1) равен единице, если $0 \leqslant -\alpha < \beta$ или $-\alpha < \beta \leqslant 0$. При $-\alpha < 0 < \beta$ решение разделяется на случаи, зависящие от соотношения величин $\alpha, \beta$ и $\sigma^2$, и представлено в табл. 1.

ТАБЛИЦА 1

$-\alpha < 0 < \beta, \quad \alpha_0 = \min\{\alpha, \beta\}$

| Случай | Условие | Максимум (3.1) | Спектр экстремального распределения |
|---|---|---|---|
| (i) | $\alpha\beta \leqslant \sigma^2$ | 1 | |
| (ii) | $\alpha\beta - \alpha_0^2 \leqslant 2\sigma^2 \leqslant 2\alpha\beta$ | $\dfrac{(\alpha-\beta)^2 + 4\sigma^2}{(\alpha+\beta)^2}$ | $\mu-\alpha, \mu+\beta$ и $\mu + 2^{-1}(\beta-\alpha)$ |
| (iii) | $2\sigma^2 \leqslant \alpha\beta - \alpha_0^2$ | $\dfrac{\sigma^2}{\sigma^2 + \alpha_0^2}$ | $\mu-\alpha$ и $\mu+\alpha^{-1}\sigma^2$, если $\beta \geqslant \alpha$; $\mu+\beta$ и $\mu-\beta^{-1}\sigma^2$, если $\beta \leqslant \alpha$ |

Замечание 3.1. Односторонние чебышевские неравенства. Неравенства

$$\Pr(X \leqslant \mu - \alpha) \leqslant \frac{\sigma^2}{\sigma^2 + \alpha^2}, \qquad \alpha > 0, \tag{3.2}$$

$$\Pr(X \leqslant \mu + \beta) \geqslant 1 - \frac{\sigma^2}{\sigma^2 + \beta^2}, \qquad \beta > 0, \tag{3.3}$$

возникают, если устремить $\beta \to \infty$ и $\alpha \to \infty$.

[1]) За всеми историческими дискуссиями и соответствующими ссылками мы направляем читателя к книге Гудвина [1964] и обзорным статьям Гудвина [1955] и Севиджа [1961].

Для доказательства результата Селберга мы можем принять, сдвигая случайную величину на $-\mu$, что $\mu = 0$, и тогда верхняя грань берется по всему классу случайных величин со средним 0 и дисперсией $\sigma^2$.

Мы сначала покажем, что максимум (3.1) есть единица, если $0 \leqslant -\alpha < \beta$ или $-\alpha < \beta \leqslant 0$. Так как эти условия симметричны, мы исследуем только первый случай. Чтобы установить, что максимум равен единице, достаточно указать распределение, сосредоточенное в двух точках с нулевым средним и дисперсией $\sigma^2$, не имеющее массы в открытом интервале $(-\alpha, \beta)$. Такое распределение существует при условии, что система уравнений

$$\begin{aligned} px + (1-p)\,y &= 0, \\ px^2 + (1-p)\,y^2 &= \sigma^2 \end{aligned} \tag{3.4}$$

имеет решение, удовлетворяющее условиям $x \leqslant -\alpha$, $y \geqslant \beta$ и $0 \leqslant p \leqslant 1$. Такое решение получаем, полагая для произвольного $y \geqslant \beta$ $x = -\sigma^2/y$ и $p = \sigma^2/(\sigma^2 + x^2)$.

Остается рассмотреть более интересный случай, когда $-\alpha < 0 < \beta$. Мы будем касаться только возможности $\beta > \alpha$, так как соответствующие заключения для $\beta < \alpha$ вытекают из очевидной симметрии задачи.

Решение разделяется на три случая:

(1) $\alpha\beta \leqslant \sigma^2$,

(2) $\alpha(\beta - \alpha) \leqslant 2\sigma^2 \leqslant 2\alpha\beta$,

(3) $2\sigma^2 \leqslant \alpha(\beta - \alpha)$.

Когда выполняется неравенство $\alpha\beta \leqslant \sigma^2$, система уравнений (3.4) допускает решение $y \in [\beta, \sigma^2/\alpha]$, $x = -\sigma^2/y \leqslant -\alpha$, $p = \sigma^2/(\sigma^2 + x^2)$. Максимум (3.1) снова равен единице.

Для случаев (2) и (3) мы применяем теорему 2.1. Возьмем $u_i(t) = t^i$; $i = 0, 1, 2$; $(c_0^0, c_1^0, c_2^0) = (1, 0, \sigma^2)$ и применим (2.5), где $\Omega(t)$ есть характеристическая функция $\Omega_E(t)$ множества $E = (-\infty, -\alpha] \cup [\beta, \infty)$. Искомый максимум (3.1) есть тогда минимальное значение

$$a\sigma^2 + c, \tag{3.5}$$

оцениваемое по отношению к множеству $\mathcal{P}_+$, всех многочленов, удовлетворяющих условию

$$Q(t) = at^2 + bt + c \geqslant \Omega_E(t). \tag{3.6}$$

Рис. 13 (см. также рис. 14) иллюстрирует функцию $\Omega_E(t)$ и типичный многочлен $Q(t)$, удовлетворяющий (3.6). Задача теперь состоит в том, чтобы «понизить» многочлен $Q(t)$ насколько это возможно при сохранении условия (3.6) до тех пор, пока минимум

$$\int_{-\infty}^{\infty} Q(t)\,dF = a\sigma^2 + c, \quad F \in V(1, 0, \sigma^2) \tag{3.7}$$

не будет достигнут.

Для получения нетривиального минимума (меньшего чем 1) значение $a = 0$ непригодно, поэтому мы можем сосредоточить внимание на многочленах $Q(t)$ следующего вида:

$$Q(t) = \delta(t-\eta)^2 + \lambda \geqslant \Omega_E(t), \qquad \delta > 0, \tag{3.8}$$

и минимизировать выражение

$$\delta(\sigma^2 + \eta^2) + \lambda. \tag{3.9}$$

Если (3.8) выполняется, тогда необходимо $\delta, \lambda \geqslant 0$ и, так как $\Pr(X \in E) \leqslant 1$, мы можем принять $\lambda \leqslant 1$ и $-\alpha < \eta < \beta$.

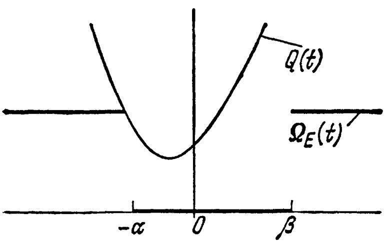



Рис. 13

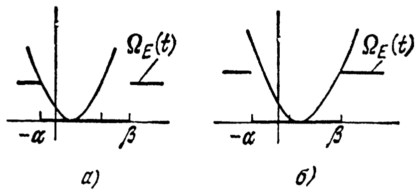



Рис 14.

Кроме того, чтобы получить нетривиальную границу, которая строго меньше единицы, мы потребуем, чтобы $\lambda = 0$. Чтобы доказать это утверждение, мы отмечаем, что теорема 2.1 гарантирует существование значений $\delta_0, \eta_0$ и $\lambda_0$, для которых (3.8) удовлетворяется и

$$I_{\max} = \sup_X \Pr(X \in E) = \delta_0(\sigma^2 + \eta_0^2) + \lambda_0, \tag{3.10}$$

где супремум берется по случайным величинам с нулевым средним и дисперсией $\sigma^2$.

Если $I_{\max} < 1$, тогда, очевидно, $\lambda_0 < 1$. Предположим, что $0 < \lambda_0 < 1$. Так как $\delta_0(t-\eta_0)^2 + \lambda_0 \geqslant \Omega_E(t)$ и $\Omega_E(t) = 0$ или 1, то отсюда следует, что

$$\frac{\delta_0(t-\eta_0)^2}{1-\lambda_0} \geqslant \Omega_E(t). \tag{3.11}$$

Однако если $I_{\max} < 1$, то $\delta_0(\sigma^2 + \eta_0^2) + \lambda_0 < 1$ и, следовательно,

$$\frac{\delta_0(\sigma^2 + \eta_0^2)}{1-\lambda_0} = \delta_0(\sigma^2 + \eta_0^2) + \frac{\lambda_0\delta_0(\sigma^2 + \eta_0^2)}{1-\lambda_0} \leqslant \delta_0(\sigma^2 + \eta_0^2) + \lambda_0.$$

Таким образом, если $\lambda_0 > 0$, то многочлен в (3.11) доставляет (3.9) меньшее значение, чем многочлен $\delta_0(t-\eta_0)^2+\lambda_0$. Следовательно, $\lambda_0 = 0$, как утверждалось.

Уменьшим значение $\delta$ так, чтобы функция $\delta(t-\eta)^2$ имела значение, равное единице, в точке $-\alpha$ или в точке $\beta$.

Теперь мы можем ограничить рассмотрение многочленами вида

$$\begin{aligned} &\frac{(t-\eta)^2}{(\eta+\alpha)^2}, \text{ где } -\alpha < \eta \leqslant \frac{\beta-\alpha}{2}, \\ &\frac{(t-\eta)^2}{(\eta-\beta)^2}, \text{ где } \frac{\beta-\alpha}{2} \leqslant \eta < \beta. \end{aligned} \tag{3.12}$$

Минимизирующий многочлен, следовательно, примет вид, показанный на рис. 14, *а*, *б*.

Если $(\beta-\alpha)/2 \leqslant \eta < \beta$, то $(\sigma^2+\eta^2)/(\eta-\beta)^2$ имеет минимальное значение $[(\alpha-\beta)^2+4\sigma^2]/(\alpha+\beta)^2$, которое достигается при $\eta = (\beta-\alpha)/2$.

При $-\alpha < \eta \leqslant (\beta-\alpha)/2$ функция $[\sigma^2+\eta^2]/(\eta+\alpha)^2$ убывает при $-\alpha < \eta < \sigma^2\alpha^{-1}$ и возрастает при $\eta \geqslant \sigma^2\alpha^{-1}$.

Пусть имеет место случай (2), т. е. $\alpha(\beta-\alpha) \leqslant 2\sigma^2 \leqslant 2\alpha\beta$.

Тогда минимум (3.5) для обоих типов многочленов, представленных в (3.12), достигается при $\eta = 2^{-1}(\beta-\alpha)$.

Минимум в этом случае равен $[(\alpha-\beta)^2+4\sigma^2]/(\alpha+\beta)^2$.

Если $\sigma^2\alpha^{-1} \leqslant 2^{-1}(\beta-\alpha)$, т. е. имеет место случай (3), то минимум (3.5) есть $\sigma^2(\sigma^2+\alpha^2)^{-1}$ и доставляется многочленом $(t-\eta)^2/(\eta+\alpha)^2$, где $\eta = \sigma^2/\alpha$.

Мы теперь рассмотрим вопрос о существовании распределений, достигающих верхней границы (3.1). Случаи $0 \leqslant -\alpha < \beta$, $-\alpha < \beta \leqslant 0$ и $\sigma^2 \geqslant \alpha\beta$ при $\alpha, \beta > 0$ уже рассмотрены.

Для оставшихся двух случаев экстремальные распределения суть следующие. В случае (2) мы помещаем массы $p_0 = [\beta(\beta-\alpha)+2\sigma^2]/(\beta+\alpha)^2$, $q_0 = 4(\beta\alpha-\sigma^2)/(\beta+\alpha)^2$ и $1-p_0-q_0$ в точки $-\alpha$, $(\beta-\alpha)/2$ и $\beta$ соответственно. Если имеет место случай (3), то мы помещаем массы $p_1 = \sigma^2(\alpha^2+\sigma^2)^{-1}$ и $1-p_1$ в точки $-\alpha$ и $\sigma^2\alpha^{-1}$ соответственно.

Пример 3.2. Неравенство Гутмана. Задача этого примера состоит в отыскании максимума

$$\Pr(X \in E) = \Pr(X \in (-\infty, \mu-\beta] \cup [\mu-\alpha, \mu+\alpha] \cup [\mu+\beta, \infty)), \quad 0 \leqslant \alpha < \beta, \tag{3.13}$$

по отношению к множеству всех случайных величин $X$ на $(-\infty, \infty)$ со средним $\mu$ и абсолютными моментами $E|X-\mu|^r = \mu_r$ и $E|X-\mu|^{2r} = \mu_{2r}$ ($r \geqslant 1$ фиксировано).

В отличие от примера 3.1, у нас теперь три интервала. Однако условия для моментов и симметрия множества $E$ позволяют нам свести эту задачу к задаче примера 3.1.

Максимум (3.13) равен единице, если $\mu_r \notin (\alpha^r, \beta^r)$, а если $\mu_r \in (\alpha^r, \beta^r)$ и $\alpha_0^2 = \min\{(\beta^r - \mu_r)^2, (\mu_r - \alpha^r)^2\}$, то решение представлено в табл. 2.

ТАБЛИЦА 2

$\mu_r \in (\alpha^r, \beta^r)$, $\alpha_0^2 = \min\{(\beta^r - \mu_r)^2, (\mu_r - \alpha^r)^2\}$

| Случай | Условие |
|---|---|
| (i) | $\mu_{2r} - \mu_r^2 \geqslant (\beta^r - \mu_r)(\mu_r - \alpha^r)$ |
| (ii) | $2^{-1}\{(\beta^r - \mu_r)(\mu_r - \alpha^r) - \alpha_0^2\} \leqslant (\mu_{2r} - \mu_r^2) \leqslant (\beta^r - \mu_r)(\mu_r - \alpha^r)$ |
| (iii) | $\mu_{2r} - \mu_r^2 \leqslant 2^{-1}\{(\beta^r - \mu_r)(\mu_r - \alpha^r) - \alpha_0^2\}$ |
| | **Максимум (3.13)** |
| (i) | 1 |
| (ii) | $\dfrac{[(\beta^r - \mu_r) - (\mu_r - \alpha^r)]^2 + 4(\mu_{2r} - \mu_r^2)}{[(\beta^r - \mu_r) + (\mu_r - \alpha^r)]^2}$ |
| (iii) | $\dfrac{\mu_{2r} - \mu_r^2}{(\mu_{2r} - \mu_r^2) + \alpha_0^2}$ |
| | **Спектр экстремального распределения** |
| (i), (ii) | $\mu \pm \alpha,\quad \mu \pm \beta$ и $\mu \pm \left(\dfrac{\alpha^r + \beta^r}{2}\right)^{1/r}$ |
| (iii) | $\mu \pm \alpha$ и $\mu \pm \left(\dfrac{\mu_{2r} - \alpha^r \mu_r}{\beta^r - \mu_r}\right)^{1/r}$ |

Мы можем принять без потери общности, что $\mu = 0$. Кроме того, рассмотрим симметризованное распределение $F(t)$, определенное как

$$G(t) = \begin{cases} [F(t) + 1 - F(-t-0)]/2, & t \leqslant 0, \\ G(-t), & t > 0. \end{cases} \tag{3.14}$$

Очевидно, что $F(t)$ и $G(t)$ дают одно и то же значение для $\Pr(X \in E)$ и что $F(t)$ и $G(t)$ имеют одни и те же абсолютные моменты. Задача сводится к нахождению максимума $\Pr(X \in E \cap [0, \infty))$ для неотрицательных случайных величин $X$, удовлетворяющих

$EX^r = \mu_r$ и $EX^{2r} = \mu_{2r}$. Наконец, переходя от случайной величины $X$ к случайной величине $Y = X^r$, получим, что задача эквивалентна максимизации

$$\Pr\{Y \in [0, \alpha^r] \cup [\beta^r, \infty)\} \tag{3.15}$$

по всем случайным величинам $Y$ таким, что $EY = \mu_r$ и $E(Y - \mu_r)^2 = \mu_{2r} - \mu_r^2$.

Принимая во внимание неравенство Селберга, мы видим, что экстремальное распределение всегда неотрицательно, если $\mu - \alpha \geqslant 0$ ($a > 0$), так что решение, представленное в табл. 1, также максимизирует (3.15) при условии, что мы заменим $\mu, \sigma^2$, $\alpha$ и $\beta$ на $\mu_r$, $\mu_{2r} - \mu_r^2$, $\mu_r - \alpha^r$ и $\beta^r - \mu_r$ соответственно.

Пример 3.3. Неравенство Глассера. Пусть $X$ — случайная величина на $(-\infty, \infty)$ с $EX = \mu$ и $E|X - \mu| = \delta > 0$. Неравенство Глессера утверждает, что

$$\Pr(-t_1\delta < X - \mu < t_2\delta) \geqslant \max\left\{0, 1 - \frac{1}{2}\left(\frac{1}{t_1} + \frac{1}{t_2}\right)\right\}. \tag{3.16}$$

Если $2^{-1}(t_1^{-1} + t_2^{-1}) \leqslant 1$, то нижняя граница достигается функцией распределения, которая сосредоточивает массы $(2t_1)^{-1}$ и $(2t_2)^{-1}$ в точках $\mu - t_1\delta$ и $\mu + t_2\delta$ соответственно, а оставшуюся часть в точке $\mu$.

Неравенство (3.16) эквивалентно максимизации

$$\Pr(Y \in (-\infty, -t_1] \cup [t_2, \infty)) \tag{3.17}$$

для случайных величин $Y$, удовлетворяющих условиям $EY = 0$ и $E|Y| = 1$.

Это есть задача, в которой ограничения на моменты включают функции 1, $t$ и $|t|$.

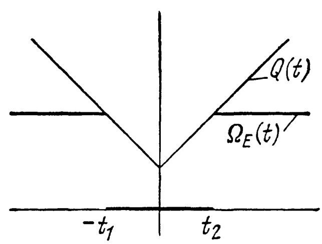


Рис. 15.

По теореме 2.1 максимум (3.17) равен минимуму $a + c$, взятому по всем многочленам $Q(t) = a|t| + bt + c$, которые удовлетворяют условию

$$Q(t) = a|t| + bt + c \geqslant \Omega_E(t), \tag{3.18}$$

где $\Omega_E$ есть характеристическая функция множества $E = (-\infty, -t_1] \cup [t_2, \infty)$.

Для того чтобы (3.18) выполнялось, мы должны иметь $Q(0) = c \geqslant 0$. Кроме того, так как $Q(t)$ линейна при $t \geqslant 0$ и $t \leqslant 0$ (рис. 15), мы можем принять, что $c \leqslant 1$; иначе $\int\limits_{-\infty}^{\infty} Q\,dF \geqslant 1$. Теперь для всякого $c \in [0, 1]$ мы можем всегда «понизить» линейные сегменты $Q(t)$, выбирая $a$ и $b$ такими, что $Q(-t_1) = Q(t_2) = 1$. Поэтому многочлены

$Q(t)$, которые следует рассмотреть, имеют вид

$$Q(t)=\frac{1-c}{2}\left(\frac{1}{t_1}+\frac{1}{t_2}\right)|t|+\frac{1-c}{2}\left(\frac{1}{t_2}-\frac{1}{t_1}\right)t+c.$$

Многочлен такого типа изображен на рис. 15. Минимум выражения

$$\int_{-\infty}^{\infty} Q(t)\,dF(t)=q+c(1-q),\quad q=\frac{1}{2}\left(\frac{1}{t_1}+\frac{1}{t_2}\right),$$

как легко видеть, равен $q$, если $q\leqslant 1$, и достигается при $c=0$. Когда $q\geqslant 1$, значение $c=1$ дает минимальное значение, равное единице.

## § 4. Чебышевские неравенства для классов унимодальных распределений

В этом параграфе мы изучаем несколько классов чебышевских неравенств, в которых функции распределения унимодальны. Под унимодальной функцией распределения $F$ с модой $m$ мы подразумеваем распределение вида

$$F(x)=\begin{cases}\displaystyle\int_{-\infty}^{x}\varphi(t)\,dt, & x<m,\\ \displaystyle F_m+\int_{-\infty}^{x}\varphi(t)\,dt, & x\geqslant m,\end{cases}\tag{4.1}$$

где функция $\varphi(t)$ не убывает при $t\leqslant m$ и не возрастает при $t\geqslant m$ и $F_m=F(m+)-F(m-)$. Вообще говоря, распределение $F$ имеет функцию плотности $\varphi$, у которой есть только один «пик» в точке $m$, и, следовательно, $\varphi(t)$ убывает при удалении $t$ от $m$.

Метод изучения задач, включающих унимодальные распределения, состоит в том, чтобы перейти от распределения $F$ к вспомогательной функции $H$ и соответственно использовать унимодальный характер $F$. Моментные ограничения на распределение $F$ эквивалентно выражаются как подходящие моментные условия на $H$. Ассоциированная задача в терминах функции $H$ будет попадать в сферу действия общего метода теоремы 2.1. Переход совершается простым использованием подходящей формы интегрирования по частям.

Пример 4.1. Неравенство Гаусса. В этом примере мы хотим оценить максимальное значение

$$\Pr\{X\in(-\infty,\mu-d]\cup[\mu+d,+\infty)\},\qquad d>0,\tag{4.2}$$

над классом унимодальных функций распределения с модой и средним, сосредоточенным в $\mu$, и дисперсией, равной $\sigma^2$.

ТАБЛИЦА 3

| Случай | Условие | Максимум | экстремальное распределение |
|---|---|---|---|
| (i) | $d^2 \leqslant \frac{4}{3}\sigma^2$ | $1-\frac{d}{\sqrt{3}\sigma}$ | прямоугольное на $[\mu-\sqrt{3}\sigma,\ \mu+\sqrt{3}\sigma]$ |
| (ii) | $d^2 \geqslant \frac{4}{3}\sigma^2$ | $-\frac{4}{9}\frac{\sigma^2}{d^2}$ | прямоугольное на $\left[\mu-\frac{3}{2}d,\ \mu+\frac{3}{2}d\right]$ с массой в $\mu$ |

Решение дано в табл. 3.

Замечание 4.1. Под прямоугольным распределением на $[a, b]$ мы подразумеваем распределение $F$, плотность которого есть

$$F'(t)=\begin{cases}\dfrac{1}{b-a}, & t\in[a, b],\\ 0, & t\notin[a, b].\end{cases}$$

Распределение может иметь «прямоугольную часть» на $[a, b]$. В этом случае константа $1/(b-a)$ соответственно уменьшается так, чтобы общая масса оставалась равной единице.

Замечание 4.2. Хотя этот специальный пример может трактоваться более непосредственно другими средствами (см., например, Барлоу и Маршалл [1964], стр. 1236), развиваемый ниже метод распространяется естественным образом на случай унимодальности высших порядков даже при наличии некоторых моментных ограничений (см. § 5). Метод конструктивен и опирается на подходящую форму интегрирования по частям для того, чтобы свести задачу к обычной задаче чебышевского неравенства со стандартными моментными ограничениями. Эта техника является весьма мощной (см. § 5) и, по-видимому, нова. Мы будем ее использовать при рассмотрении неравенства Гаусса, усеченного варианта этого неравенства, неравенства Роудена (пример 4.3) и некоторых обобщений работы Маллоу [1963].

В случае примеров 4.1 и 4.3 результаты можно эквивалентным образом интерпретировать как чебышевские неравенства для положительных случайных величин с убывающими плотностями.

При максимизации (4.2) мы можем принять без ограничения общности, что $\mu=0$ и $F$ симметрично (ср. метод примера 3.2). Следовательно, достаточно решить эквивалентную задачу нахождения максимума

$$\Pr\{X\in[d, \infty)\}, \tag{4.3}$$

где максимум берется по классу $\mathcal{F}$, состоящему из всех унимодальных распределений, удовлетворяющих условиям

$$F(0-)=0,\quad F(x)=F_0+\int_0^x \varphi(t)\,dt,\quad 0\leqslant x<\infty,\quad \int_0^\infty x^2\,dF(x)=\sigma^2 \tag{4.4}$$

($\varphi$ не убывает). Без потери общности мы принимаем, что $\varphi(t)$ непрерывна справа.

Рассмотрим сначала проблему максимизации (4.3) над подклассом $\mathcal{F}_0 \subset \mathcal{F}$ функций $F$, подчиненных дополнительному ограничению $F_0 = 0$.

Для всякого $F \in \mathcal{F}_0$ положим $dH(t) = -d\varphi(t)$. Тогда, интегрируя по частям, получаем

$$\int_0^\infty \varphi(t)\,dt = 1 = \int_0^\infty t\,dH(t),$$
$$\int_0^\infty t^2 \varphi(t)\,dt = \sigma^2 = \int_0^\infty \frac{t^3}{3}\,dH(t) \tag{4.5}$$

и

$$\int_d^\infty \varphi(t)\,dt = \int_d^\infty (t-d)\,dH(t).$$

Из теоремы 2.1 следует, что максимум (4.3) над классом $\mathcal{F}_0$ равен минимуму величины $a + 3b\sigma^2$, взятому по отношению ко всем многочленам $P(t)$, удовлетворяющим условию

$$P(t) = at + bt^3 \geqslant \psi(t),$$

где $\psi(t) = \begin{cases} 0, & 0 \leqslant t \leqslant d, \\ t-d, & t \geqslant d. \end{cases}$

Умножая многочлены на положительный масштабный множитель $\gamma \in (0, 1]$, мы отмечаем, что $a + 3b\sigma^2 \geqslant \gamma(a + 3b\sigma^2)$, так как обе части положительны. Отсюда следует, что мы можем сосредоточить внимание на тех многочленах $P(t)$, которые касаются ненулевой части $\psi(t)$, т. е. $P(\eta) = \eta - d$ для некоторого $\eta > d$. Для этих многочленов $P'(\eta) = 1$, и поэтому

$$a\eta + b\eta^3 = \eta - d, \quad a + 3b\eta^2 = 1,$$

или

$$a = 1 - \frac{3d}{2\eta}, \quad b = \frac{d}{2\eta^3}.$$

Для минимизации функции

$$g(\eta) = a + 3b\sigma^2 = \left(1 - \frac{3d}{2\eta}\right) + \frac{3d}{2\eta^3}\sigma^2 \tag{4.6}$$

мы находим, что ее производная отрицательна при $0 < \eta < \sqrt{3}\,\sigma$ и положительна при $\eta > \sqrt{3}\,\sigma$. Так как $P(t) \geqslant 0$ для малых $t$, мы должны иметь $a \geqslant 0$ или $\eta \geqslant (3/2)\,d$, так что минимум $a + 3b\sigma^2$

достигается при $\eta = \sqrt{3}\sigma$, если $4\sigma^2 \geqslant d^2$ и при $\eta = (3/2)d$, если $4\sigma^2 < 3d^2$. Эти значения дают решение задачи, рассматр ıваемой над классом $\mathcal{F}_0$.

Для того чтобы расширить значимость вышеупомянутых результатов на класс распределений, допускающих массу $F_0$ в нуле, мы видоизменим $F$ так, чтобы масса в нуле стала прямоугольным куском на $[0, \varepsilon]$ с весом $F_0/\varepsilon$. Затем мы применяем вышеупомянутые результаты и устремляем $\varepsilon$ к нулю.

Л гко можно проверить, что границы действительно достигаются, первая — прямоугольным распределением на $[\mu-\sqrt{3}\sigma, \mu+\sqrt{3}\sigma]$ и вторая — прямоугольным распределением на $[\mu - 3/2\,d, \mu + 3/2d]$ с массой в $\mu$. Для того чтобы получить эти распределения, заметим, что если мера $dH(t)$ достигает границы при ограничениях (4.5), то она должна сосредоточивать всю массу в точках 0 и $\sqrt{3}\sigma$, когда $4\sigma^2 \geqslant 3d^2$ и в точках 0 и $3d/2$, когда $4\sigma^2 < 3d^2$. Следовательно, мы заключаем из соотношения $dH(t) = -d\varphi(t)$, что функция распределения $F$ в (4.4), которая доставляет максимум, должна быть такой, что $\varphi(t)$ постоянна на интервале $[0, \sqrt{3}\sigma]$, когда $4\sigma^2 \geqslant 3d^2$, и постоянна на $[0, 3d/2]$, когда $4\sigma^2 < 3d^2$. Вне этих интервалов $\varphi(t)$ должна равняться нулю. Постоянное значение $\varphi(t)$ может быть определено из условия

$$\int_0^\infty t^2\varphi(t)\,dt = \sigma^2, \tag{4.7}$$

а значение $F_0$ — из соотношения $F_0 + \int_0^\infty \varphi(t)\,dt = 1$.

Усеченная задача. Усеченный вариант поставленной выше задачи легко может быть истолкован с использованием предшествующих результатов. Допустим, мы хотим максимизировать

$$\Pr\{X \in (-\infty, \mu - d] \cup [\mu + d, \infty)\}, \qquad d > 0 \tag{4.8}$$

для распределений типа (4.1), которые усечены в точке $Q$, т. е. $\varphi(\mu + Q) = \varphi(\mu - Q) = 0$. В этой ситуации достаточно максимизировать

$$\Pr(X \in [d, \infty)), \tag{4.9}$$

где распределения $F_X$ имеют вид (4.4) и удовлетворяют ограничениям (4.5) и дополнительному условию $\varphi(Q) = 0$.

Проверим сначала, что если $\varphi(t)$ не возрастает и удовлетворяет соотношениям $\int_0^Q \varphi(t)\,dt \leqslant 1$ и $\int_0^Q t^2\varphi(t)\,dt = \sigma^2$, то

$$3\sigma^2 \leqslant Q^2. \tag{4.10}$$

Чтобы установить (4.10), мы определяем $\alpha = \min\{t \mid \varphi(t) < (1/Q)\}$.

Если $h(t)$ — произвольная неубывающая функция, то

$$\int_0^Q [h(t) - h(\alpha)]\left[\varphi(t) - \frac{1}{Q}\right] dt \leqslant 0. \tag{4.11}$$

Если дополнительно $h(t) \geqslant 0$, то неравенство (4.11) сводится к

$$\int_0^Q h(t)\varphi(t)\,dt \leqslant \int_0^Q \frac{h(t)}{Q}\,dt, \tag{4.12}$$

так как $\int\limits_0^Q \left[\varphi(t) - \frac{1}{Q}\right] dt \leqslant 0.$

Неравенство $3\sigma^2 \leqslant Q^2$ получается теперь подстановкой $h(t) = t^2$. Но (4.12) есть специальный случай общего результата (теорема 5.4), обсуждаемого в гл. XI.

Из неравенства $3\sigma^2 \leqslant Q^2$ следует, что если $3d^2 \leqslant 4\sigma^2$, то остается верным неусеченное решение и максимум (4.8) дается случаем (i) табл. 3. А в случае, когда $3d^2 \geqslant 4\sigma^2$, неусеченное решение останется верным при условии, что $3d/2 \leqslant Q$.

Остается рассмотреть случай, когда $3d^2 \geqslant 4\sigma^2$ и $a \leqslant Q \leqslant 3d/2$. Для применения теоремы 2.1 необходимо рассмотреть все те многочлены, которые ограничены $\psi(t)$ сверху на конечном интервале $[0, Q]$.

Как и ранее, многочлены $P(t) \geqslant \psi(t)$ $(0 \leqslant t \leqslant Q)$, которые касаются ненулевой части $\psi(t)$ при $t$, не превосходящем $Q$, рассматривается как возможные экстремальные многочлены в усеченной задаче. Кроме того, многочлены $P(t) \geqslant \psi(t)$ $(0 \leqslant t \leqslant Q)$, для которых $P(Q) = \psi(Q)$ (они могут пересекаться в этой точке) также могут доставлять минимум. Допустим, что $P(t)$ удовлетворяет неравенству $P(t) \geqslant \psi(t)$ $(0 \leqslant t \leqslant Q)$ и $aQ + bQ^3 = Q - d$, где $a \geqslant 0$. Тогда

$$a + 3b\sigma^2 = \left(1 - \frac{3\sigma^2}{Q^2}\right)a + 3\left(1 - \frac{d}{Q}\right)\frac{\sigma^2}{Q^2} \geqslant 3\left(1 - \frac{d}{Q}\right)\frac{\sigma^2}{Q^2}.$$

Значение $3\sigma^2 Q^{-3}(Q - d)$ есть максимум (4.8), так как

$$\frac{3\sigma^2}{Q^3}(Q - d) \leqslant \frac{4}{9}\frac{\sigma^2}{d^2} \quad \text{при } \frac{3}{2}d \geqslant Q > d.$$

Итак, единственное изменение в табл. 3 для усеченной задачи получается, когда $3d^2 > 4\sigma^2$ и $d \leqslant Q \leqslant 3/2d$. В этих условиях максимум (4.8) есть $3\sigma^2(1 - d/Q)/Q^2$. Это значение может быть достигнуто для прямоугольного распределения на $[\mu - Q, \mu + Q]$ с некоторой дополнительной массой в $\mu$.

Для полной строгости нам следовало бы сначала рассмотреть усеченный подкласс $\mathcal{F}_0$, а затем распространить результаты так, как мы сделали в неусеченной задаче. Формальное доказательство предоставляется проделать читателю.

Пример 4.2. Мы хотим определить максимум

$$\Pr\{X \in [d, \infty)\}, \quad d > 0, \tag{4.13}$$

по отношению к классу неотрицательных унимодальных распределений с модой в $m > 0$ и заданным средним $\mu > 0$. Более точно будут рассматриваться распределения вида

$$F(x) = \begin{cases} 0, & x < 0, \\ \int_0^x \varphi(t)\,dt, & 0 < x < m, \\ F_m + \int_0^x \varphi(t)\,dt, & x \geqslant m, \end{cases} \tag{4.14}$$

где $\varphi$ не убывает на $[0, m]$ и не возрастает на $(m, \infty)$, $F_m$ есть скачок $F$ в точке $m$, и $F$ удовлетворяет моментному условию $\int_0^\infty x\,dF(x) = \mu$.

Решение представлено в табл. 4.

ТАБЛИЦА 4

| Случай | Условие | Максимум | Экстремальное распределение |
|---|---|---|---|
| (i) | $d \geqslant m$ и $2\mu - m \geqslant d + \sqrt{d^2 - md}$ | $1 - \dfrac{d-m}{2(\mu - m)}$ | Прямоугольное на $[m, 2\mu - m]$. |
| (ii) | $d \geqslant m$ и $2\mu - m \leqslant d + \sqrt{d^2 - md}$ | $\dfrac{2\mu - m}{(\sqrt{d} + \sqrt{d-m})^2}$ | Прямоугольное на $[0, m]$ и на $[m, d + \sqrt{d^2 - md}]$. |
| (iii) | $d \leqslant m$ и $2\mu - m \leqslant d$ | $\dfrac{2\mu - d}{m}$ | Прямоугольное на $[0, d]$ и на $[d, m]$. |
| (iv) | $d \leqslant m$ и $2\mu - m \geqslant d$ | 1 | |

Чтобы проверить указанное решение, мы используем незначительную модификацию метода из примера 4.1. Сначала рассмотрим распределения $F$ вида (4.14), которые не допускают скачков в точке $m$. Затем, полагая

$$dH(t) = -(t - m)\,d\varphi(t) \tag{4.15}$$

и интегрируя по частям, мы получаем

$$\int_0^\infty \varphi(t)\,dt = 1 = \int_0^\infty dH(t), \quad \int_0^\infty t\varphi(t)\,dt = \mu = \frac{m}{2} + \int_0^\infty \frac{t}{2}\,dH(t)$$

(отсюда $2\mu \geqslant m$) и

$$\int_0^\infty \varphi(t)\,dt = \begin{cases} \int_0^\infty \frac{t-d}{t-m}\,dH(t), & \text{если } d \geqslant m, \\ \int_0^\infty \frac{m-d}{m-t}\,dH(t) + \int_0^\infty dH(t), & \text{если } d < m. \end{cases}$$

Для $d \geqslant m$ положим

$$\psi_1(t) = \begin{cases} 0, & 0 \leqslant t \leqslant d, \\ \frac{t-d}{t-m}, & t \geqslant d. \end{cases}$$

Задача состоит теперь в минимизации $b_1 + b_2(2\mu - m)$ по отношению ко всем многочленам, удовлетворяющим условию $b_1 + b_2 t \geqslant \geqslant \psi_1(t)$. Ясно, что минимум достигается на прямой линии, которая касательна к ненулевой части $\psi_1$, т. е. для некоторого $\eta > d$ экстремальная прямая удовлетворяет условиям

$$b_1 + b_2\eta = \frac{\eta - d}{\eta - m},$$

$$b_2 = \frac{d-m}{(\eta - m)^2}.$$

В этом случае

$$b_1 = \frac{\eta^2 - 2d\eta + md}{(\eta - m)^2} \geqslant 0,$$

откуда следует, что $\eta \geqslant d + \sqrt{d^2 - md}$. Принимая во внимание это условие, для получения первого из двух вариантов решения мы минимизируем выражение

$$b_1 + b_2(2\mu - m) = 1 - \frac{2(d-m)(\eta - \mu)}{(\eta - m)^2}.$$

Отметим, что мы еще должны доказать пригодность указанных границ для распределений $F$, допускающих скачок в точке $m$. Для определения экстремального распределения мы замечаем, что $H$ состоит из скачков в точках, для которых минимизирующий многочлен удовлетворяет равенству $b_1 + b_2 t = \psi_1(t)$, а затем используем (4.15).

Для $d < m$ мы определим

$$\psi_2(t) = \begin{cases} \frac{m-d}{m-t}, & 0 \leqslant t \leqslant d, \\ 1, & t \geqslant d, \end{cases}$$

и минимизируем $b_1 + b_2(2\mu - m)$ среди многочленов, удовлетворяющих условию $b_1 + b_2 t \geqslant \psi_2(t)$. Минимум достигается при $b_1 + b_2 d = 1$, где $(m-d)/m \leqslant b_1 \leqslant 1$, так как функция $(m-t)^{-1}(m-d)$ выпукла, а $b_2(2\mu - m)$ возрастает в точке $b_2$. Тогда $b_1 + b_2(2\mu - m) = (2\mu - m)/d - b_1(2\mu - m - d)$, и для того, чтобы вывести второй вариант решения, мы выбираем $b_1 = 1$, если $2\mu - m > d$ и $b_1 = (m - d)/m$, если $2\mu - m < d$.

Чтобы распространить упомянутые результаты на распределения, допускающие скачок в точке $m$, мы действуем так же, как в в примере 4.1.

Неизвестная мода. Используем теперь решение задачи для известной моды, рассмотренной выше, для максимизации над более обширным классом всех унимодальных распределений вида (4.14), где мода не задана. С этой целью мы теперь рассматриваем моду как дополнительную переменную. Для моды, расположенной в $m$, мы знаем максимум $\Pr\{X \in [d, \infty)\}$, указанный в табл. 4. Остается максимизировать эти значения по $m$. Анализ решений из табл. 4 показывает, что

$$\Pr(X \in [d, \infty)) \leqslant \begin{cases} 1, & d \leqslant \mu, \\ \dfrac{2\mu}{d} - 1, & \mu \leqslant d \leqslant \dfrac{3}{2}\mu, \\ \dfrac{\mu}{2d}, & d \geqslant \dfrac{3}{2}\mu. \end{cases} \tag{4.16}$$

Эти три границы достигаются при условиях $d \leqslant m \leqslant 2\mu - d$, $m = d$ и $m = 0$ соответственно. Обе нетривиальные границы получены из второй строки табл. 4, где экстремальное распределение состоит из двух прямоугольных кусков. Таким образом, при $m = 0$ или $m = d$ в моде сосредоточена некоторая масса. При $m = d$ эта масса есть $2\mu/d - 1$, а при $m = 0$ она равна $1 - \mu/d$.

Если $d = 3\mu/2$, то ни одна из масс не равна нулю, так что максимум достигается при $m = 0$ или $d$, но не достигается ни для какой моды в промежутке между ними. Таким образом, в нетривиальных случаях максимизирующее значение $m$ либо единственно, либо есть одна из двух точек.

Пример 4.3. Неравенство Роудена. В этом примере мы хотим определить максимум

$$\Pr\{X \in (-\infty, -d) \cup [d, \infty)\} \tag{4.17}$$

над классом унимодальных распределений с модой в нуле, удовлетворяющих моментным условиям

$$\int_{-\infty}^{\infty} |x|\, dF(x) = \nu_1, \quad \int_{-\infty}^{\infty} x^2\, dF(x) = \nu_2 \qquad (\nu_1 > 0). \tag{4.18}$$

Решение дано ниже.

ТАБЛИЦА 5[1])

| Случай | Условие | Максимум | Экстремальное распределение |
|---|---|---|---|
| (i) | $0 \leqslant d \leqslant \nu_1$ | $1-\dfrac{d}{2\nu_1}$ | Прямоугольное на $[0, 2\nu_1]$ с «массой» в $\infty$ |
| (ii) | $\nu_1 \leqslant d \leqslant \dfrac{3\nu_2}{4\nu_1}$ | $\dfrac{\nu_1}{2d}$ | Прямоугольное на $[0, 2d]$ с массой в 0 и «массой» в $\infty$ |
| (iii) | $\dfrac{3\nu_2}{4\nu_1} \leqslant d \leqslant \dfrac{\nu_2}{\nu_1}$ | $\dfrac{4\nu_1}{3\nu_2}-\dfrac{8\nu_1^3 d}{9\nu_2^2}$ | Прямоугольное на $[0, 3\nu_2/2\nu_1]$ с массой в 0 |
| (iv) | $\dfrac{\nu_2}{\nu_1} \leqslant d$ | $\dfrac{3\nu_2-4\nu_1^2}{3\eta^2-8\nu_1\eta+3\nu_2}$ | Прямоугольное на $[0, 3\eta-2\eta^2 d^{-1}]$ и на $[3\eta-2\eta^2 d^{-1}, \eta]$ с массой в 0 |

[1]) $\eta$—наибольший корень для $2\eta^3-(3d+4\nu_1)\eta^2+8\nu_1 d\eta-3\nu_2 d=0$.

Нам представляется удобным описать экстремальные распределения в табл. 5 с массами, сосредоточенными на неотрицательной полуоси. Чтобы перевести решение на всю ось, масса на неотрицательной полуоси распределяется симметричным образом по отношению к началу координат. Некоторые разъяснения необходимы по отношению к понятию «массы» на бесконечности. Для каждой строки в табл. 5 с «массой» на бесконечности можно представить последовательность функций распределения $G_k$, определенных на $[0, \infty)$, удовлетворяющих моментным ограничениям (4.18) и таких, что предел

$$\lim_{k\to\infty} (1-G_k(d-))$$

равен своему максимальному значению.

Применяя принцип выбора Хелли по отношению к $\{G_k\}$, получим предельное распределение $G(x)$. К сожалению, сходимость вторых моментов

$$\int_0^\infty t^2\, dG_k(t) \to \int_0^\infty t^2\, dG(t) \tag{4.19}$$

может нарушаться, поэтому $G$ не удовлетворяет автоматически требуемым моментным условиям.

Эту трудность можно обойти следующим образом. Пусть $C_2\,[0, \infty)$ обозначает класс непрерывных функций $\varphi(t)$, определенных на $[0, \infty)$, таких, что предел

$$\lim_{t\to\infty} \frac{\varphi(t)}{t^2} = \alpha$$

существует и конечен. Распределение $G$ с массой $\lambda$ рассматривается как линейный функционал $L$, определенный на $C_2\,[0, \infty)$

посредством выражения

$$L(\varphi) = \int_0^\infty \varphi(t)\,dG(t) + \lambda \lim_{t\to\infty} \frac{\varphi(t)}{t^2}. \tag{4.20}$$

Здесь мы предполагаем, что $\int_0^\infty t^2 dG(t) < \infty$. Заметим, что если $\varphi(t) = 1$ или $\varphi(t) = t$, то $L(\varphi)$ сводится к обычному нулевому и первому моментам. «Масса» $\lambda$ на бесконечности компенсирует ту часть второго момента, которая теряется при предельном переходе в (4.19).

Мы теперь проведем рассуждение для случая (i) с некоторыми деталями. Прямоугольное распределение $G$ на $[0, 2\nu_1]$ с массой $\lambda = \nu_2 - (4/3)\nu_1^2$ на бесконечности удовлетворяет моментным ограничениям (4.18) и максимизирует (4.17). Пусть $F$ обозначает прямоугольную часть $G$ на $[0, 2\nu_1]$, а $F_k$ — представляют собой распределения, состоящие из прямоугольного куска на $[0, k]$ с весом $\varepsilon = 2\nu_1/k^2$ и массами $1 - \varepsilon k$ в нуле.

Тогда последовательность $\{G_k\}$ определяется как $G_k = \alpha_k F + (1 - \alpha_k) F_k$, где

$$\alpha_k = \frac{k - \frac{3}{2}\frac{\nu_2}{\nu_1}}{k - 2\nu_1}.$$

Распределения $G_k$ унимодальны и удовлетворяют соотношениям (4.18). Кроме того, можно показать, что для всякой функции $\varphi \in C_2[0, \infty)$ мы имеем

$$\lim_{k\to\infty} \int_0^\infty \varphi(t)\,dG_k(t) = \int_0^\infty \varphi(t)\,dF(t) + \left(\nu_2 - \frac{4}{3}\nu_1^2\right) \lim_{t\to\infty} \frac{\varphi(t)}{t^2}.$$

В этом случае последовательность $\{G_k\}$ «сходится» к прямоугольному распределению $G$ на $[0, 2\nu_1]$ с массой $\lambda = \nu_2 - \frac{4}{3}\nu_1^2$ на бесконечности.

Проверим теперь результаты из табл. 5.

Так как множество $(-\infty, d] \cup [d, \infty)$ симметрично относительно моды, мы можем свести, как и в примере 3.2, рассматриваемую задачу к односторонней задаче максимизации $\Pr(X \in [d, \infty))$ над классом распределений, удовлетворяющих (4.18) и имеющих вид

$$F(x) = \begin{cases} F_0 + \int_0^x \varphi(t)\,dt, & x \geqslant 0, \\ 0, & x < 0, \end{cases}$$

где $\varphi$ не возрастает.

Мы примем для начала, что $F_0 = 0$; распространение на более широкий класс осуществляется так же, как и в примере 4.1.

Определив меру $dH(t) = -td\varphi(t)$ и выполнив интегрирование по частям, мы приходим к соотношениям

$$
\begin{aligned}
1 &= \int_0^\infty \varphi(t)\,dt = \int_0^\infty dH(t),\\
\nu_1 &= \int_0^\infty t\varphi(t)\,dt = \int_0^\infty \frac{t}{2}\,dH(t),\\
\nu_2 &= \int_0^\infty t^2\varphi(t)\,dt = \int_0^\infty \frac{t^2}{3}\,dH(t),\\
&\int_d^\infty \varphi(t)\,dt = \int_d^\infty \left(1-\frac{d}{t}\right)dH(t).
\end{aligned}
\tag{4.21}
$$

Если теперь мы определим

$$
\psi(t) = \begin{cases} 0, & 0 \leqslant t \leqslant d,\\ 1-\dfrac{d}{t}, & t \geqslant d,\end{cases}
$$

и соответственно применим теорему 2.1, то задача сведется к задаче определения минимума

$$
a + 2b\nu_1 + 3c\nu_2, \tag{4.22}
$$

где $a$, $b$ и $c$ изменяются произвольно, но так, чтобы выполнялись условия

$$
P(t) = a + bt + ct^2 \geqslant \psi(t), \qquad t \in [0, \infty).
$$

Умножая многочлен $P(t)$ на положительный масштабный множитель, получим, что необходимо рассмотреть только многочлены $P(t)$ такие, что существует некоторое $\eta > d$, удовлетворяющее уравнениям

$$
P(\eta) = a + b\eta + c\eta^2 = 1 - \frac{d}{\eta}, \quad P'(\eta) = b + 2c\eta = \frac{d}{\eta^2}. \tag{4.23}
$$

Решая уравнения (4.23) для $a$ и $b$ в терминах $c$ и $\eta$, получим

$$
a = 1 - \frac{2d}{\eta} + c\eta^2, \quad b = \frac{d}{\eta^2} - 2c\eta. \tag{4.24}
$$

В этом случае

$$
a + 2b\nu_1 + 3c\nu_2 = (\eta^2 - 4\nu_1\eta + 3\nu_2 c) + \frac{1}{\eta^2}(\eta^2 - 2d\eta + 2\nu_1 d). \tag{4.25}
$$

Подобным образом, решая эти уравнения для $b$ и $c$ в терминах $a$ и $\eta$, мы получим

$$b=\frac{1}{\eta}\left(2-\frac{3d}{\eta}-2a\right),\quad c=\frac{1}{\eta^2}\left(\frac{2d}{\eta}-1+a\right)$$

и, следовательно,

$$a+2b\nu_1+3c\nu_2=\frac{1}{\eta^2}(\eta^2-4\nu_1\eta+3\nu_2)a+$$
$$+\frac{1}{\eta^3}[4\nu_1\eta^2-(3\nu_2+6d\nu_1)\eta+6\nu_2 d]. \qquad (4.26)$$

Из неравенства Шварца и выражений для $\nu_1$ и $\nu_2$ в (4.21) заключаем, что $3\nu_1\geqslant 4\nu_1$, откуда $\eta^2-4\nu_1\eta+3\nu_2\geqslant 0$ для всех вещественных $\eta$. Выражение $\eta^2-4\nu_1\eta+3\nu_2$ появляется в первом члене правых частей (4.25) и (4.26).

Теперь займемся минимизацией (4.22) относительно $\eta$ для трех отдельных интервалов $\eta\geqslant 2d$, $3d/2\leqslant\eta\leqslant 2d$ и $d<\eta\leqslant 3d/2$. Когда условия на $\nu_1$, $\nu_2$ и $d$ в отдельных минимизациях перекрываются, будем сравнивать эти частные минимумы. Если $\eta\geqslant 2d$, то из (4.24) следует, что $a\geqslant 0$ и $b\geqslant 0$ при $c=0$.

Следовательно, мы можем выбрать $c=0$ в (4.25) и минимизировать выражение

$$\frac{1}{\eta^2}(\eta^2-2d\eta+2\nu_1 d),\qquad \eta\geqslant 2d.$$

При $\nu_1\geqslant d$ минимум равен $1-d/2\nu_1$ и достигается при $\eta=2\nu_1$, а если $\nu_1\leqslant d$, то минимум равен $\nu_1/2d$ и достигается при $\eta=2d$.

С помощью аналогичного анализа мы минимизируем (4.26) в области $3d/2\leqslant\eta\leqslant 2d$, замечая, что в этих условиях $b\geqslant 0$ и $c\geqslant 0$ при $a=0$, и, следовательно, мы можем принять $a=0$ в (4.26) и минимизировать второй член правой части. В этом случае мы получаем результаты, представленные в табл. 6.

ТАБЛИЦА 6

| Случай | Условие | Минимум | Достигается для |
|---|---|---|---|
| (i) | $0\leqslant d\leqslant\frac{3\nu_2}{4\nu_1}$ | $\frac{\nu_1}{2d}$ | $\eta=2d$ |
| (ii) | $\frac{3\nu_2}{4\nu_1}\leqslant d\leqslant\frac{\nu_2}{\nu_2}$ | $\frac{4\nu_1^2}{3\nu_2}-\frac{8\nu_1^3 d}{9\nu_2^2}$ | $\eta=\frac{3\nu_2}{2\nu_2}$ |
| (iii) | $\frac{\nu_2}{\nu_1}\leqslant d$ | $\frac{4\nu_2}{9d^2}$ | $\eta=\frac{3}{2}d$ |

Сравнивая минимум, полученный при $\eta\geqslant 2d$ и $3d/2\leqslant\eta\leqslant 2d$, видим, что первые три варианта решения в табл. 5 подтверждены.

Однако остается еще минимизировать (4.22) в области $d < \eta \leqslant 3/2d$. В этом случае $a + bt + ct^2 \geqslant \psi(t)$ тогда и только тогда, когда $c \geqslant d^2/[4\eta^3(\eta - d)] = c^*$. (Результативный многочлен $a^* + b^*t + c^*t^2$ обращается в нуль при $t = 3\eta - 2\eta^2 d^{-1}$.) Так как (4.25) убывает по $c$, мы полагаем $c = c^*$. При таком выборе $c$ имеем

$$a + 2b\nu_1 + 3c\nu_2 = \frac{4\eta^4 - 12d\eta^3 + (8\nu_1 d + 9d^2)\eta^2 - 12\nu_1 d^2\eta + 3\nu_2 d^2}{4\eta^3(\eta - d)}.$$

Числитель производной от правой части по $\eta$ равен

$$4\eta^2 d(4\eta - 3d)[2\eta^3 - (3d + 4\nu_1)\eta^2 + 8\nu_1 d\eta - 3\nu_2 d];$$

это выражение имеет четыре положительных нуля и отрицательно между двумя наибольшими корнями. Исследуя производные числителя, мы находим, что наибольший корень не превосходит $3d/2$ тогда и только тогда, когда $d \geqslant \nu_2/\nu_1$. Таким образом, мы устанавливаем, что первые три пункта действительно доставляют решение при $d \leqslant \nu_2/\nu_1$ и решение в последнем случае, когда $d \geqslant \nu_2/\nu_1$ такое, как указано.

## § 5. Унимодальность высшего порядка

Исследования этого параграфа, касающиеся чебышевских неравенств для обобщенных унимодальных распределений, были начаты Гауссом. Наше изложение совершенствует и обобщает недавнюю работу Маллоу [1963] (см. также Маллоу [1956] и сопровождающую дискуссию).

Пример 5.1. *Функции распределения s-го порядка гладкости с границами* $\lambda_0, \lambda_1, \ldots, \lambda_s$.

Определение 5.1. Пусть $\lambda_0, \lambda_1, \ldots, \lambda_s$ — непрерывные положительные функции на открытом конечном или бесконечном интервале $(a, b)$. Говорят, что функция распределения $F$ обладает *гладкостью s-го порядка с границами* $\lambda_0, \lambda_1, \ldots, \lambda_s$, если:

(1) $(s+1)$-я производная от $F$ всюду существует,

(2) для некоторых $x_1, x_2, \ldots, x_s$, удовлетворяющих неравенствам $a = x_0 < x_1 < \ldots < x_s < x_{s+1} = b$, имеем $0 \leqslant (-1)^r F^{(s+1)}(x) \leqslant \lambda_r(x)$, $x_r \leqslant x \leqslant x_{r+1}$, $r = 0, 1, \ldots, s$.

Полезно исследовать смысл этого определения при специальном выборе $s$. Распределение нулевого порядка гладкости имеет функцию плотности $F'$, которая ограничена $\lambda_0$ на $(a, b)$. Определение гладкости порядка $s = 1$ означает, что плотность имеет производную, которая неотрицательна и ограничена $\lambda_0$ на некотором интервале $(a, x_1]$ и неположительна и ограничена снизу функцией $-\lambda_1$ на $[x_1, b)$. Это означает, в частности, что $F$ унимодальна с модой в $x_1$. Отметим, что случай $s = 1$ не включает все унимодальные функции распределения, как мы их определили в § 4. Однако произвольное унимодальное распределение может быть получено как предел распределений первого порядка гладкости, если устремить $\lambda_0(x)$ и $\lambda_1(x)$ к бесконечности и допустить массы, сосредоточенные в $x_1$. Для случая гладкой функции распределения $F$ второго порядка плотность $F'$ имеет две точки перегиба, $x_1$ и $x_2$ и третья производная $F^{(3)}$ имеет чередующиеся знаки на трех интервалах $(a, x_1)$, $(x_1, x_2)$ и $(x_2, b)$ и ограничена на этих интервалах $\lambda_0$, $-\lambda_1$ и $\lambda_2$ соответственно.

Для последовательности чисел $\mu_0, \mu_1, \ldots, \mu_n$ мы хотим определить точные границы на $F(t)$, где $F$ принадлежит классу $\mathcal{F}$ функций распределения, которые обладают гладкостью $s$-го порядка с заданными границами $\lambda_0, \lambda_1, \ldots, \lambda_s$ и удовлетворяют моментным условиям

$$\int_a^b t^k dF(t) = \mu_k, \qquad k = 0, 1, \ldots, n. \tag{5.1}$$

Узловые точки $x_1, x_2, \ldots, x_s$ выбраны заранее. Далее будем предполагать, что класс $\mathcal{F}$ не является пустым.

Чтобы решить эту задачу, мы действуем следующим образом. Интегрирование (5.1) по частям дает соотношения

$$\int_a^b t^k F^{(s+1)}(t)\,dt = \begin{cases} (-1)^s \dfrac{k!}{(k-s)!}\,\mu_{k-s}, & k = s, s+1, \ldots, s+n, \\ 0, & k = 0, 1, \ldots, s-1, \end{cases}$$
$$F(t_0) = \int_a^{t_0} dF(t) = \frac{(-1)^s}{s!}\int_a^{t_0} (t-t_0)^s F^{(s+1)}(t)\,dt. \tag{5.2}$$

Теперь можно применить теорему 8.1 гл. VIII. В самом деле, если мы положим

$$d\nu(x) = \begin{cases} \lambda_{2r}(x)\,dx, & x_{2r} \leqslant x \leqslant x_{2r+1}, \\ 0, & x_{2r+1} \leqslant x \leqslant x_{2r+2}, \end{cases} \qquad r=0, 1, \ldots,$$

$$d\lambda(x) = \begin{cases} 0, & x_{2r} \leqslant x \leqslant x_{2r+1}, \\ -\lambda_{2r+1}(x)\,dx, & x_{2r+1} \leqslant x \leqslant x_{2r+2}, \end{cases} \qquad r = 0, 1, \ldots,$$

то $d\nu$ и $d\lambda$ суть конечные (обобщенны ) меры на $(a, b)$.

Рассмотрим пространство моментов

$$\sum_{n+s+1}(\nu, \lambda) =$$
$$= \left\{ \mathbf{c} = (c_0, c_1, \ldots, c_{n+s}) \mid c_k = \int_a^b t^k d\sigma(t), \qquad k = 0, 1, \ldots, n+s \right\},$$

где $\sigma$ подчинено ограничениям

$$d\lambda \leqslant d\sigma \leqslant d\nu. \tag{5.3}$$

Отметим, что этот класс мер $d\sigma$ включает класс $\mathcal{F}$. Определим точку $\mathbf{c}^0 = (c_0^0, c_1^0, \ldots, c_{s+n}^0)$ как

$$c_k^0 = \begin{cases} 0, & k = 0, 1, \ldots, s-1, \\ \dfrac{(-1)^s\, k!}{(k-s)!}\,\mu_{k-s}, & k = s, s+1, \ldots, s+n. \end{cases}$$

Если точка $\mathbf{c}^0$ есть граничная точка $\sum\limits_{n+s+1}(\nu, \lambda)$, то $\mathcal{F}$ состоит из одного элемента, и задача тривиальна. Если точка $\mathbf{c}^0$ содержится во внутренности $\sum\limits_{n+s+1}(\nu, \lambda)$, мы определяем

$$\Omega_\varepsilon(t) = (-1)^s \frac{(t-t_0)^s}{s!} + \varepsilon t^{n+s+1}, \quad a < t < b, \ \varepsilon > 0,$$

при этом функции $1, t, t^2, \ldots, t^{n+s}$, $\Omega_\varepsilon(t)$ удовлетворяют условиям теоремы 8.1 гл. VIII.

Теперь мы заключаем, что для произвольной меры $d\sigma$, представляющей $\mathbf{c}^0$, справедливо

$$\int_a^{t_0} \Omega_\varepsilon(t)\, d\underline{\sigma}_{t_0}(t) \leqslant \int_a^{t_0} \Omega_\varepsilon(t)\, d\sigma(t) \leqslant \int_a^{t_0} \Omega_\varepsilon(t)\, d\bar{\sigma}_{t_0}(t). \tag{5.4}$$

Меры $d\underline{\sigma}_{t_0}$ и $d\bar{\sigma}_{t_0}$ суть канонические меры, представляющие $\mathbf{c}^0$, которые равны $d\nu$ и $d\lambda$ на чередующихся интервалах. Индекс числа этих интервалов есть $2^{-1}(n+s+1)$ или $2^{-1}(n+s+2)$ и $d\underline{\sigma}_{t_0}(d\bar{\sigma}_{t_0})$ начинает (кончает) интервал равенства с $d\nu$ в точке $t_0$.

Отметим, что члены в (5.4) не зависят от $\varepsilon$. Фактически функция $\Omega_\varepsilon$ может быть заменена любой функцией, удовлетворяющей условиям теоремы 8.1 гл. VIII, и (5.4) будет выполняться.

Устремляя $\varepsilon$ к нулю, из (5.4) получаем, что

$$\int_a^{t_0} \Omega\, d\underline{\sigma}_{t_0} \leqslant F(t_0) \leqslant \int_a^{t_0} \Omega\, d\bar{\sigma}_{t_0},$$

где $\Omega(t) = \dfrac{(-1)^s (t-t_0)^s}{s!}$, $\quad a < t \leqslant t_0$.

Обозначим

$$\underline{F}_{t_0}(t) = \int_a^t \Omega\, d\underline{\sigma}_{t_0}, \quad \bar{F}_{t_0}(t) = \int_a^t \Omega\, d\bar{\sigma}_{t_0}, \tag{5.5}$$

тогда $\underline{F}_{t_0}(t_0) \leqslant F(t_0) \leqslant \bar{F}_{t_0}(t_0)$, $F \in \mathcal{F}$.

Отметим, что функции $\underline{F}_{t_0}$ и $\bar{F}_{t_0}$ принадлежат классу $\mathcal{F}$. Этот факт может быть легко установлен исследованием вида $d\underline{\sigma}_{t_0}$ и $d\bar{\sigma}_{t_0}$. Проведенное рассуждение доказывает следующую теорему.

Теорема 5.1. *Пусть $s$ и $n$ — неотрицательные целые числа, а $x_1, x_2, \ldots, x_s$ удовлетворяют условию $-\infty \leqslant a = x_0 < x_1 < \ldots \ldots < x_s < x_{s+1} = b \leqslant \infty$. Пусть $\mathcal{F}$ есть класс функций распределения $F$ на $(a, b)$, которые обладают гладкостью $s$-го порядка с границами $\lambda_0, \lambda_1, \ldots, \lambda_s$ и узловыми точками изменения $x_1, x_2, \ldots \ldots, x_s$ и таковы, что*

$$\int_a^b t^k dF(t) = \mu_k, \qquad k = 0, 1, \ldots, n.$$

*Тогда для любой функции $F \in \mathcal{F}$ имеем*

$$\underline{F}_{t_0}(t_0) \leqslant F(t_0) \leqslant \overline{F}_{t_0}(t_0),$$

*где $\underline{F}_{t_0}$ и $\overline{F}_{t_0}$ определены в* (5.5).

Пример 5.2. *Колоколообразные распределения.*

Определение 5.2. Говорят, что функция распределения $F$ является *колоколообразной $s$-го порядка с узловыми точками $x_1, x_2, \ldots, x_s$*, если:

(1) $F^{(s)}(x)$ существует и непрерывна для всех точек, за исключением конечного числа, и производные $s$-го порядка слева и справа $F_{-}^{(s)}(x)$ и $F_{+}^{(s)}(x)$ всюду существуют и являются соответственно непрерывными слева и справа;

(2) $F^{(s)}(x)$ непрерывна в точках $x_1, x_2, \ldots, x_s$ и $(-1)^r F_{-}^{(s)}(x)$ и $(-1)^r F_{+}^{(s)}(x)$ не убывают на множестве $x_r < x < x_{r+1}$, $r = 0, 1, \ldots, s$, где $x_0 = -\infty$ и $x_{s+1} = +\infty$.

Мы обозначим этот класс через $\mathcal{G}$.

Из сравнения определений 5.2 и 5.1 очевидно, что всякая функция распределения $F$, которая обладает гладкостью $s$-го порядка, является колоколообразной, но не наоборот.

В этом примере мы хотим максимизировать

$$1 - F(d) = \int_d^\infty dF(t) \tag{5.6}$$

над классом функций распределения $F \in \mathcal{G}$, удовлетворяющих моментным условиям

$$\int_{-\infty}^{\infty} t^k dF(t) = \mu_k, \qquad k = 0, 1, \ldots, n. \tag{5.7}$$

Этот подчиненный ограничениям класс обозначим $\mathcal{G}(\mu_0, \ldots, \mu_n)$.

Для любой функции $F \in \mathcal{G}(\mu_0, \ldots, \mu_n)$ мера

$$dH(t) = (-1)^s (t - x_1) \ldots (t - x_s)\, dF_{+}^{(s)}(t)$$

определена и неотрицательна.

Мы будем предполагать, что $\mathcal{G}(\mu_0, \ldots, \mu_n)$ содержит по крайней мере одно распределение $F_+$ такое, что точка с координатами

$$c_k = \int_{-\infty}^{\infty} t^k dH(t), \qquad k = 0, 1, \ldots, n, \tag{5.8}$$

содержится во внутренности пространства моментов $\mathcal{M}_{n+1}$, образованного системой $\{t_i\}_0^n$ на $(-\infty, \infty)$. Значения $c_k$, $k = 0, 1, \ldots, n$, определены и конечны, так как

$$c_k = (-1)^s \int_{-\infty}^{\infty} t^k (t - x_1) \ldots (t - x_s)\, dF^{(s)}(t)$$

и

$$\int_{-\infty}^{\infty} t^m dF^{(s)}(t) = \begin{cases} 0, & m = 0, 1, \ldots, s-1, \\ \dfrac{(-1)^s\, m!}{(m-s)!}\, \mu_{m-s}, & m = s, s+1, \ldots, s+n. \end{cases}$$

В частности, $c_0 = s!\,\mu_0$.

Для применения теоремы 2.1 необходимо, чтобы (5.6) имело вид

$$\int_{d}^{\infty} dF(t) = \int_{-\infty}^{\infty} \Omega(t)\, dH(t) \tag{5.9}$$

при подходящем выборе $\Omega(t)$.

При $d > x_s$ мы можем проинтегрировать (5.6) по частям и получить

$$\int_{d}^{\infty} dF = \int_{d}^{\infty} \frac{(t-d)^s}{s!\,(t-x_1)\ldots(t-x_s)}\, dH(t),$$

а при $d < x_1$ мы можем написать

$$\int_{d}^{\infty} dF = 1 - \int_{-\infty}^{d} dF(t) = \int_{-\infty}^{\infty} \frac{dH(t)}{\mu_0 s!} - \int_{-\infty}^{d} \frac{(t-d)^s}{s!\,(t-x_1)\ldots(t-x_s)}\, dH(t).$$

Следовательно, для $d > x_s$ мы можем определить $\Omega$ в (5.9) как

$$\Omega(t) = \begin{cases} 0, & t < d, \\ \dfrac{(t-d)^s}{s!\,(t-x_1)\ldots(t-x_s)}, & t \geqslant d, \end{cases} \tag{5.10}$$

а для $d < x_1$

$$\Omega(t) = \begin{cases} (s!\,\mu_0)^{-1}, & t \geqslant d, \\ \dfrac{1}{\mu_0 s!} - \dfrac{(t-d)^s}{s!\,(t-x_1)\ldots(t-x_s)}, & t < d. \end{cases} \tag{5.11}$$

В общем случае, когда $x_k < d < x_{k+1}$ для некоторого $k = 1, \ldots$ $\ldots, s-1$, мы покажем, что существуют многочлены $P$ и $Q$ такие, что

$$\int_d^\infty dF(t) = \int_d^\infty \frac{P(t)}{(t-x_1)\ldots(t-x_k)}\, dH(t) + \int_{-\infty}^d \frac{Q(t)}{(t-x_{k+1})\ldots(t-x_s)}\, dH(t).$$

С этой целью положим

$$\begin{aligned} R(t) &= P(t)(t-x_{k+1})\ldots(t-x_s), \\ S(t) &= Q(t)(t-x_1)\ldots(t-x_k). \end{aligned} \tag{5.12}$$

Определим $R$ как многочлен степени $s$ с главным коэффициентом $(-1)^s/s!$ и $S$ как многочлен степени $s-1$. Отметим, что

$$\begin{aligned} S(x_i) &= 0, \qquad i = 1, \ldots, k, \\ R(x_i) &= 0, \qquad i = k+1, \ldots, s. \end{aligned} \tag{5.13}$$

Дополнительно потребуем, чтобы

$$S^{(i)}(d) = R^{(i)}(d), \qquad i = 0, 1, \ldots, s-1. \tag{5.14}$$

Если положить

$$\begin{aligned} R(t) &= \frac{(-1)^s}{s!} t^s + a_{s-1}t^{s-1} + \ldots + a_0, \\ S(t) &= b_{s-1}t^{s-1} + \ldots + b_0, \end{aligned}$$

то уравнения (5.13) и (5.14) относительно неизвестных $\mathbf{z}=(a_{s-1}, \ldots$ $\ldots, a_0, b_{s-1}, \ldots, b_0)$ имеют единственное решение. Мы запишем эти уравнения кратко в матричных обозначениях как

$$\mathbf{Az} = \mathbf{c},$$

где $\mathbf{z}$ и $\mathbf{c}$ — векторы с $2s$ компонентами, а $\mathbf{A}$ — квадратная $2s \times 2s$-матрица. Легко видеть, что вектор $\mathbf{c}$ не является нулевым ни при каком выборе $d$ и что определитель $\mathbf{A}$ может быть представлен как пропорциональный к

$$\pm \begin{vmatrix} x_1^{s-1} & x_1^{s-2} & \ldots & 1 \\ x_2^{s-1} & x_2^{s-2} & \ldots & 1 \\ \cdot & \cdot & \cdot & \cdot \\ x_s^{s-1} & x_s^{s-2} & \ldots & 1 \end{vmatrix}$$

заменой в $\mathbf{A}$ $i$-го столбца $(i = 1, 2, \ldots, s)$ на $i$-й плюс $(i+s)$-й столбец.

Определим теперь многочлены $P$ и $Q$ посредством (5.12) и запишем (5.6) в виде

$$\int_{d}^{\infty} dF = \int_{-\infty}^{\infty} \Omega(t)\, dH(t),$$

где

$$\Omega(t) = \begin{cases} \dfrac{P(t)}{(t-x_1)\ldots(t-x_k)}, & t \geqslant d, \\ \dfrac{Q(t)}{(t-x_{k+1})\ldots(t-x_s)}, & t < d. \end{cases}$$

Мы можем получить теперь точные границы для интеграла (5.15)

$$\int^{\infty} dF(t), \tag{5.15}$$

применяя конструктивный метод теоремы 2.1. Проделанная процедура, которая позволяет записать интеграл (5.15) как интеграл по $(-\infty, \infty)$ в терминах некоторой функции $\Omega(t)$, дает в наше распоряжение средства вычисления границ для интегралов по более общим множителям, например

$$\int_{\alpha}^{\beta} dF(t) = \int_{-\infty}^{\infty} \Omega_1(t)\, dt - \int_{-\infty}^{\infty} \Omega_2(t)\, dt,$$

где

$$\int_{\alpha}^{\infty} dF(t) = \int_{-\infty}^{\infty} \Omega_1(t)\, dt,$$

$$\int_{\beta}^{\infty} dF(t) = \int_{-\infty}^{\infty} \Omega_2(t)\, dt.$$

Глава XIII

# МНОГОМЕРНЫЕ ЧЕБЫШЕВСКИЕ НЕРАВЕНСТВА

## § 1. Введение

Главный принцип, который лежит в основе вывода многомерных неравенств, есть по-прежнему метод теоремы 2.1 гл. XII. При получении конкретных результатов технические детали более громоздки, как это обычно бывает при переходе к случаю нескольких переменных.

На протяжении этой главы $\mathbf{x}$ будет обозначать $n$-мерный вектор. Рассмотрим пространство моментов, образованное функциями

$$\begin{gathered} u_0(\mathbf{x}) = 1, \qquad u_i(\mathbf{x}) = x_i, \qquad i = 1, \ldots, n, \\ u_{ij}(\mathbf{x}) = x_i x_j, \qquad 1 \leqslant i \leqslant j \leqslant n. \end{gathered} \tag{1.1}$$

В этой ситуации пространство моментов имеет размерность $(n+1)(n+2)/2$, а $T$ отождествляется с $R^n$.

Рассмотрим множество $V(\mathbf{c})$ распределений или мер, которое связано с моментной точкой

$$\mathbf{c} = (c_0, c_1, \ldots, c_n, \pi_{11}, \pi_{12}, \ldots, \pi_{nn}), \tag{1.2}$$

и положим, что $c_0 = 1$, $c_i = 0$ $(i = 1, \ldots, n)$, а матрица $\mathbf{\Pi} = \| \pi_{ij} \|$ положительно определена. Это значит, что мы ограничиваем внимание множеством всех распределений $\sigma$, для которых

$$\begin{gathered} \int x_i \, d\sigma(\mathbf{x}) = E(X_i) = 0, \\ \int x_i x_j \, d\sigma(\mathbf{x}) = E(X_i X_j) = \pi_{ij}, \qquad i, j = 1, \ldots, n. \end{gathered}$$

Мы утверждаем, что (1.2) определяет внутреннюю точку пространства моментов. Допустим, напротив, что это не так. Тогда можно найти нетривиальный многочлен (т. е. «многочлен» по отношению к функциям (1.1)), удовлетворяющий соотношениям

$$a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i,j=1}^{n} a_{ij} x_i x_j \geqslant 0 \tag{1.3}$$

для всех $x_1, \ldots, x_n$ и

$$a_0 + \sum_{i,j=1}^{n} a_{ij}\pi_{ij} = 0. \qquad (1.4)$$

Но из (1.3) следует, что $a_0 \geqslant 0$ (подставьте $x_i = 0$, $i=1, \ldots, n$) и что $\mathbf{A} = \| a_{ij} \|$ положительно полуопределена. Более того, так как $\mathbf{\Pi} = \| \pi_{ij} \|$ положительно определена, а $\mathbf{A}$ полуопределена, то $\operatorname{tr}\{\mathbf{A\Pi}\} \geqslant 0$ и равенство достигается, только если $\mathbf{A} \equiv 0$ (через $\operatorname{tr}\mathbf{A}$ обозначается, как обычно, след матрицы $\mathbf{A}$).

Принимая во внимание эти факты и (1.4), мы заключаем, что $a_0 = 0$ и $\mathbf{A} \equiv 0$. Но в этом случае неравенство (1.3) возможно, только если $a_i = 0$, $i = 1, \ldots, n$. Из этого противоречия следует, что (1.2) образует, как утверждалось, внутреннюю точку пространства моментов.

Пусть $\Omega(\mathbf{x})$ обозначает произвольную неотрицательную симметричную функцию (т. е. $\Omega(\mathbf{x}) = \Omega(-\mathbf{x})$), которая ограничена в компактной области, и пусть $\Omega(\mathbf{x})/(\mathbf{Ax}, \mathbf{x}) \to 0$ при $|\mathbf{x}|^2 = \sum_{i=1}^{n} x_i^2 \to \infty$ для любого $\mathbf{A} > 0$ (т. е. $\mathbf{A}$ положительно определена)[1]. Наша настоящая цель состоит в определении точных границ для $\int \Omega(\mathbf{x}) d\sigma(\mathbf{x})$ при $\sigma \in V(\mathbf{c})$. Исследование, как и ранее, основывается на общем методе теоремы 2.1 гл. XII.

Рассмотрим произвольный многочлен, удовлетворяющий условию

$$a_0 + \sum_{i=1}^{n} a_i x_i + \sum_{i,j=1}^{n} a_{ij} x_i x_j \geqslant \Omega(\mathbf{x}), \qquad (\mathbf{x} \in R^n). \qquad (1.5)$$

Также имеем

$$a_0 - \sum_{i=1}^{n} a_i x_i + \sum_{i,j=1}^{n} a_{ij} x_{ij} \geqslant \Omega(\mathbf{x}) \qquad (\mathbf{x} \in R^n) \qquad (1.6)$$

вследствие симметрии $\Omega(\mathbf{x})$. Складывая (1.5) и (1.6), получим

$$a_0 + \sum a_{ij} x_i x_j \geqslant \Omega(\mathbf{x}) \qquad (\mathbf{x} \in R^n). \qquad (1.7)$$

Более того, математические ожидания величин, стоящих слева в (1.5) — (1.7) по $\sigma \in V(\mathbf{c}^0)$, совпадают.

Это означает, что без потери общности мы можем ограничить класс многочленов многочленами вида (1.7). Ясно, что из (1.7) следует, что $a_0 \geqslant 0$ (достаточно положить $x_i = 0$) и что $\mathbf{A} = \| a_{ij} \|$ положительно полуопределена. Более того, теорема 2.1 гл. XII

[1]) $(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} x_i y_i$ обозначает обычное скалярное произведение векторов $\mathbf{x}$ и $\mathbf{y}$.

утверждает, что

$$\inf [a_0 + \operatorname{tr} \{\mathbf{A\Pi}\}] = \sup_{\sigma \in V(\mathbf{c})} \int \Omega(\mathbf{x})\, d\sigma(\mathbf{x}) = I_{\max}, \tag{1.8}$$

где нижняя грань берется по множеству всех многочленов, удовлетворяющих (1.7).

В особом случае, когда $\Omega(\mathbf{x})$ есть характеристическая функция симметричного множества $\Gamma$, возможно следующее упрощение. В этом случае, если значение (1.8) строго меньше единицы, мы можем взят $a_0$ равным нулю. В самом деле, допустим $I_{\max} < 1$, и пусть $a_0$ и $\mathbf{A}_0$ таковы, что $I_{\max} = a_0 + \operatorname{tr} \mathbf{A}_0\mathbf{\Pi}$. Ясно, что $0 \leqslant a_0 < 1$. Следовательно, так как $(\mathbf{A}_0\mathbf{x}, \mathbf{x}) + a_0 \geqslant \Omega(\mathbf{x})$ и $\Omega(\mathbf{x}) = 0$ или 1, мы заключаем, что

$$\frac{(\mathbf{A}_0\mathbf{x}, \mathbf{x})}{1 - a_0} \geqslant \Omega(\mathbf{x}).$$

Кроме того, если $I_{\max} = a_0 + \operatorname{tr} \mathbf{A}_0\mathbf{\Pi} < 1$, мы можем заключить, что

$$\frac{\operatorname{tr} \mathbf{A}_0\mathbf{\Pi}}{1 - a_0} \leqslant a_0 + \operatorname{tr} \mathbf{A}_0\mathbf{\Pi}$$

и неравенство будет строгим, если $a \neq 0$. Таким образом, $a_0$ может быть выбрано равным нулю, как утверждалось.

Определим

$$\mathcal{A} = \{\mathbf{A} \mid (\mathbf{A}\mathbf{x}, \mathbf{x}) \geqslant \Omega(\mathbf{x}),\ \ \mathbf{x} \in R^n,\ \ \mathbf{A} > 0\} \tag{1.9}$$

и суммируем полученные результаты в следующей теореме.

Теорема 1.1. (1) *Пусть* $\Omega(\mathbf{x})$ — *характеристическая функция симметричного множества* $\Gamma$, *и пусть* $\mathbf{c}$ *задано соотношением* (1.2). *Тогда*

$$\sup_{\sigma \in V(\mathbf{c})} \int \Omega(\mathbf{x})\, d\sigma(\mathbf{x}) \leq \min \operatorname{tr} \{\mathbf{A\Pi}\}. \tag{1.10}$$

*где минимум берется по всем* $\mathbf{A} \in \mathcal{A}$. *Равенство в* (1.10) *возникает всякий раз, когда правая часть не превосходит единицы.*

(2) *Если* $\Omega(\mathbf{x})$ *есть произвольная неотрицательная симметрическая функция, для которой* $\Omega(\mathbf{x})/|\mathbf{x}|^2 \to 0$ $(|\mathbf{x}| \to \infty)$ *и* $\Omega(\mathbf{x})$ *ограничена на компактной области, то*

$$\sup_{\sigma \in V(\mathbf{c})} \int \Omega(\mathbf{x})\, d\sigma(\mathbf{x}) = \min_{\mathcal{P}} [a + \operatorname{tr} \mathbf{A\Pi}],$$

*где* $\mathcal{P}$ *состоит из многочленов, для которых*

$$a + (\mathbf{A}\mathbf{x}, \mathbf{x}) \geqslant \Omega(\mathbf{x})$$

*для всех* $\mathbf{x} = (x_1, \ldots, x_n)$.

Эта теорема была независимо и одновременно открыта Олкином и Праттом [1958] и Уиттлом [1958 b]. Частные случаи были ранее изучены Берге [1937] и Лалом [1955]. Наш анализ базируется на геометрии пространств моментов, исследуемой общим методом теоремы 1.2 гл. XII.

Замечание 1.1. На протяжении этой главы мы изучаем чебышевские неравенства при условии, что все моменты (1.2) заданы. Есть исследования для случая, когда задана только часть моментов; см., например, работу Маршалла [1960]. Конечно, многочлены в этом случае понимаются как линейные комбинации только тех функций, моменты которых заданы. В этой связи предположим, что мы не задаем полного множества $\mathbf{c} = (c_0, c_1, \ldots, c_n, \pi_{11}, \pi_{12}, \ldots \ldots, \pi_{nn})$, как в (1.2) с $c_i = 0$, $i = 1, \ldots, n$, а задаем только $\mathbf{c}^* = (c_0, \pi_{11}, \pi_{12}, \ldots, \pi_{nn})$, т. е. моменты второго порядка. Пространство моментов, которое теперь рассматривается, образовано функциями $u_0(\mathbf{x}) = 1$, $u_{ij}(\mathbf{x}) = x_i x_j$, $1 \leqslant i \leqslant j \leqslant n$. Тогда из теоремы 1.2 гл. XII непосредственно следует, что теорема 1.1 настоящего параграфа выполняется, причем условие симметрии, накладываемое на $\Omega(\mathbf{x})$, может быть опущено, а $V(\mathbf{c})$ заменяется на $V(\mathbf{c}^*)$. Теорема 1.1 утверждает, что при условии симметрии границы в (1.10) и (1.11) являются точными, даже если верхняя грань берется по меньшему классу мер $V(\mathbf{c})$.

Подобные замечания можно сделать также для некоторых дальнейших теорем этой главы.

## § 2. Крайние точки множества $\mathcal{A}$

Мы изучаем далее природу множества $\mathcal{A}$ всех матриц $\mathbf{A}$, для которых

$$(\mathbf{A}\mathbf{x}, \mathbf{x}) \geqslant \Omega(\mathbf{x}), \tag{2.1}$$

где $\Omega(\mathbf{x})$ есть характеристическая функция симметричного множества $\Gamma$, дополнение которого есть выпуклое открытое ограниченное тело.

Непосредственно ясно, что $\mathbf{A} > 0$ и, следовательно, $\mathbf{B} = \mathbf{A}^{-1} > 0$. Охарактеризуем множество всех крайних точек для

$$\mathcal{B} = \{\mathbf{B} = \mathbf{A}^{-1} \mid \mathbf{A} \in \mathcal{A}\}. \tag{2.2}$$

Так как $\Lambda = \Gamma^c$ (через $\Gamma^c$ обозначено дополнение $\Gamma$) есть симметричное открытое ограниченное выпуклое тело, то мы можем считать, что $\bar{\Lambda}$ (через $\bar{\Lambda}$ обозначено замыкание $\Lambda$) — единичная сфера. Пусть $\Lambda^*$ обозначает сопряженную единичную сферу, т. е. положим $\Lambda^* = \{\mathbf{y} \mid |(\mathbf{y}, \mathbf{x})| \leqslant 1$ для всех $\mathbf{x} \in \bar{\Lambda}\}$. Ясно, что $\Lambda^*$ есть также симметричное выпуклое ограниченное тело в $R^n$. Точка $\mathbf{y}$ лежит на границе $\Lambda^*$ тогда и только тогда, когда $(\mathbf{y}, \mathbf{x})^2 \leqslant 1$ для всех $\mathbf{x} \in \bar{\Lambda}$ и равенство достигается для некоторого $\mathbf{x} \in \bar{\Lambda}$.

Следующая лемма дает простой критерий определения принадлежности $\mathbf{B}$ к множеству $\mathcal{B}$.

Лемма 2.1. *Положительно определенная матрица* $\mathbf{B}$ *принадлежит множеству* $\mathcal{B}$ *тогда и только тогда, когда*

$$(\mathbf{B}\mathbf{y}, \mathbf{y}) \leqslant 1 \tag{2.3}$$

*для всех* $\mathbf{y} \in \mathcal{D}\Lambda^*$.

Доказательство. Предположим сначала, что (2.3) выполняется. Равенство

$$\sup_{\mathbf{x} \neq 0} \frac{(\mathbf{y}, \mathbf{x})^2}{(\mathbf{A}\mathbf{x}, \mathbf{x})} = (\mathbf{A}^{-1}\mathbf{y}, \mathbf{y}) = (\mathbf{B}\mathbf{y}, \mathbf{y}) \tag{2.4}$$

верно для любого $\mathbf{y}$. Выберем $\mathbf{x}$ произвольно на границе $\Lambda$ и определим $\mathbf{y}_\mathbf{x} \in \mathcal{D}\Lambda^*$, удовлетворяющее $(\mathbf{y}_\mathbf{x}, \mathbf{x})^2 = 1$. Сравнивая (2.3) и (2.4), мы заключаем, что $(\mathbf{A}\mathbf{x}, \mathbf{x}) \geqslant 1$, и, следовательно, $\mathbf{A} \in \mathcal{A}$ или, эквивалентно, $\mathbf{B} \in \mathcal{B}$.

Наоборот, если $\mathbf{A} \in \mathcal{A}$, то $(\mathbf{A}\mathbf{x}, \mathbf{x}) \geqslant 1$ для любого $\mathbf{x} \in \mathcal{D}\Lambda$. Известно также, что для всякого $\mathbf{y} \in \mathcal{D}\Lambda^*$, $(\mathbf{x}, \mathbf{y})^2 \leqslant 1$ для всех $x \in \Lambda$. Следовательно, для $\mathbf{x} \in \mathcal{D}\Lambda$

$$\frac{(\mathbf{x}, \mathbf{y})^2}{(\mathbf{A}\mathbf{x}, \mathbf{y})} \leqq 1$$

и, используя свойство однородности, мы выводим (2.3) из (2.4).

Простым следствием характеризации (2.3) множества $\mathcal{B}$ является

Следствие 2.1. *Множество* $\mathcal{B}$ *выпукло и ограничено.*

Мы теперь сделаем отступление, касающееся некоторой задачи максимизации, решение которой потребуется в дальнейшем.

Пусть $\mathbf{P}$ и $\mathbf{Q}$ — матрицы, причем $\mathbf{P}$ положительно определена, а $\mathbf{Q}$—симметрическая. Рассмотрим задачу максимизации $(\mathbf{Q}\mathbf{x}, \mathbf{x})$ при условии, что $(\mathbf{P}\mathbf{x}, \mathbf{x}) = 1$. Ее решение хорошо известно. А именно, $\max(\mathbf{Q}\mathbf{x}, \mathbf{x}) = \lambda$, где $\lambda$ есть максимальное собственное число $\mathbf{Q}$ относительно $\mathbf{P}$. Более того, максимум достигается на любом векторе $\mathbf{x}_0$, который является собственным вектором $\mathbf{Q}$ относительно $\mathbf{P}$, соответствующим собственному числу $\lambda$, т. е. $\mathbf{Q}\mathbf{x}_0 = \lambda\mathbf{P}\mathbf{x}_0$.

Мы применим этот результат к специальному случаю, когда $\mathbf{Q}$ — матрица ортогонального проектирования (ее ранг равен единице):

$$\mathbf{Q}\mathbf{x} = (\mathbf{x}, \mathbf{y}_0)\,\mathbf{y}_0, \tag{2.5}$$

где $(\mathbf{y}_0, \mathbf{y}_0) = 1$. Тогда $(\mathbf{Q}\mathbf{x}, \mathbf{x}) = (\mathbf{x}, \mathbf{y}_0)^2$.

В этом специальном случае решение задачи на максимум, поставленной выше, следующее. Используя тот факт, что $\mathbf{Q}$—матрица единичного ранга, мы получаем, что $\mathbf{P}\mathbf{x}_0 = \mu\mathbf{y}_0$ или $\mathbf{x}_0 = \mu\mathbf{P}^{-1}\mathbf{y}_0$. Накладывая условие нормировки $(\mathbf{P}\mathbf{x}_0, \mathbf{x}_0) = 1$, находим

$$\mathbf{x}_0 = \frac{\mathbf{P}^{-1}\mathbf{y}_0}{\sqrt{(\mathbf{P}^{-1}\mathbf{y}_0, \mathbf{y}_0)}}, \tag{2.6}$$

так что $\mathbf{x}_0$ определено единственным образом.

Теперь мы подготовлены к описанию крайних точек множества $\mathcal{A}$.

Теорема 2.1. *Если $\Lambda^*$ имеет конечное число крайних точек, т. е. является многогранником, то матрица $\mathbf{A} \in \mathcal{A}$ есть крайняя точка множества $\mathcal{A}$ в том и только в том случае, когда существует $n$ линейно независимых векторов $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$, принадлежащих $\Lambda^*$, таких, что $\mathbf{B} = \mathbf{A}^{-1} \in \mathcal{B}$ удовлетворяет*

$$(\mathbf{B}\mathbf{y}_i, \mathbf{y}_i) = 1, \qquad i = 1, \dots, n, \tag{2.7}$$

*или, эквивалентным образом, существует $n$ линейно независимых векторов $\mathbf{x}_1, \dots, \mathbf{x}_n$, принадлежащих $\Lambda$, таких, что*

$$(\mathbf{A}\mathbf{x}_i, \mathbf{x}_i) = 1, \qquad i = 1, 2, \dots, n. \tag{2.8}$$

Замечание 2.1. Легко видеть, что если $\mathbf{y}$ принадлежит $\Lambda^*$ и удовлетворяет условию $(\mathbf{B}\mathbf{y}, \mathbf{y}) = 1$, то $\mathbf{y}$ является крайней точкой $\Lambda$, так как $(\mathbf{B}\mathbf{y}, \mathbf{y})$ строго выпукло по $\mathbf{y}$.

Замечание 2.2. Отправным пунктом для этой теоремы послужила работа Олкина и Пратта [1958], которые исследовали специальный случай, когда $\Lambda$ является единичным кубом с центром в начале координат. Нетрудно показать на примере, что если $\Lambda^*$ не является многогранником, то часть утверждения теоремы, касающаяся необходимости, неверна. В этом случае приобретает значение степень касания элипсоида, вписанного в $\Lambda$, с поверхностью $\Lambda$. В самом деле, рассмотрим в двумерном случае квадратичную форму, определенную уравнением окружности $x^2 + y^2 = 1$, т. е. $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Пусть $\Lambda$ — симметричная выпуклая область, определенная для достаточно малых положительных $\delta$ и $\varepsilon$ следующими уравнениями:

$$|y| = \begin{cases} g(x), & |x| \leqslant 1 - \varepsilon, \\ l(|x|), & 1 - \varepsilon \leqslant |x| \leqslant x_0, \end{cases}$$

где $g(x) = \sqrt{1 - x^2} + \delta x^4$ и $l(x)$ есть отрезок прямой линии, проходящей через точку $(1 - \varepsilon,\ g(1 - \varepsilon))$ и имеющей наклон $g'(1 - \varepsilon)$, а $x_0$ определяется уравнением $l(x_0) = 0$. Тогда $\mathbf{A}$ — крайняя точка, хотя соответствующий эллипсоид касается $\Lambda$ только один раз.

Доказательство. Достаточность. Пусть $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ принадлежат $\Lambda^*$, линейно независимы и удовлетворяют (2.7), где $\mathbf{B} \in \mathcal{B}$. Предположим, в противоречии с утверждением теоремы, что $\mathbf{A}$ не есть крайняя точка. Тогда мы можем найти представление

$$\mathbf{A} = \frac{\mathbf{A}_1 + \mathbf{A}_2}{2}, \quad \mathbf{A}_i \in \mathcal{A} \quad (i = 1, 2), \quad \mathbf{A}_1 \neq \mathbf{A}_2. \tag{2.9}$$

Для произвольного $\mathbf{x}$ выполняется

$$\frac{(\mathbf{x}, \mathbf{y}_j)^2}{(\mathbf{A}\mathbf{x}, \mathbf{x})} \leqslant \frac{1}{2}\left\{\frac{(\mathbf{x}, \mathbf{y}_j)^2}{(\mathbf{A}_1\mathbf{x}, \mathbf{x})} + \frac{(\mathbf{x}, \mathbf{y}_j)^2}{(\mathbf{A}_2\mathbf{x}, \mathbf{x})}\right\} \qquad (j = 1, \dots, n), \tag{2.10}$$

причем равенство возни ает только в случае

$$(\mathbf{A}\mathbf{x}, \mathbf{x}) = (\mathbf{A}_1\mathbf{x}, \mathbf{x}) = (\mathbf{A}_2\mathbf{x}, \mathbf{x}).$$

Беря верхнюю грань по $\mathbf{y}$, получим

$$1 = (\mathbf{B}\mathbf{y}_j, \mathbf{y}_j) \leqslant \frac{1}{2}\{(\mathbf{B}_1\mathbf{y}_j, \mathbf{y}_j) + (\mathbf{B}_2\mathbf{y}_j, \mathbf{y}_j)\}. \tag{2.11}$$

По лемме 2.1 каждое слагаемое в (2.11) не превосходит единицы, и, следовательно, всюду имеет место равенство.

Более того, принимая во внимание (2.5) и (2.6), мы заключаем что верхние грани

$$\sup_{\mathbf{x}\neq 0}\frac{(\mathbf{x}, \mathbf{y}_j)^2}{(\mathbf{A}\mathbf{x}, \mathbf{x})}, \quad \sup_{\mathbf{x}\neq 0}\frac{(\mathbf{x}, \mathbf{y}_j)^2}{(\mathbf{A}_1\mathbf{x}, \mathbf{x})}, \quad \sup_{\mathbf{x}\neq 0}\frac{(\mathbf{x}, \mathbf{y}_j)^2}{(\mathbf{A}_2\mathbf{x}, \mathbf{x})}$$

достигаются на одном и том же векторе. Так как $(\mathbf{y}_j, \mathbf{B}\mathbf{y}_j) = (\mathbf{y}_j, \mathbf{B}_1\mathbf{y}_j) = (\mathbf{y}_j, \mathbf{B}_2\mathbf{y}_j) = 1$, то из (2.11) получаем, что $\mathbf{B}\mathbf{y}_j = \mathbf{B}_1\mathbf{y}_j = \mathbf{B}_2\mathbf{y}_j$, $j = 1, \ldots, n$. Из линейной независимости векторов $\mathbf{y}_1, \ldots, \mathbf{y}_n$ следует $\mathbf{B}_1 = \mathbf{B}_2$, и, следовательно, $\mathbf{A}_1 = \mathbf{A}_2$. Полученное противоречие доказывает требуемый результат.

Необходимость. Допустим, что множество $E$ всех векторов $\mathbf{y} \in \Lambda^*$, удовлетворяющих условию $(\mathbf{B}\mathbf{y}, \mathbf{y}) = 1$, образует пространство размерности меньше чем $n$. Как было замечено ранее, $E$ содержит только крайние точки $\Lambda^*$. Определим $\tilde{E}$ как любое $(n-1)$-мерное пространство, содержащее $E$. Пусть $\mathbf{z}$ — вектор, ортогональный к $\tilde{E}$. Определим проекцию отображения $\mathbf{P}$ на направление $\mathbf{z}$.

А именно, для любого вектора $\boldsymbol{\omega}$ имеем единственное разложение $\boldsymbol{\omega} = a\mathbf{z} + \mathbf{u}$, где $\mathbf{u} \in \tilde{E}$. Положим $\mathbf{P}\boldsymbol{\omega} = a\mathbf{z}$. Сформируем $\tilde{\mathbf{B}} = \mathbf{B} + \eta\mathbf{P}$, где $\eta$ — положительное и достаточно малое — будет определено подходящим образом далее. Рассмотрим произвольную крайнюю точку $\tilde{\mathbf{y}} \in \Lambda^*$, $\tilde{\mathbf{y}} \notin E$. Для $\tilde{\mathbf{y}}$ имеет место разложение $\tilde{y} = a\mathbf{z} + \mathbf{u}$. Отсюда следует, что $(\tilde{\mathbf{B}}\tilde{\mathbf{y}}, \tilde{\mathbf{y}}) = (\mathbf{B}\tilde{\mathbf{y}}, \tilde{\mathbf{y}}) + \eta a^2(\mathbf{P}\mathbf{z}, \mathbf{z})$.

Так как $(\mathbf{B}\tilde{\mathbf{y}}, \tilde{\mathbf{y}}) < 1$, мы можем определить $\eta$ положительным и достаточно малым, так что выполняется

$$(\tilde{\mathbf{B}}\tilde{\mathbf{y}}, \tilde{\mathbf{y}}) \leqslant 1. \tag{2.12}$$

Множество всех крайних точек $\tilde{\mathbf{y}}$ в $\Lambda^* \cap E^c$, по предположению, конечно. Отсюда следует, что $\eta > 0$ может быть выбрано так, что (2.12) выполняется для всех крайних точек $\tilde{\mathbf{y}} \in \Lambda^* \cap E^c$. По построению $(\tilde{\mathbf{B}}\mathbf{y}, \mathbf{y}) = (\mathbf{B}\mathbf{y}, \mathbf{y}) = 1$ для $\mathbf{y} \in E$. Далее $\Lambda^*$ порождается выпуклыми комбинациями своих крайних точек, и, следовательно, (2.12) выполняется для всех $\mathbf{y} \in \Lambda^*$.

Очевидно,

$$\tilde{\mathbf{B}} = \mathbf{B} + \eta\mathbf{P} > \mathbf{B}, \tag{2.13}$$

откуда

$$\widetilde{\mathbf{A}} = \widetilde{\mathbf{B}}^{-1} < \mathbf{A} = \mathbf{B}^{-1}. \qquad (2.14)$$

Далее

$$\mathbf{A} = \frac{\widetilde{\mathbf{A}} + (2\mathbf{A} - \widetilde{\mathbf{A}})}{2}. \qquad (2.15)$$

С помощью (2.13) и (2.14) легко проверить, что $\widetilde{\mathbf{A}}$ и $2\mathbf{A} - \widetilde{\mathbf{A}}$ различны и принадлежат $\mathcal{A}$. Соотношение (2.15) представляет $\mathbf{A}$ как выпуклую комбинацию двух различных элементов $\mathcal{A}$, и значит, $\mathbf{A}$ не есть крайняя точка. Доказательство необходимости закончено.

Замечание 2.3. Мы можем следующим образом сформулировать утверждение теоремы 2.1 в геометрических терминах. Условие $(\mathbf{A}\mathbf{x}, \mathbf{x}) \geqslant \Omega(\mathbf{x})$ эквивалентно утверждению о том, что эллипсоид $(\mathbf{A}\mathbf{x}, \mathbf{x}) = 1$ вписывается в выпуклое тело $\Lambda$. Теорема 2.1 утверждает, что $\mathbf{A}$ является крайней точкой тогда и только тогда, когда ассоциированный эллипсоид касается $\Lambda$ вдоль по крайней мере $n$ независимых направлений. Здесь мы снова обращаем внимание читателя на замечание 2.2, подчеркивающее то обстоятельство, что обратное утверждение для произвольных выпуклых тел неверно.

## § 3. Чебышевские неравенства для $\Pr\left(\max\limits_{1\leqslant i\leqslant n} |X_i| \geqslant 1\right)$

Следующие два параграфа посвящены нескольким важным примерам многомерных неравенств, относящихся к специальному выбору $\Omega(x)$. Мы продолжим наше обсуждение для случая $\Omega(\mathbf{x})$ общего вида в § 5.

В данном параграфе мы применяем предшествующие результаты к определению верхней границы для

$$\Pr\left(\max_{1\leqslant i\leqslant n} |Y_i|/k_i\sigma_i \geqslant 1\right), \qquad (3.1)$$

где $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ есть случайный вектор со средним 0 и неособенной ковариационной матрицей $\mathbf{\Sigma} = \|\sigma_{ij}\|$, а $\sigma_i^2 = \sigma_{ii}$. Формулировка задачи упрощается, если сделать замену $X_i = Y_i/k_i\sigma_i$. Случайный вектор $\mathbf{X} = (X_1, \ldots, X_n)$ имеет среднее нуль и ковариационную матрицу $\mathbf{\Pi} = \mathbf{K}^{-1}\mathbf{R}\mathbf{K}^{-1}$, где $\mathbf{R} = \|\rho_{ij}\|$ — корреляционная матрица $\mathbf{Y}$, $\pi_{ij} = \sigma_{ij}/(\sigma_i\sigma_j k_i k_j) = \rho_{ij}/(k_i k_j)$, а $\mathbf{K}$ — диагональная матрица с элементами $k_1, \ldots, k_n$. После этой замены переменных (3.1) принимает вид

$$\Pr\left(\max_{1\leqslant i\leqslant n} |Y_i|/k_i\sigma_i \geqslant 1\right) = \Pr\left(\max_{1\leqslant i\leqslant n} |X_i| \geqslant 1\right). \qquad (3.2)$$

В данном случае множество $\Lambda$ (см. § 2) имеет вид

$$\Lambda = \{\mathbf{x} \mid \max_{1\leqslant i\leqslant n} |x_i| < 1\}. \qquad (3.3)$$

При таком определении $\Lambda$ легко проверить, что $\Lambda^* = \{\mathbf{y} \mid \Sigma |y_i| \leqslant 1\}$.

Крайние точки $\Lambda^*$, очевидно, есть вектора $(0, \ldots, 0, 1, 0, \ldots, 0)$ с одной единичной компонентной и остальными нулевыми. Применение теоремы 2.1 дает следующий результат.

Теорема 3.1. (Олкин и Пратт [1958] и Уиттл [1958b].) *Если* $\Lambda$ *задано посредством* (3.3), *то* $\mathbf{A} \in \mathcal{A}$ (*см.* (1.9)) *является крайней точкой тогда и только тогда, когда матрица* $\mathbf{B} = \mathbf{A}^{-1}$ *положительно определена и* $b_{ii} = 1$ $(i = 1, 2, \ldots, n)$.

Подкласс $\mathcal{B}$ (см. (2.2)), удовлетворяющий условиям теоремы 3.1, обозначим через $\mathcal{B}_0 \cdot \mathcal{B}_0$, очевидно, является выпуклым, ограниченным и имеет внутренность. По теореме 1.1 мы имеем

$$\Pr\left(\max_{1 \leqslant i \leqslant n} |X_i| \geqslant 1\right) \leqslant \min \operatorname{tr} \mathbf{A}\Pi,$$

где минимум берется по множеству $\mathcal{A}$, образованному матрицами $\mathbf{A}$, удовлетворяющими условию

$$(\mathbf{A}\mathbf{x}, \mathbf{x}) \geqslant \Omega(\mathbf{x}), \tag{3.4}$$

где $\Omega(\mathbf{x})$ — есть характеристическая функция множества $\Lambda^c$, а $\Lambda$ определено в (3.3). Из неравенства (3.4), очевидно, следует, что матрица $\mathbf{A}$ положительно определена.

Так как $\operatorname{tr} \mathbf{A}\Pi$ является линейной по $\mathbf{A}$, то ее минимум над множеством $A$ достигается на крайней точке. Множество крайних точек полностью охарактеризовано в теореме 3.1. Минимизирующая матрица $\mathbf{A}$ или $\mathbf{B} = \mathbf{A}^{-1}$ получает более точное выражение, когда $\mathbf{A}$ определяется соотношением (3.3). Мы приступаем к задаче получения этой дополнительной структуры. Для этой цели нам потребуется две следующие леммы.

Лемма 3.1. *Функция* $\operatorname{tr} \mathbf{B}^{-1}\mathbf{\Pi}$ *строго выпукла по* $\mathbf{B} \in \mathcal{B}$ *и имеет единственный минимум, расположенный во внутренней точке* $\mathbf{B}_0 \in \mathcal{B}_0$.

Замечание 3.1. Трудно приписать лемму 3.1 какому-нибудь единственному автору. Утверждение о выпуклости возникает во многих контекстах математической статистики и программирования и составило за некоторое время часть фольклора теории матриц.

Доказательство. Доказательство выпуклости может быть основано на известном факте, что $\mathbf{B}^{-1}$ как функция матрицы выпукла, когда $\mathbf{B}$ пробегает множество положительно определенных матриц в том смысле, что $\lambda \mathbf{B}_1^{-1} + (1 - \lambda) \mathbf{B}_2^{-1} \geqslant (\lambda \mathbf{B}_1 + (1 - \lambda) \mathbf{B}_2)^{-1}$ при $0 < \lambda < 1$ и равенство достигается тогда и только тогда, когда $\mathbf{B}_1 = \mathbf{B}_2$. Умножение слева и справа на $\mathbf{\Pi}^{1/2}$ ничего не изменяет. Вычислив след от обеих частей, получим, что функция $\operatorname{tr} \mathbf{B}^{-1}\mathbf{\Pi}$ строго выпукла, а матрица $\mathbf{B}_0$, на которой достигается минимум, единственна.

То обстоятельство, что минимум $\operatorname{tr} \mathbf{B}^{-1}\mathbf{\Pi}$ расположен во внутренней точке $\mathscr{B}$, вытекает из соотношения

$$\operatorname{tr} \mathbf{B}^{-1}\Pi \geqslant \nu \operatorname{tr} \mathbf{B}^{-1}, \tag{3.5}$$

где $\nu$ — наименьшее собственное число $\mathbf{\Pi}$.

Действительно, из (3.5) следует, что $\operatorname{tr} \mathbf{B}^{-1}\mathbf{\Pi} \to \infty$, если $\mathbf{B}$ приближается к границе $\mathscr{B}_0$ и становится особенной.

Следующая лемма также потребуется нам далее и вместе с тем имеет некоторый самостоятельный интерес. Она обобщает результат Олкина и Пратта (см. следствие 3.2 ниже).

Лемма 3.2. *Пусть* $y_1, \dots, y_m$ $(n \leqslant m \leqslant n(n+1)/2)$ *—векторы, среди которых есть по крайней мере $n$ независимых, и пусть*

$$\tilde{\mathscr{B}}_0 = \{\mathbf{B} \mid \mathbf{B} > 0,\ (\mathbf{B}y_i, y_i) = 1,\ i = 1, \dots, m\}. \tag{3.6}$$

*Допустим, что $\mathscr{B}_0$ не пусто (это заведомо так в случае $m = n$). Если*

$$\mathscr{E} = \{\mathbf{E} \mid \mathbf{E} \text{ симметрично},\ (\mathbf{E}y_i, y_i) = 0,\ i = 1, \dots, m\},$$

*то существуют единственное $\mathbf{B}_0 \in \mathscr{B}_0$ и $\mathbf{\Delta}_0$, такие, что*

$$\operatorname{tr} \mathbf{B}\mathbf{\Delta}_0 = 0 \tag{3.7}$$

*для всех* $\mathbf{E} \in \mathscr{E}$ *и*

$$\mathbf{\Pi} = \mathbf{B}_0\mathbf{\Delta}_0\mathbf{B}_0. \tag{3.8}$$

Доказательство. Множество $\tilde{\mathscr{B}}_0$, очевидно, выпукло и ограничено и минимум

$$\min_{\tilde{\mathscr{B}}_0} \operatorname{tr} \mathbf{B}^{-1}\mathbf{\Pi}$$

достигается в единственной внутренней точке $\mathbf{B}_0$ в $\tilde{\mathscr{B}}_0$. Пусть $\tilde{\mathscr{B}}_1$ — компактное выпуклое подмножество $\tilde{\mathscr{B}}_0$, которое содержит $\mathbf{B}_0$ в своей внутренности, и пусть

$$\tilde{\mathscr{C}}_1 = \{\mathbf{C} \mid \mathbf{C} = \mathbf{B}^{-1}\mathbf{\Pi}^{1/2},\ \mathbf{B} \in \tilde{\mathscr{B}}_1\}.$$

Рассмотрим теперь игру (см. замечание 2.1 гл. X) с ядром

$$K(\mu, \mathbf{B}) = \int \frac{\operatorname{tr}^2 \mathbf{\Pi}^{1/2}\mathbf{C}}{\operatorname{tr} \mathbf{C}'\mathbf{B}\mathbf{C}}\, d\mu(\mathbf{C}),$$

где $\mu$ — вероятностная мера на $\tilde{\mathscr{C}}_1$ и $\mathbf{B}$ пробегает множество $\tilde{\mathscr{B}}_1$. Игрок I должен максимизировать это ядро по $\mu$, а игрок II — минимизировать по $\mathbf{B} \in \tilde{\mathscr{B}}_1$. Ядро $K(\mu, \mathbf{B})$ строго выпукло по $\mathbf{B}$ и линейно по $\mu$. Цена этой игры равна

$$\inf_{\mathbf{B}} \sup_{\mu} K(\mu, \mathbf{B}) = \inf_{\mathbf{B}} \sup_{\mathbf{C}} \frac{\operatorname{tr}^2 \mathbf{\Pi}^{1/2}\mathbf{C}}{\operatorname{tr} \mathbf{C}'\mathbf{B}\mathbf{C}} = \inf_{\mathbf{B}} \operatorname{tr} \mathbf{B}^{-1}\mathbf{\Pi} = \operatorname{tr}\mathbf{B}_0^{-1}\mathbf{\Pi}.$$

Игрок II имеет единственную оптимальную стратегию $\mathbf{B}_0$, так что игрок I имеет единственную оптимальную стратегию $\mu_0$, причем эта мера сосредоточена в точке $\mathbf{C}_0 = \mathbf{B}_0^{-1}\mathbf{\Pi}^{1/2}$.

В этом случае

$$\frac{\operatorname{tr}^2 \mathbf{\Pi}^{1/2}\mathbf{C}_0}{\operatorname{tr} \mathbf{C}_0'\mathbf{B}\mathbf{C}_0} \geqslant \operatorname{tr} \mathbf{B}_0^{-1}\mathbf{\Pi}$$

для всех $\mathbf{B} \in \tilde{\mathscr{B}}_1$, так что

$$\operatorname{tr} \mathbf{\Pi}\mathbf{B}_0^{-} \geqslant \mathbf{B}(\mathbf{B}_0^{-1}\mathbf{\Pi}\mathbf{B}_0^{-1}) \tag{3.9}$$

для всех $\mathbf{B} \in \tilde{\mathscr{B}}_1$. Это показывает, что $\mathbf{B} = \mathbf{B}_0$ максимизирует функционал $\operatorname{tr} \mathbf{B}(\mathbf{B}_0^{-1}\mathbf{\Pi}\mathbf{B}_0^{-1})$ над множеством $\tilde{\mathscr{B}}_1$.

Если $\mathbf{B} = \mathbf{B}_0 + \varepsilon\mathbf{E}$, где $\mathbf{E} \in \mathscr{E}$, то $\mathbf{B}$ принадлежит $\tilde{\mathscr{B}}_1$ при условии, что $|\varepsilon|$ достаточно мало, откуда можно заключить, что $\operatorname{tr} \mathbf{E}(\mathbf{B}_0^{-1}\mathbf{\Pi}\mathbf{B}_0^{-1}) = 0$. Поэтому, $\Delta_0 = \mathbf{B}_0^{-1}\mathbf{\Pi}\mathbf{B}_0^{-1}$ удовлетворяет (3.7), и мы имеем представление $\mathbf{\Pi} = \mathbf{B}_0\Delta_0\mathbf{B}_0$, как и утверждает лемма. Чтобы доказать единственность представления (3.8), допустим, что $\mathbf{\Pi} = \mathbf{B}_1\Delta_1\mathbf{B}_1$, где $\mathbf{B}_1 \in \tilde{\mathscr{B}}_0$ и $\operatorname{tr} \mathbf{E}\Delta_1 = \operatorname{tr} \mathbf{E}(\mathbf{B}_1^{-1}\mathbf{\Pi}\mathbf{B}_1^{-1}) = 0$ для всех $\mathbf{E} \in \mathscr{E}$. Но для $\mathbf{B} \in \mathscr{B}_0$, $\mathbf{B} - \mathbf{B}_1$ принадлежит $\mathscr{E}$, так что $\operatorname{tr} \mathbf{\Pi}\mathbf{B}_1^{-1} = \operatorname{tr} \mathbf{B}\mathbf{B}_1^{-1}\mathbf{\Pi}\mathbf{B}_1^{-1}$ или

$$\frac{\operatorname{tr} \mathbf{\Pi}\mathbf{B}_1^{-1}}{\operatorname{tr} \mathbf{B}\mathbf{B}_1^{-1}\mathbf{\Pi}\mathbf{B}_1^{-1}} = \operatorname{tr} \mathbf{B}_1^{-1}\mathbf{\Pi} \geqslant \operatorname{tr} \mathbf{B}_0^{-1}\mathbf{\Pi}$$

для всех $\mathbf{B} \in \tilde{\mathscr{B}}_0$, так как $\mathbf{B}_0$ минимизирует $\operatorname{tr} \mathbf{B}^{-1}\mathbf{\Pi}$ над $\tilde{\mathscr{B}}_0$. В этом случае $\mathbf{C}_1 = \mathbf{B}_1^{-1}\mathbf{\Pi}^{1/2}$ есть оптимальная стратегия для игрока I, и, таким образом, в силу единственности мы заключаем, что $\mathbf{B}_1 = \mathbf{B}_0$. Тогда, очевидно, $\Delta_0 = \Delta_1$.

В особом случае, когда $\tilde{\mathscr{B}}_0 = \mathscr{B}_0$ (см. стр. 505) лемма 3.2 имеет следующий вид.

Следствие 3.2. (1) *Для всякой положительно определенной матрицы* $\mathbf{\Pi}$ *существует единственная положительно определенная матрица* $\mathbf{B}_0$ *с единицами на главной диагонали такая, что матрица* $\mathbf{B}_0^{-1}\mathbf{\Pi}\mathbf{B}_0^{-1}$ *диагональна.*

(2) *Минимум* $\operatorname{tr} \mathbf{B}^{-1}\mathbf{\Pi}$ *над* $\tilde{\mathscr{B}}_0$ *достигается на единственной матрице* $\mathbf{B}_0$ *такой, что матрица* $\mathbf{B}_0^{-1}\mathbf{\Pi}\mathbf{B}_0^{-1}$ *диагональна.*

Доказательство. (1) Если $y_i$ $(i = 1, \ldots, n)$ — вектор с $i$-й единичной компонентой и нулевыми остальными, то $\mathscr{E}$, очевидно, состоит из всех симметричных матриц с нулевыми диагональными элементами. В этом случае матрица $\Delta_0 = \mathbf{B}_0^{-1}\mathbf{\Pi}\mathbf{B}_0^{-1}$ диагональна.

(2) По лемме 3.2 минимум $\operatorname{tr} \mathbf{B}^{-1}\mathbf{\Pi}$ имеет место для единственной матрицы $\mathbf{B}_0$. Диагональная форма матрицы также вытекает из этой

леммы. Соединяя теорему 1.1 и следствие 3.2, получаем следующую теорему.

Теорема 3.2 (Олкин и Пратт [1958]). *Пусть $X$ — случайный вектор с нулевым средним и ковариационной матрицей* $\mathbf{\Pi}$. *Если* $\mathbf{B}_0$ — *единственная положительно определенная матрица с единицами на главной диагонали, такая, что матрица* $\mathbf{B}_0^{-1}\mathbf{\Pi}\mathbf{B}_0^{-1}$ *диагональна, то*

$$\Pr(\max_{1\leqslant i\leqslant n}|X_i|\geqslant 1)\leqslant \operatorname{tr}\mathbf{B}_0^{-1}\mathbf{\Pi}=\operatorname{tr}\mathbf{B}_0^{-1}\mathbf{\Pi}\mathbf{B}_0^{-1}, \qquad (3.10)$$

*и равенство достигается в* (3.10) *тогда и только тогда, когда*

$$\operatorname{tr}\mathbf{B}_0^{-1}\mathbf{\Pi}\leqslant 1.$$

Доказательство. Единственная часть теоремы, которая еще не доказана, есть утверждение: $\operatorname{tr}\mathbf{B}_0^{-1}\mathbf{\Pi}=\operatorname{tr}\mathbf{B}_0^{-1}\mathbf{\Pi}\mathbf{B}_0^{-1}$. Однако если $\mathbf{B}_0$ имеет единицы на главной диагонали и $\mathbf{B}_0^{-1}\mathbf{\Pi}\mathbf{B}_0^{-1}$ — диагональная матрица, то, очевидно,

$$\operatorname{tr}\mathbf{B}_0^{-1}\mathbf{\Pi}\mathbf{B}_0^{-1}=\operatorname{tr}\mathbf{B}_0^{-1}\mathbf{\Pi}\mathbf{B}_0^{-1}\mathbf{B}_0=\operatorname{tr}\mathbf{B}_0^{-1}\mathbf{\Pi}.$$

Пусть $\Theta=\mathbf{B}_0^{-1}\mathbf{\Pi}\mathbf{B}_0^{-1}$ и $\theta_1,\theta_2,\ldots,\theta_n$ обозначают диагональные элементы матрицы $\Theta$. Единственное распределение, достигающее верной границы в (3.10), может быть выражено через $\theta_1,\theta_2,\ldots,\theta_n$ следующим образом.

Теорема 3.3 (Олкин и Пратт [1958]). (1) *Если* $\sum_{i=1}^{n}\theta_i\leqslant 1$, *то равенство достигается в* (3.10) *тогда и только тогда, когда*

$$\Pr(\mathbf{X}=\mathbf{b}^i)=\Pr(\mathbf{X}=-\mathbf{b}^i)=\frac{\theta_i}{2},\qquad i=1,2,\ldots,n,$$
$$\Pr(\mathbf{X}=0)=1-\sum_{i=1}^{n}\theta_i, \qquad (3.11)$$

*где векторы* $\mathbf{b}^1,\ldots,\mathbf{b}^n$ *есть столбцы матрицы* $\mathbf{B}_0$.

(2) *Если* $\sum_{i=1}^{n}\theta_i>1$, *то*

$$\Pr(\max_{1\leqslant i\leqslant n}|X_i|\geqslant 1)=1$$

*при условии, что*

$$\Pr(\mathbf{X}=s\mathbf{b}^i)=\Pr(\mathbf{X}=-s\mathbf{b}^i)=\frac{\theta_i}{2s^2}, \qquad (3.12)$$

*где* $s^2=\sum_{i=1}^{n}\theta_i$.

Доказательство. Равенство в (3.10) достигается как на распределении, задаваемом (3.11), так и на распределении, задаваемом (3.12), и $EX=0$,

$EXX' = \Pi$. Следовательно, утверждение (2) и часть, касающаяся необходимости в утверждении (1), верны. Доказательство достаточности в п. (1) элементарно и частично опирается на замечание 2.2 гл. XII. Мы опускаем технические детали (см. Олкин и Пратт [1958]).

## § 4. Обобщения неравенства Берге

В этом параграфе мы даем конкретное применение теоремы 1.1. Для любого случайного вектора $\mathbf{X} = (X_1, X_2, \ldots, X_n)$, удовлетворяющего условиям $E\mathbf{X} = 0$ и $E\mathbf{XX}' = \mathbf{\Pi}$ теоремы 1.1, следует, что

$$\Pr\left(\max_{1\leqslant i\leqslant n} |X_i| \geqslant 1\right) \leqslant \min_{\mathbf{A}\in\mathcal{A}} \operatorname{tr} \mathbf{A\Pi}, \tag{4.1}$$

где $\mathcal{A}$ содержит все положительно определенные матрицы $\mathbf{A}$, для которых $(\mathbf{Ax}, \mathbf{x}) \geqslant 1$ всякий раз, когда $\max\limits_{1\leqslant i\leqslant n} |x_i| \geqslant 1$, $\mathbf{x} = (x_1, \ldots, x_n)$. В § 3 были описаны некоторые характеризации для матриц $\mathbf{A} \in \mathcal{A}$, доставляющих минимум в (4.1).

Есть существенная трудность в получении явной оценки для наилучшей границы в (4.1). Однако мы можем получить явную границу (не обязательно точную) для $\Pr(\max\limits_{1\leqslant i\leqslant n} |X_i| \geqslant 1)$, минимизируя $\operatorname{tr} \mathbf{A\Pi}$ над некоторым подмножеством крайних точек $\mathcal{A}$. Это подмножество состоит из матриц $\mathbf{A}$ специального вида

$$\mathbf{A} = [(1-a)\,\mathbf{I} - a\mathbf{ee}']^{-1}, \quad -\frac{1}{n-1} < a < 1, \tag{4.2}$$

где $\mathbf{e} = (1, 1, \ldots, 1)$ (здесь $a\mathbf{ee}'$ есть матрица размера $n \times n$, все элементы которой равны $a$).

Матрица $a\mathbf{ee}'$ имеет собственное число $na$ единичной кратности и собственное число нуль кратности $n-1$.

Следовательно, собственные значения матрицы $(1-a)\,\mathbf{I} + a\mathbf{ee}'$ есть $1 + (n-1)\,a, 1-a, 1-a, \ldots, 1-a$.

Таким образом, условие $-\frac{1}{n-1} < a < 1$ эквивалентно положительной определенности матрицы $\mathbf{A}$. Кроме того, матрицы $\mathbf{A}$ вида (4.2) содержатся в $\mathcal{A}$ и в силу теоремы 3.1 есть крайние точки $\mathcal{A}$.

Для случая $n = 2$ матрицы, задаваемые (4.1), имеют вид

$$\mathbf{A} = \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}^{-1}, \quad |a| < 1. \tag{4.3}$$

В этом случае, если в (3.1) мы положим $k_1 = k_2 = k$, то минимизируя $\operatorname{tr} \mathbf{A\Pi}$, где

$$\mathbf{\Pi} = k^{-2} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

получим

$$\Pr(|Y_1| \geqslant k\sigma_1 \text{ или } |Y_2| \geqslant k\sigma_2) \leqslant \frac{1 + \sqrt{1-\rho^2}}{k^2}. \tag{4.4}$$

Здесь $\sigma_1^2$ и $\sigma_2^2$ обозначают соответственно дисперсии $Y_1$ и $Y_2$. Результат (4.4) получен Берге [1937] и является точным всякий раз, когда правая часть не превосходит единицы.

Следующая теорема, полученная Олкином и Праттом [1958], обобщает это неравенство.

Теорема 4.1. *Пусть* $\mathbf{Y}$ — *случайный вектор такой, что* $E\mathbf{Y} = 0$ $E\mathbf{YY}' = \mathbf{\Sigma}$ *и пусть* $X_i = Y_i/k_i\sigma_i$, *где* $\sigma_i^2 = \sigma_{ii}$. *Тогда*

$$\Pr(\max_{1\leqslant i\leqslant n} |Y_i|/k_i\sigma_i \geqslant 1) = \Pr(\max_{1\leqslant i\leqslant n} |X_i| \geqslant 1) \leqslant \frac{n-1}{n} t - \frac{n-2}{n^2} u + \frac{2}{n^2}\sqrt{u(nt-u)(n-1)} = n^{-2}[\sqrt{u} + \sqrt{(nt-u)(n-1)}]^2,$$

*где* $t = \operatorname{tr}\mathbf{\Pi}$, $u = (\mathbf{e}, \mathbf{\Pi e})$ *и* $E\mathbf{XX}' = \mathbf{\Pi}$.

Замечание 4.1. Для $n = 2$ мы получаем неравенство

$$\Pr(|Y_1| \geqslant k_1\sigma_1 \text{ или } |Y_2| \geqslant k_2\sigma_2) \leqslant \leqslant \frac{1}{2k_1^2k_2^2}\left[k_1^2 + k_2^2 + \sqrt{(k_1^2+k_2^2) - 4\rho^2k_1^2k_2^2}\right],$$

которое принадлежит Лалу [1955] и обобщает результат Берге.

Замечание 4.2. Условия, при которых граница в упомянутой теореме является точной, даны у Олкина и Пратта [1958].

Доказательство. Записывая выражение

$$\operatorname{tr}[(1-a)\mathbf{I} + a\mathbf{ee}']^{-1}\mathbf{\Pi} = \frac{\operatorname{tr}(\mathbf{I} - \alpha\mathbf{ee}')\mathbf{\Pi}}{1-a} = \frac{t-\alpha u}{1-a}, \quad \alpha = \frac{a}{1+(n-1)a},$$

и вычисляя его производную по $a$, найдем, что производная обращается в нуль в точках

$$a = \frac{t \pm \sqrt{u(nt-u)/(n-1)}}{u-(n-1)t}. \tag{4.5}$$

Из условия $-(n-1)^{-1} < a < 1$ следует, что только знак (—) обеспечивает допустимое значение $a$. Это значение дает минимум, так как $(t-\alpha u)/(1-a) \to \infty$ при $a \to 1$ или $a \to -(n-1)^{-1}$.

Подстановка (4.5) в выражение $(1-a)^{-1}(t-\alpha u)$ завершает доказательство теоремы.

## § 5. Чебышевские неравенства, в которых $\Omega(\mathbf{x})$ — симметрическая характеристическая функция

Вернемся теперь к ситуации, когда $\Omega_\Gamma(\mathbf{x})$ — характеристическая функция некоторого множества общего вида $\Gamma$. Как и в предыдущих разделах $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ обозначает случайный вектор такой, что $E\mathbf{X} = 0$ и $E\mathbf{XX}' = \mathbf{\Pi}$.

Будем предполагать на протяжении этого параграфа, что $\Omega(\mathbf{x})$ таково, что из соотношения

$$(\mathbf{A}\mathbf{x}, \mathbf{x}) + (\mathbf{b}, \mathbf{x}) + c \geqslant \Omega_{\Gamma}(\mathbf{x}), \qquad c < 1, \tag{5.1}$$

вытекает положительная определенность матрицы $\mathbf{A}$. Например, это будет иметь место, когда множество $\Gamma$ является дополнением ограниченного множества $\Gamma^c$.

Когда $\mathbf{A}$ положительно определена, (5.1) может быть записана в виде

$$(\mathbf{A}(\mathbf{x}-\mathbf{a}), \mathbf{x}-\mathbf{a}) + \delta \geqslant \Omega_{\Gamma}(\mathbf{x}), \tag{5.2}$$

где $\mathbf{a} = -2^{-1}\mathbf{A}^{-1}\mathbf{b}$ и $\delta = c - (\mathbf{A}\mathbf{a}, \mathbf{a})$. Повторяя, рассуждение, проведенное на стр. 499, можно показать, что без потери общности мы можем всегда выбрать $\delta = 0$.

Замечание 5.1. В этом случае справедливо неравенство

$$\Pr(X \in \Gamma) \leqslant \min\{\operatorname{tr} A(\boldsymbol{\Pi} + aa')\},$$

где минимум распространяется по всем матрицам $\mathbf{A}$ и векторам $\mathbf{a}$, удовлетворяющим (5.2) при $\delta = 0$.

Далее мы примем, что функция $\Omega_{\Gamma}(\mathbf{x})$ и матрица $\boldsymbol{\Pi}$ инвариантны по отношению к некоторой группе преобразований $S \in \mathcal{C}$, действующей на $\mathbf{x}$, т. е. постулируем, что

$$\begin{gathered} \Omega_{\Gamma}(\mathbf{S}\mathbf{x}) = \Omega_{\Gamma}(\mathbf{x}), \\ \mathbf{S}\boldsymbol{\Pi}\mathbf{S}' = \boldsymbol{\Pi} \ (\mathbf{S} \in \mathcal{C}). \end{gathered} \tag{5.3}$$

Тот интуитивно ясный факт, что симметрия задачи должна отражаться в симметрии класса минимизирующих многочленов, формализуется в следующей теореме.

Теорема 5.1. *Если выполняются условия* (5.3), *то мы можем ограничить внимание при минимизации*

$$\operatorname{tr} \mathbf{A}\{\boldsymbol{\Pi} + \mathbf{a}\mathbf{a}'\} \tag{5.4}$$

*векторами* $\mathbf{a}$, *инвариантными по отношению к* $\mathcal{C}$, *т.е. такими векторами, что* $\mathbf{S}\mathbf{a} = \mathbf{a}$ *для всех* $\mathbf{S} \in \mathcal{C}$. *Кроме того, если группа* $\mathcal{C}$ *компактна, абелева или, более общим образом, разрешима, то мы можем ограничить внимание матрицами* $\mathbf{A}$, *удовлетворяющими* (5.2) *и инвариантными относительно* $\mathcal{C}$, *т.е. такими, что* $\mathbf{S}'\mathbf{A}\mathbf{S} = \mathbf{A}$.

Доказательство. Предположим, что $\sup \Pr(\mathbf{X} \in \Gamma) < 1$, и пусть вектор $\mathbf{a}_0$ и матрица $\mathbf{A}_0$ минимизируют (5.4). Тогда (5.3) дает

$$(\mathbf{S}(\mathbf{x} - \mathbf{S}^{-1}\mathbf{a}_0), \mathbf{A}_0\mathbf{S}(\mathbf{x} - \mathbf{S}^{-1}\mathbf{a}_0)) \geqslant \Omega(\mathbf{x}).$$

Рассмотрим

$$f(\mathbf{x}) = \frac{1}{2}\left[((\mathbf{x}-\mathbf{a}_0), \mathbf{A}_0(\mathbf{x}-\mathbf{a})) + ((\mathbf{x}-\mathbf{S}^{-1}\mathbf{a}_0), \mathbf{S}'\mathbf{A}_0\mathbf{S}(\mathbf{x}-\mathbf{S}^{-1}\mathbf{a}_0))\right].$$

Обращаясь ко второму соотношению в (5.3), получим

$$Ef(\mathbf{X}) = \operatorname{tr}\mathbf{A}_0[\mathbf{\Pi} + \mathbf{a}_0\mathbf{a}_0']. \tag{5.5}$$

Если $\mathbf{S}\mathbf{a}_0 \neq \mathbf{a}_0$, то функция $f(\mathbf{x})$ не обращается в нуль и может быть записана в виде (5.2) при $\delta > 0$. Однако, когда $\sup \Pr(\mathbf{X} \in \Gamma) < 1$, минимизирующая квадратичная форма должна иметь $\delta = 0$. Следовательно, $\mathbf{S}\mathbf{a}_0 = \mathbf{a}_0$.

Если группа $\mathcal{G}$ компактна, то положим

$$\widetilde{\mathbf{A}} = \int \mathbf{S}'\mathbf{A}_0\mathbf{S}\, d\mu(\mathbf{S}), \tag{5.6}$$

где $d\mu$ обозначает меру Хаара на $\mathcal{G}$, нормированную так, что мера $\mathcal{G}$ равна единице. Тогда

$$((\mathbf{x}-\mathbf{a}_0), \widetilde{\mathbf{A}}(\mathbf{x}-\mathbf{a}_0)) \geqslant \Omega(\mathbf{x}), \tag{5.7}$$

$$E((\mathbf{X}-\mathbf{a}_0), \widetilde{\mathbf{A}}(\mathbf{X}-\mathbf{a}_0)) = \int E((\mathbf{X}-\mathbf{a}_0), \mathbf{S}'\mathbf{A}_0\mathbf{S}(\mathbf{X}-\mathbf{a}_0))\, d\mu(\mathbf{S}) =$$
$$= \operatorname{tr}\mathbf{A}_0(\mathbf{\Pi} + \mathbf{a}_0\mathbf{a}_0').$$

Матрица $\widetilde{\mathbf{A}}$ удовлетворяет соотношению $\mathbf{S}'\widetilde{\mathbf{A}}\mathbf{S} = \mathbf{A}'$ для всех $\mathbf{S} \in \mathcal{G}$ и достигает минимальной границы. Таким образом, доказательство теоремы завершено для случая, когда $\mathcal{G}$ — компакт.

В случае, когда группа $\mathcal{G}$ абелева, мы действуем следующим образом. Пусть $\mathcal{A}_0$ — класс положительно определенных матриц, удовлетворяющих неравенству $((\mathbf{x}-\mathbf{a}_0), \mathbf{A}(\mathbf{x}-\mathbf{a}_0)) \geqslant \Omega_\Gamma(\mathbf{x})$, которые доставляют минимум выражению $\operatorname{tr}\mathbf{A}\{\mathbf{\Pi} + \mathbf{a}_0\mathbf{a}_0'\}$. Отметим, что $\mathcal{A}_0$ выпукло, и на основе (3.5) мы можем заключить, что $\mathcal{A}_0$ ограничено. Кроме того, так как неравенство

$$((\mathbf{x}-\mathbf{a}_0), \mathbf{A}(\mathbf{x}-\mathbf{a}_0)) \geqslant \Omega_\Gamma(\mathbf{x}) \tag{5.8}$$

влечет $\mathbf{A} > 0$, то $\mathcal{A}_0$ замкнуто.

Далее для фиксированного $\mathbf{A}_0 \in \mathcal{A}$ и $\mathbf{S}_1 \in \mathcal{G}$ положим

$$\mathbf{B}_m = \frac{1}{m}\sum_{k=0}^{m} (\mathbf{S}_1')^k \mathbf{A}_0 (\mathbf{S}_1)^k.$$

Ясно, что $\mathbf{B}_m$ содержится в $\mathcal{A}_0$.

Так как $\mathcal{A}_0$ компактно, мы можем выбрать подпоследовательность $\mathbf{B}_{m_l}$, которая сходится к некоторому $\mathbf{A}_1$ в $\mathcal{A}_0$. Более того, так как последовательность

$$\mathbf{S}'\mathbf{B}_{m_l}\mathbf{S} - \mathbf{B}_{m_l} = \frac{1}{m_l}\{(\mathbf{S}')^{m_l+1}\mathbf{A}_0(\mathbf{S})^{m_l+1} - \mathbf{A}_0\}$$

сходится к нулевой матрице, то

$$\mathbf{S}'\mathbf{A}_1\mathbf{S} = \mathbf{A}_1.$$

Пусть $\mathcal{A}_1$ — подкласс $\mathcal{A}_0$, определенный как

$$\mathcal{A}_1 = \{\mathbf{A} \mid \mathbf{A} \in \mathcal{A}_0, \mathbf{S}_1'\mathbf{A}\mathbf{S}_1 = \mathbf{A}\}.$$

Этот подкласс является выпуклым и компактным и, согласно предыдущим рассуждениям, непустым.

Теперь для $\mathbf{S}_2 \in \mathcal{G}$ положим $A_2 = \{\mathbf{A} \mid \mathbf{A} \in \mathcal{A}_1, \mathbf{S}_2'\mathbf{A}\mathbf{S}_2 = \mathbf{A}\}$; $\mathcal{A}_2$ также является выпуклым и компактным.

Мы хотим показать, что $\mathcal{A}_2$ не пусто. Для фиксированного $\mathbf{A}_1$ выберем подпоследовательность из

$$\frac{1}{m}\sum_{k=0}^{m} (\mathbf{S}_2')^k \mathbf{A}_1 \mathbf{S}_2^k,$$

которая сходится, скажем, к $\mathbf{A}_2$. Из того, что $\mathcal{G}$ — абелева, заключаем, что $(\mathbf{S}_2')^k\mathbf{A}_1\mathbf{S}_2^k \in \mathcal{A}_1$ и, следовательно, $\mathbf{A}_2 \in \mathcal{A}_2$. Итак, $\mathcal{A}_2$ не пусто.

Действуя подобным образом, мы получаем, что для всякого конечного множества элементов $\mathbf{S}_i \in \mathcal{G}$, $i = 1, 2, \ldots, n$, множество

$$\mathcal{A}(\mathbf{S}_1, \ldots, \mathbf{S}_n) = \{\mathbf{A} \mid \mathbf{A} \in \mathcal{A}_0, \mathbf{S}_i'\mathbf{A}\mathbf{S}_i = \mathbf{A},\ i = 1, \ldots, n\}$$

не пусто. Так как эти множества являются также компактными, то из свойства конечных пересечений и компактности $\mathcal{A}_0$ мы заключаем, что существует элемент $\mathbf{A} \in \mathcal{A}_0$, удовлетворяющий соотношению $\mathbf{S}'\mathbf{A}\mathbf{S} = \mathbf{A}$ для всех $\mathbf{S} \in \mathcal{G}$. Теорема, таким образом, установлена в случае, когда $\mathcal{G}$ — абелева.

Используя полученный для абелевых и компактных групп результат, распространение на разрешимые группы можно осуществить стандартными рассуждениями (см., например, работу Карлина [1953 b]). Мы опускаем детали.

Параграфы 6 и 7 дают примеры использования доказанной теоремы. Можно проиллюстрировать результат, взяв в качестве $\mathcal{G}$ циклическую группу, порожденную матрицей

$$\mathbf{S} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

При таком выборе $\mathcal{G}$ матрица $\mathbf{\Pi}$ инвариантна, т. е. $\mathbf{S}'\mathbf{\Pi}\mathbf{S} = \mathbf{\Pi}$ для всех $\mathbf{S} \in \mathcal{G}$ тогда и только тогда, когда $\mathbf{\Pi}$ имеет вид

$$\mathbf{\Pi} = \begin{pmatrix} \alpha & \beta & \gamma & \beta \\ \beta & \alpha & \beta & \gamma \\ \gamma & \beta & \alpha & \beta \\ \beta & \gamma & \beta & \alpha \end{pmatrix}. \tag{5.9}$$

Теорема 5.1 утверждает, что для минимизации (5.4) достаточно ограничить внимание векторами **a**, все компоненты котороых одинаковы, и мы можем выбирать **A** того же вида, что и **П**. Конечно, это верно только при условии, что области $\Gamma$ и $\Gamma^c$ инвариантны. Простая область такого типа определяется неравенствами

$$x_1 + x_2 + x_3 < l, \quad x_2 + x_3 + x_4 < l,$$
$$x_3 + x_4 + x_1 < l, \quad x_4 + x_1 + x_2 < l.$$

## § 6. Чебышевские неравенства для прямоугольника

Материал этого параграфа опирается на работу Исии [1959 b].

Дискуссия § 3 была посвящена задаче характеризации и вычисления точных верхних границ для вероятности

$$\Pr\left(\max_{1 \leqslant i \leqslant n} |X_i| \geqslant 1\right), \tag{6.1}$$

где $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ имеет заданный первый и второй моменты. Достигнутые выводы были существенно связаны с предположением о том, что $E\mathbf{X} = 0$ и симметрией области $\Gamma$. В данном параграфе мы проанализируем особый случай этой задачи, когда эти ограничения сняты.

А именно, мы хотим определить точную верхнюю границу для вероятности

$$\Pr(\mathbf{X} \in \Gamma) \tag{6.2}$$

по отношению ко всем случайным векторам $\mathbf{X} = (X_1, X_2)$ с вектором средних $\mu$ $(\mu_1, \mu_2)$ и ковариационной матрицей $\Sigma$. Здесь множество $\Gamma$ является дополнением открытого прямоугольника, стороны которого не обязательно равной длины. Примем отношение длин сторон, одной, параллельной оси $x_1$, и другой, параллельной оси $x_2$, равным $\sigma_1/\sigma_2$. Мы рассматриваем только случай, когда $\mu$ лежит на диагонали и является внутренней точкой множества $\Gamma^c$.

Сдвигом начала координат, изменением масштаба и, возможно, вращением координатных осей задача сводится к оценке максимума вероятности попадания во внешность прямоугольника $\Gamma_0^c$ с вершинами $(-\alpha, -\alpha)$, $(-\alpha, \beta)$, $(\beta, \beta)$ и $(\beta, -\alpha)$, $(\alpha > 0, \beta > 0)$ для случайных векторов, удовлетворяющих условиям $E\mathbf{X} = 0$ и

$$E\mathbf{X}\mathbf{X}' = \mathbf{\Pi} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

Методы § 1 показывают, что

$$\Pr(\mathbf{X} \in \Gamma_0) \leqslant \min \operatorname{tr} \mathbf{A}[\mathbf{\Pi} + \mathbf{a}\mathbf{a}']. \tag{6.3}$$

Минимум оценивается по отношению ко всем **A** и **a**, удовлетворяющим условию

$$((\mathbf{x} - \mathbf{a}), \mathbf{A}(\mathbf{x} - \mathbf{a})) \geqslant \Omega_{\Gamma_0}(\mathbf{x}), \tag{6.4}$$

где матрица $\mathbf{A}$ положительно определена, а $\Omega_{\Gamma_0}(\mathbf{x})$ обозначает характеристическую функцию $\Gamma_0$. Граница в (6.3) достигается при некотором выборе $\mathbf{x}$ всякий раз, когда правая часть не превосходит единицы.

Пусть $\mathcal{C}$ — группа из двух элементов, порожденная

$$\mathbf{S} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix};$$

применяя теорему 5.1, мы получаем, что достаточно ограничить внимание векторами $\mathbf{a}$, для которых $a_1 = a_2 = m$ и матрицами $\mathbf{A}$ вида [1])

$$\mathbf{A} = \begin{vmatrix} d & -dt \\ -dt & d \end{vmatrix}, \quad d > 0, \quad |t| < 1. \tag{6.5}$$

Мы поэтому хотим минимизировать $\operatorname{tr} \mathbf{A}[\mathbf{\Pi} + \mathbf{aa}']$, где $\mathbf{A}$ имеет вид (6.5), а $\mathbf{a} = (m, m)$. Простые преобразования показывают, что

$$\operatorname{tr} \mathbf{A}[\mathbf{\Pi} + \mathbf{aa}'] = \lambda_1 u_1 + \lambda_2 u_2 + \lambda_1 \xi^2, \tag{6.6}$$

где

$$\lambda_1 = d(1 - t), \quad \lambda_2 = d(1 + t), \quad u_1 = 1 + \rho, \ u_2 = 1 - \rho, \quad \xi = \sqrt{2}m.$$

Условие $((\mathbf{x} - \mathbf{a}), \mathbf{A}(\mathbf{x} - \mathbf{a})) \geqslant \Omega_{\Gamma_0}(\mathbf{x})$ сводится к

$$\frac{2\lambda_1\lambda_2}{\lambda_1 + \lambda_2}(\alpha + m)^2 \geqslant 1, \quad \frac{2\lambda_1\lambda_2}{\lambda_1 + \lambda_2}(\beta - m)^2 \geqslant 1. \tag{6.7}$$

Мы предположим теперь для определенности, что $\alpha \leqslant \beta$ и минимизируем (6.6) при ограничении (6.7). Окончательное вычисление приводит к следующему результату.

Теорема 6.1. *Пусть* $\mathbf{X} = (X_1, X_2)$ — *случайный вектор такой, что* $E\mathbf{X} = 0$ *и* $E\mathbf{XX}' = \mathbf{\Pi} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. *Пусть* $\Gamma_0$ — *множество, дополнение которого есть открытый прямоугольник с вершинами* $(-\alpha, -\alpha)$, $(-\alpha, \beta)$, $(\beta, \beta)$ *и* $(\beta, -\alpha)$, *где* $0 < \alpha \leqslant \beta$.

(1) *Если* $\beta - \alpha \geqslant \sqrt{2}\lambda$ *и* $2\alpha^2 > 1 - \rho$, где

$$\lambda = \frac{\sqrt{2}\alpha(1 + \rho) + \sqrt{2(1 - \rho^2)(\alpha^2 + \rho)}}{2\alpha^2 - (1 - \rho)},$$

то

$$\Pr(\mathbf{X} \in \Gamma_0) \leqslant \frac{\lambda^2}{\lambda^2 + (1 + \rho)}.$$

(2) *Если условие* (1) *не удовлетворяется*, $\alpha, \beta > 1$ *и* $2(\alpha\beta - 1)^2 \geqslant 2(1 - \rho^2) + (1 - \rho)(\beta - \alpha)^2$, *то*

$$\Pr(\mathbf{X} \in \Gamma_0) \leqslant \frac{(\beta - \alpha)^2 + 4 + \sqrt{16(1 - \rho^2) + 8(1 - \rho)(\beta - \alpha)^2}}{(\alpha + \beta)^2}.$$

[1]) Сведение к этому классу матриц может быть получено непосредственно из соображений симметрии без обращения к теореме 5.1.

(3) *Если не удовлетворяются оба условия* (1) *и* (2), *то*

$$\Pr(\mathbf{X} \in \Gamma_0) \leqslant 1.$$

*В каждом случае граница является точной.*

Замечание 6.1. (1) Подстановкой $\rho = 1$ неравенство сводится к результату Селберга (см. § 3 гл. XII).

(2) Если $\alpha = \beta$, мы получаем результат Берге (см. (4.4)).

## § 7. Неравенства для $\Pr\left(\max\limits_{1 \leqslant i \leqslant n} \mathbf{X}_i \geqslant 1\right)$

В §§ 3 и 4 мы обсуждали задачу максимизации $\Pr\left(\max\limits_{1 \leqslant i \leqslant n} |X_i| \geqslant 1\right)$. В этом параграфе мы получим некоторые вероятностные границы (см. Маршалл и Олкин [1960 b]) для односторонней области $\Gamma = \{x \mid \max\limits_{1 \leqslant i \leqslant n} x_i \geqslant 1\}$. Рассмотрим

$$\Pr(\mathbf{X} \in \Gamma) = \Pr\left(\max_{1 \leqslant i \leqslant n} X_i \geqslant 1\right), \tag{7.1}$$

где $\mathbf{X} = (X_1, \ldots, X_n)$ — случайный вектор, удовлетворяющий условиям $E\mathbf{X} = 0$, $E\mathbf{X}\mathbf{X}' = \mathbf{\Pi}$, где $\mathbf{\Pi}$ имеет вид

$$\mathbf{\Pi} = \sigma^2 [(1 - \rho)\,\mathbf{I} + \rho \mathbf{e}\mathbf{e}'], \qquad -(n-1)^{-1} < \rho < 1 \tag{7.2}$$

(ср. (4.2))

Наша задача снова сводится к задаче минимизации выражения

$$\operatorname{tr} \mathbf{A}[\mathbf{\Pi} + \mathbf{a}\mathbf{a}']. \tag{7.3}$$

где $\mathbf{A}$ и $\mathbf{a}$ удовлетворяют ограничению

$$((\mathbf{x} - \mathbf{a}), \mathbf{A}(\mathbf{x} - \mathbf{a})) \geqslant \Omega_\Gamma(\mathbf{x}). \tag{7.4}$$

Как и в § 5, мы можем предположить, что $\mathbf{A} > 0$.

Применяя теорему 5.1 ($\mathcal{G}$, которое следует взять, есть группа матриц перестановок размера $n \times n$), получаем, что достаточно ограничить внимание векторами $\mathbf{a}$ с равными компонентами и матрицами $\mathbf{A}$ с одинаковыми элементами на диагонали и одинаковыми элементами вне диагонали, т. е. матрицами $\mathbf{A}$ вида (7.2).

Теперь, если $\mathbf{a} = m\mathbf{e}$ $(m < 1, \mathbf{e} = (1, 1, \ldots, 1))$, то $\operatorname{tr} \mathbf{A}(\mathbf{\Pi} + \mathbf{a}\mathbf{a}') = (1 - m)^{-2} \operatorname{tr} \mathbf{B}^{-1}(\mathbf{\Pi} + m^2 \mathbf{e}\mathbf{e}')$, где $\mathbf{B}^{-1} = (1 - m)^2 \mathbf{A}$.

Подстановка $\mathbf{z} = (1 - m)^{-1}(\mathbf{x} - \mathbf{a})$ приводит (7.4) к виду

$$(\mathbf{z}, \mathbf{B}^{-1}\mathbf{z}) \geqslant \Omega_\Gamma(m\mathbf{e} + (1 - m)\,\mathbf{z}) = \Omega_\Gamma(\mathbf{z}).$$

Обращаясь к теореме 3.1, мы получаем возможность рассматривать только те матрицы $\mathbf{B}$, на главной диагонали которых все элементы равны единице. Если мы положим, что $\mathbf{B} = (1 - b)\,\mathbf{I} + b\mathbf{e}\mathbf{e}'$, а $\mathbf{U}$ — произвольная ортогональная матрица с первым столбцом, равным $\mathbf{e}/\sqrt{n}$, то (7.3) принимает вид

$$H(m, b) = \frac{\operatorname{tr}\{(\mathbf{UBU}')^{-1}[\mathbf{U\Pi U}' + m^2\mathbf{U}\mathbf{ee}'\mathbf{U}']\}}{(1-m)^2} =$$

$$= \frac{n[m^2 + \sigma^2 + b(\sigma^2 t - m^2)]}{(1-m^2)(1-b)[1+(n-1)b]},$$

где $t = (n-1)(1-\rho) - 1$.

Мы опускаем утомительные выкладки, участвующие в минимизации $H(m, b)$ и просто записываем результат.

Теорема 7.1. *Пусть* $\mathbf{X} = (X_1, \ldots, X_n)$ — *случайный вектор с* $E\mathbf{X} = 0$ *и* $E\mathbf{XX}' = \mathbf{\Pi}$, *где* $\mathbf{\Pi}$ *имеет вид* (7.2). *Если* $1 - \sigma^2 t > 0$ *и* $n \geqslant \sigma^2(n-1)(1+t)$, *то следующее неравенство является точным*:

$$P = \Pr(\max_{1 \leqslant i \leqslant n} X_i \geqslant 1) \leqslant$$

$$\leqslant \frac{n\sigma^2\{\sqrt{(1+(n-1)\rho)[1+\sigma^2-\sigma^2(n-1)(1-\rho)]} + (n-1)\sqrt{1-\rho}\}^2}{\{n+\sigma^2[1+(n-1)\rho]\}^2}.$$

*Если упомянутые условия не выполняются, то* $P \leqslant 1$.

Замечание 7.1. (1) Если $\rho = 1$ или $n = 1$, то получаем одномерное одностороннее неравенство $\Pr(\mathbf{X} \geqslant 1) \leqslant \sigma^2/(\sigma^2+1)$.

(2) Маршаллом и Олкином [1960 b] даны конкретные иллюстрации, представляющие случаи равенства в упомянутой теореме.

## § 8. Чебышевские неравенства для случая, когда $\Omega$ есть характеристическая функция конечного объединения выпуклых областей

Пусть $\Omega_+(\mathbf{x})$ есть характеристическая функция выпуклого множества $S_+$, замыкание которого не содержит начала координат. Определим $S_-$ как симметрический образ $S_+$ и множество $S = S_- \bigcup S_+$.

Как и прежде, мы задаем первые и вторые моменты, т. е.

$$E(X_i) = 0, \quad E(X_i X_j) = \pi_{ij}, \qquad i, j = 1, \ldots, n, \tag{8.1}$$

и хотим определить

$$\sup_{\sigma \in V(\mathbf{c})} E(\Omega(x_1, \ldots, x_n)) = I_\Omega, \tag{8.2}$$

где $\Omega$ — характеристическая функция симметричного множества $S$, а $V(\mathbf{c})$ состоит из вероятностных мер, удовлетворяющих (8.1). Используя теорему 2.1 гл. XII, получаем

$$I_\Omega \leqslant \min \operatorname{tr} \mathbf{A\Pi}, \tag{8.3}$$

где минимум берется по множеству всех многочленов, удовлетворяющих условию

$$(\mathbf{Ax}, \mathbf{x}) \geqslant \Omega(\mathbf{x}), \quad \mathbf{x} = (x_1, \ldots, x_n). \tag{8.4}$$

Очевидно, что достаточно ограничить внимание теми матрицами **A**, для которых равенство выполняется при некотором **x** на границе $S$.

Покажем теперь, что в данном случае множество всех матриц **A**, которые уместно рассматривать при оценивании (8.3), может быть подвергнуто дальнейшему существенному сокращению. По существу, достаточно рассматривать только матрицы **A** единичного ранга, удовлетворяющие (8.4).

Теорема 8.1. *При данных выше условиях имеем*

$$I_\Omega = \min(\mathbf{a}, \mathbf{\Pi a}) \tag{8.5}$$

*при условии, что* $I_\Omega < 1$, *где минимум распространяется по множеству* $\mathcal{A}^*$ *всех векторов* **a**, *для которых*

$$(\mathbf{a}, \mathbf{x})^2 \geqslant 1 \tag{8.6}$$

*для всех* $\mathbf{x} \in S$.

Доказательство. Рассмотрим произвольную матрицу **A**, удовлетворяющую (8.4). Рассмотрим множество $U$ всех **x**, для которых $(\mathbf{Ax}, \mathbf{x}) \leqslant 1$. Ясно, что $U$ и $S_+$ — выпуклые множества, не содержащие внутренних точек в пересечении. Заметим также, что начало координат есть внутренняя точка $U$. В соответствии с этим существует разделяющая гиперплоскость, т. е. существует вектор **a** такой, что

$$\begin{aligned} (\mathbf{a}, \mathbf{x}) &\leqslant 1 \quad \mathbf{x} \in U, \\ (\mathbf{a}, \mathbf{x}) &\geqslant 1, \quad \mathbf{x} \in S^+. \end{aligned} \tag{8.7}$$

Так как $\mathbf{x} \in U$ и, следовательно, $-\mathbf{x} \in U$, то отсюда следует, что $(\mathbf{a}, \mathbf{x})^2 \leqslant 1$ для всех **x**, удовлетворяющих условию $(\mathbf{Ax}, \mathbf{x}) = 1$ и, следовательно,

$$(\mathbf{Ax}, \mathbf{x}) \geqslant (\mathbf{a}, \mathbf{x})^2 \tag{8.8}$$

для всех **x**.

Более того, мы заключаем на основе (8.7), что

$$(\mathbf{a}, \mathbf{x})^2 \geqslant 1 \tag{8.9}$$

для $\mathbf{x} \in S$.

Соотношения (8.9) показывают, что $(\mathbf{a}, \mathbf{x})^2$ есть допустимый многочлен, удовлетворяющий условию (8.4).

В силу (8.8) имеем

$$\operatorname{tr} \mathbf{A\Pi} \geqslant \operatorname{tr}\{\mathbf{aa}'\mathbf{\Pi}\} = (\mathbf{a}, \mathbf{\Pi a}). \tag{8.10}$$

Это выражает тот факт, что для всякого **A**, удовлетворяющего (8.4), мы можем найти допустимую матрицу единичного ранга, которая улучшает оценку $I_\Omega$. Следовательно, (8.5) имеет место, что и требовалось доказать.

Односторонняя версия теоремы 8.1 может быть доказана подобным методом.

Теорема 8.2. *Пусть* $\Omega_+(\mathbf{x})$ — *характеристическая функция множества* $S_+$; *тогда*

$$I_+ = \max_{\sigma \in V(\mathbf{c})} \int \Omega_+(\mathbf{x})\, d\sigma(\mathbf{x}) = \min_{\mathbf{a} \in \mathcal{A}_+^*} \frac{(\mathbf{a}, \Pi\mathbf{a})}{1 + (\mathbf{a}, \Pi\mathbf{a})}, \tag{8.11}$$

*где*

$$\mathcal{A}_+^* = \{\mathbf{a} \mid (\mathbf{a}, \mathbf{x}) \geqslant 1 \text{ для всех } \mathbf{x} \in S_+\}. \tag{8.12}$$

Доказательство. Из теоремы 2.1 гл. XII следует, что $I_+ = \min(\gamma + \operatorname{tr} \mathbf{A}\Pi)$, где минимум берется по отношению к множеству всех многочленов, удовлетворяющих соотношению

$$\gamma + (\mathbf{b}, \mathbf{x}) + (\mathbf{A}\mathbf{x}, \mathbf{x}) \geqslant \Omega_+(\mathbf{x}). \tag{8.13}$$

Отсюда непосредственно заключаем, что матрица $\mathbf{A}$ положительно полуопределена и $\gamma \geqslant 0$.

Многочлен в левой части описывает эллипсоид, центр которого не обязательно лежит в начале координат. Пусть $U$ обозначает область, где

$$(\tilde{\mathbf{A}}(\mathbf{x} - \boldsymbol{\xi}^0), \mathbf{x} - \boldsymbol{\xi}^0) = \gamma + (\mathbf{b}, \mathbf{x}) + (\mathbf{A}\mathbf{x}, \mathbf{x}) \leqslant 1 \tag{8.14}$$

$$(\tilde{\mathbf{A}} \geqslant 0 \text{ и } \boldsymbol{\xi}^0 \notin \bar{S}_+)$$

и $\boldsymbol{\xi}^0$ представляет центр эллипсоида. Ясно, что $U$ и $S_+$ не перекрываются и, следовательно, могут быть разделены опорной гиперплоскостью. Перенося $U$ и $S_+$ так, чтобы центр $U$ совпадал с началом координат, мы получим, как и в теореме 8.1, что $(\tilde{\mathbf{A}}(\mathbf{x} - \boldsymbol{\xi}^0), \mathbf{x} - \boldsymbol{\xi}^0) \geqslant (\delta + (\mathbf{a}, \mathbf{x}))^2$ для всех $\mathbf{x}$, и $(\delta + (\mathbf{a}, \mathbf{x})) \geqslant 1$ для $\mathbf{x} \in S_+$. (В самом деле, $\delta = (\mathbf{a}, -\boldsymbol{\xi}^0)$).

Следовательно,

$$I_+ = \min(\delta^2 + (\mathbf{a}, \Pi\mathbf{a})), \tag{8.15}$$

где минимум берется по множеству всех многочленов вида

$$(\delta + (\mathbf{a}, \mathbf{x})) \geqslant 1 \tag{8.16}$$

для $\mathbf{x} \notin S_+$. Перепишем (8.16) в виде

$$(\mathbf{a}, \mathbf{x}) \geqslant 1 - \delta \text{ или } (\mathbf{b}, \mathbf{x}) \geqslant 1 \tag{8.17}$$

для $\mathbf{x} \in S_+$, где $\mathbf{b} = \mathbf{a}/(1 - \delta)$ $(\delta < 1)$ и (8.15) превращается в

$$I_+ = \min(\delta^2 + (1 - \delta)^2 (\mathbf{b}, \Pi\mathbf{b})). \tag{8.18}$$

Минимизация по $\delta$ приводит к следующему:

$$I_+ = \min_{\mathbf{b}} \frac{(\mathbf{b}, \Pi\mathbf{b})}{1 + (\mathbf{b}, \Pi\mathbf{b})},$$

где $\mathbf{b}$ — произвольный вектор, удовлетворяющий неравенству $(\mathbf{b}, \mathbf{x}) \geqslant 1$ для всех $\mathbf{x} \in S_+$. Более того, легко видеть, что минимум действительно достигается.

Теоремы 8.1 и 8.2 принадлежат Олкину и Маршаллу [1960 а]. Данный здесь метод отличен от их метода. Теорема 8.3, приводимая ниже, по-видимому, нова.

Рассмотрим область $W_+ = \bigcap_{i=1}^{k} S_i$ где $S_i$ все выпуклы, и ни одно из этих множеств не содержит начала координат в своем замыкании. Пусть $W_-$ — симметричный образ $W_+$ и $W = W_+ \bigcup W_-$. Обозначим через $\Omega_W(t)$ характеристическую функцию $W$ и положим

$$I_W = \sup_{\sigma \in V(c)} \int \Omega_W(\mathbf{x})\, d\sigma(\mathbf{x}).$$

Пусть $\mathbf{A}^0$ — положительно определенная фиксированная матрица, удовлетворяющая неравенству $(\mathbf{A}^0\mathbf{x}, \mathbf{x}) \geqslant \Omega_W(\mathbf{x})$. Рассмотрим область $R = \{\mathbf{x} \mid (\mathbf{A}^0\mathbf{x}, \mathbf{x}) \leqslant 1\}$. Ясно, что $R^0$ (через $R^0$ обозначена внутренность $R$) не пересекается с $S_1, S_2, \ldots, S_k$. Для каждого $S_i$ существует отделяющая плоскость $(\mathbf{a}_i, \mathbf{x}) \geqslant 1$ для $\mathbf{x} \in S_i$, $(\mathbf{a}_i, \mathbf{x})^2 \leqslant 1$ при $\mathbf{x} \in R$. Пусть $\widetilde{S}_i = \{\mathbf{x} \mid (\mathbf{a}_i, \mathbf{x})^2 \geqslant 1\}$ $(i = 1, \ldots, k)$ и $\widetilde{W} = \bigcap_{i=1}^{k} \widetilde{S}_i \supset W$.

Очевидно,

$$R \subset \overline{\bigcap_{i=1}^{k} \widetilde{S}_i^c}.$$

Пусть $\mathbf{b}_1, \ldots, \mathbf{b}_l$ — минимальное множество векторов, линейно не зависящих от $\{\mathbf{a}_i\}_{i=1}^k$, такое, что $\{\mathbf{a}_i\}_{i=1}^k \bigcup \{\mathbf{b}_j\}_{j=1}^l$ порождает $R^n$. Зададим длину векторов $\mathbf{b}_j$ так, что $|\mathbf{b}_i| \leqslant \varepsilon, j = 1, \ldots, l$. Теперь рассмотрим множество точек

$$U^c = \{\mathbf{x} \mid (\mathbf{a}_i, \mathbf{x})^2 < 1,\ i = 1, \ldots, k,\quad (\mathbf{b}_j, \mathbf{x}) < 1,\quad j = 1, \ldots, l\}. \tag{8.19}$$

Если $\varepsilon > 0$ выбрать достаточно малым, то

$$R \subset \overline{U}^c. \tag{8.20}$$

Пусть $\Omega_U(t)$ — характеристическая функция множества $U$, дополнительного к (8.19). Соотношение (8.20) выражает свойство

$$(\mathbf{A}^0\mathbf{x}, \mathbf{x}) \geqslant \Omega_U(\mathbf{x}) \geqslant \Omega_W(\mathbf{x}). \tag{8.21}$$

Теперь

$$I_W \leq I_U = \sup_{\sigma \in V} \int \Omega_U(\mathbf{x})\, d\sigma(\mathbf{x}) \leq \min \operatorname{tr} \Pi\mathbf{A} \leq \operatorname{tr} \Pi\mathbf{A}^0, \tag{8.22}$$

где минимум распространяется по тем $\mathbf{A}$, которые удовлетворяют (8.21). Мы знаем по теореме 2.1, что минимум достигается в крайней точке, а эта последняя характеризуется тем свойством, что эллипсоид $(\mathbf{A}\mathbf{x}, \mathbf{x}) \leqslant 1$ касается поверхности $U^c$ в достаточно большом числе точек, содержащих $n$ линейно независимых векторов.

Таким образом, для всякого $\varepsilon > 0$ существует матрица $\mathbf{A}_\varepsilon$ такая, что

$$(\mathbf{A}_\varepsilon \mathbf{x}, \mathbf{x}) \geqslant \Omega_W(\mathbf{x}), \quad \operatorname{tr}(\mathbf{A}_\varepsilon \mathbf{\Pi}) \leqslant \operatorname{tr}(\mathbf{A}^0 \mathbf{\Pi})$$

и $(\mathbf{A}_\varepsilon \mathbf{x}, \mathbf{x}) \leqslant 1$ касается поверхности $U^c$ вдоль $n$ линейно независимых направлений. Это значит, в частности, что $(\mathbf{A}_\varepsilon \mathbf{x}, \mathbf{x}) = 1$ для каждого $j$ при некотором $\mathbf{x}$ (зависящем от $j$), удовлетворяющем соотношению $(\mathbf{b}_j, \mathbf{x}) = 1$.

Так как $|\mathbf{b}_j| \to 0$ при $\varepsilon \to 0$, то отсюда следует, что мы можем выбрать предельную матрицу $\tilde{\mathbf{A}}$ ранга, не большего чем $k$, для которой

$$(\tilde{\mathbf{A}} \mathbf{x}, \mathbf{x}) \geqslant \Omega_W(\mathbf{x}), \quad \operatorname{tr}(\tilde{\mathbf{A}} \mathbf{\Pi}) \leqslant \operatorname{tr}(\mathbf{A}^0 \mathbf{\Pi}). \tag{8.23}$$

Суммируя, мы имеем следующий результат.

Теорема 8.3. *Пусть* $W_+ = \bigcup_{i=1}^{k} S_i$, *где каждое* $S_i$ *выпукло и* $0 \notin S_i$. *Пусть* $W_-$ — *симметричный образ* $W_+$, *и определим* $W = W_+ \cup W_-$. *Тогда*

$$I_W = \sup_{\sigma \in V(c)} \int \Omega_W(\mathbf{x})\, d\sigma(\mathbf{x}) \leqslant \min \operatorname{tr} \mathbf{\Pi A}, \tag{8.24}$$

*где минимум берется по множеству всех матриц ранга, не большего, чем* $k$, *для которых* $(\mathbf{A}\mathbf{x}, \mathbf{x}) \geqslant \Omega_W(\mathbf{x})$. *Более того, если* $I_W < 1$, *то равенство в* (8.24) *достигается.*

## § 9. Границы для минимальных компонент

Результаты этого параграфа взяты из работы Маршала и Олкина [1960 b].

Пусть $S = S_+ \cup S_-$, где

$$S_+ = \{\mathbf{x} \mid \min_{1 \leqslant j \leqslant n} x_j \geqslant 1\},$$

$$S_- = \{\mathbf{x} \mid \min_{1 \leqslant j \leqslant n} (-x_j) \geqslant 1\}.$$

В качестве следствий из результатов § 8 мы докажем следующие результаты.

Теорема 9.1. *Если* $E\mathbf{X} = 0$ *и* $E\mathbf{X}\mathbf{X}' = \mathbf{\Pi}$, *то*

$$\Pr(\mathbf{X} \in S) \equiv \Pr(\min_j X_j \geqslant 1 \text{ или } \min_j(-X_j) \geqslant 1) \leqslant \min \frac{1}{(\mathbf{e}, \mathbf{\Pi}_S^{-1}\mathbf{e})}, \tag{9.1}$$

$$\Pr(\mathbf{X} \in S_+) \equiv \Pr(\min X_j \geqslant 1) \leqslant \min \frac{1}{1 + (\mathbf{e}, \mathbf{\Pi}_S^{-1}\mathbf{e})}, \tag{9.2}$$

*где минимум берется по всем главным подматрицам* $\mathbf{\Pi}_s$ *матрицы* $\mathbf{\Pi}$, *таким, что* $\mathbf{\Pi}_s^{-1}\mathbf{e} > 0$ ($\mathbf{e}$ *обозначает вектор, все элементы* *которого равны единице). Равенство может достигаться в* (9.1) *всякий раз, когда граница не больше чем* 1.

Доказательство. Чтобы доказать (9.1), достаточно показать в соответствии с теоремой 8.1, что

$$\min \frac{1}{(\mathbf{e}, \Pi_S^{-1}\mathbf{e})} = \min_{\mathcal{A}} (\mathbf{a}, \Pi\mathbf{a}), \tag{9.3}$$

где $\mathcal{A} = \{\mathbf{a} \mid \mathbf{a} \geqslant 0,\ (\mathbf{a}, \mathbf{e}) = 1\}$.

Обозначим через $\mathbf{b}$ «подвектор» ненулевых компонент вектора $\mathbf{a}$, на котором достигается минимум правой части (9.3).

Так как $(\mathbf{b}, \mathbf{e}) = 1$, то после дифференцирования $(\Pi_S\mathbf{b}, \mathbf{b}) + \lambda[(\mathbf{b}, \mathbf{e}) - 1]$ получаем, что $\mathbf{b}$ удовлетворяет уравнению $2\Pi_S\mathbf{b} + \lambda\mathbf{e} = 0$. Отсюда для $\mathbf{b}$ получаем выражение

$$\mathbf{b} = \frac{\Pi_S^{-1}\mathbf{e}}{(\mathbf{e}, \Pi_S^{-1}\mathbf{e})},$$

откуда следует (9.3). Утверждение о точности границы в (9.1) следует из теоремы 8.1.

Неравенство (9.2) выводится таким же способом. Детальное доказательство опускаем.

Замечание 9.1. Если

$$\Pi = \begin{pmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{pmatrix},$$

то

$$(\Pi^{-1}\mathbf{e}, \mathbf{e}) \geqslant (\Pi_{11}^{-1}\mathbf{e}, \mathbf{e}), \tag{9.4}$$

так что при применении теоремы 9.1 нет необходимости исследовать все подматрицы $\Pi_S$ матрицы $\Pi$, для которых $\Pi_S^{-1}\mathbf{e} > 0$.

Чтобы проверить (9.4), положим $\Pi^{-1} = \mathbf{A}$. Тогда (9.4) эквивалентно неравенству

$$(\mathbf{e}, \mathbf{A}\mathbf{e}) \geqslant (\mathbf{e}, (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})\,\mathbf{e}),$$

которое справедливо в силу очевидного неравенства

$$((\mathbf{A}_{22}\mathbf{e} + \mathbf{A}_{21}\mathbf{e}),\ \mathbf{A}_{22}^{-1}(\mathbf{A}_{22}\mathbf{e} + \mathbf{A}_{21}\mathbf{e})) \geqslant 0.$$

В следующих примерах мы даем значение $(\mathbf{e}, \Pi_S^{-1}\mathbf{e})$, которое достигает минимума в (9.1) и (9.2).

Пример 9.1. Если $n = 2$ и $EX_1^2 = \pi_{11} = \pi_1^2$, $EX_2^2 = \pi_{22} = \pi_2^2$, то

$$(\mathbf{e}, \Pi_S^{-1}\mathbf{e}) = \begin{cases} \pi_1^{-2}, & \pi_1^2 \leqslant \pi_{12}, \\ (\pi_1^2 + \pi_2^2 - 2\pi_{12})/(\pi_1^2\pi_2^2 - \pi_{12}^2), & \pi_1^2 > \pi_{12}. \end{cases}$$

Пример 9.2. Если $\Pi = \pi^2[(1 - \rho)\,\mathbf{I} + \rho\mathbf{e}\mathbf{e}']$, то $\Pi^{-1}\mathbf{e} > 0$ и $(\mathbf{e}, \Pi^{-1}\mathbf{e}) = n[\pi^2(1 + (n - 1)\rho]^{-1}$.

Глава XIV

# НЕРАВЕНСТВА ЧЕБЫШЕВСКОГО ТИПА ДЛЯ СУММ СЛУЧАЙНЫХ ВЕЛИЧИН И НЕЛИНЕЙНЫХ ЗАДАЧ

В §§ 1—3 мы выводим некоторые чебышевские неравенства для сумм независимых случайных величин. Неравенства этого типа восходят к Бернштейну [1924], хотя в более «сыром» виде встречаются и ранее в различных источниках. Результат Бернштейна помогает установить вид закона повторного логарифма для сумм испытаний Бернулли (обобщения можно найти у Лоева [1963] и Реньи [1962]). Другим классическим неравенством теории вероятностей того же типа является неравенство Колмогорова (см. Лоев [1963]). Недавно появилось много новых работ, посвященных улучшениям неравенства Бернштейна. Вклад в эту область внесли Беннет [1962], Хефдинг [1955], Хефдинг и Шрикханд [1955], Бирнбаум, Раймонд и Цукерман [1947] и Самуэльс [1965].

В § 1 мы выделяем четыре случая неравенств чебышевского типа для сумм независимых случайных величин. Доказательства приводимых здесь теорем содержатся в § 2. В § 3 обсуждается гипотеза Самуэльса.

Полученные в первых трех параграфах границы для сумм независимых случайных величин не являются, вообще говоря, точными. В § 4 мы рассматриваем обобщение результата Дубинса и Севиджа [1965 b] относительно сумм случайных величин, которые не обязательно являются независимыми.

Параграфы 6, 7 посвящены нескольким неравенствам, связанным с некоторыми нелинейными вариационными задачами Мы изучаем эти примеры, чтобы указать на применимость обобщенной леммы Неймана—Пирсона в решении некоторых задач чебышевского типа.

## § 1. Формулировка теорем

Мы хотим определить точные границы для вероятности

$$\Pr(\overline{X} - \mu \geqslant t), \qquad t > 0, \tag{1.1}$$

где $n\overline{X} = \sum_{i=1}^{n} X_i$, $X_i$ $(i = 1, 2, \ldots, n)$ — независимые случайные величины и $E\overline{X} = \mu$. Всякая граница для (1.1), очевидно, заключает в себе границу для $\Pr(-\overline{X} + \mu \geqslant t)$ и, следовательно, границу для $\Pr(|\overline{X} - \mu| \geqslant t)$.

Обращаясь к соотношениям (1.5) и (3.2) гл. XII, мы выводим, что

$$\Pr(|\overline{X} - \mu| \geqslant t) \leq \frac{\sigma^2}{nt^2}, \tag{1.2}$$

$$\Pr(|\overline{X} - \mu| \geqslant t) \leq \frac{\sigma^2}{\sigma^2 + nt^2} \tag{1.3}$$

(см. (1.4) гл, XII), где $nE(\overline{X} - \mu)^2 = \sigma^2$. При выводе (1.2) и (1.3) предполагается, что случайные величины $X_i$, $i = 1, 2, \ldots, n$, имеют конечные дисперсии. Если $X_i \geqslant 0$, $i = 1, 2, \ldots, n$, то неравенство Маркова, примененное к $\overline{X}$, приводит к результату

$$\Pr(\overline{X} - \mu \geqslant t) \leq \frac{\mu}{\mu + t}. \tag{1.4}$$

Если случайные величины не ограничены и независимы, то границы (1.2), (1.3) и (1.4) точны. Это можно проверить, взяв $X_2 = X_3 = \ldots = X_n = 0$ и определив $X_1$ подходящим образом. Например, в (1.3) равенство достигается, если положить распределение $X_1$ состоящим из масс $nt^2/(\sigma^2 + nt^2)$ и $\sigma^2/(\sigma^2 + nt^2)$, сосредоточенных в точках $(n\mu t - \sigma^2)/t$ и $n(\mu + t)$ соответственно. Если, однако, в добавление к моментным ограничениям и предположению о независимости мы наложим дальнейшее ограничение о том, что $X_i$ «собственно ограничены», то неравенства (1.2)—(1.4) уже не будут точными. Это замечание станет очевидным, если мы сравним (1.3) и (1.8) для малых значений $t$.

Методы гл. XII были подходящими при выводе неравенств чебышевского типа с ограничениями, имеющими вид моментных условий или условий гладкости. Если требовать точных границы для (1.1), то вид анализа, применявшийся до сих пор, не является подходящим для трактовки ограничений, включающих независимость $X_1, X_2, \ldots, X_n$ плюс условия ограниченности.

Метод, используемый в следующих параграфах (называемый впредь методом Бернштейна), состоит в том, что характеристическая функция множества $\overline{X} - \mu \geqslant t$ мажорируется экспоненциальной функцией $\exp[h(X - \mu - t)]$ ($h$ — положительный параметр). Таким образом, мы выводим

$$\Pr(\overline{X} - \mu \geqslant t) = \Pr(S_n - ES_n \geqslant nt) \leq E\exp[h(S_n - ES_n - nt)] =$$

$$= \exp(-hnt) \prod_{i=1}^{n} E\exp\{h(X_i - EX_i)\} \quad (S_n = n\overline{X}). \tag{1.5}$$

Последнее равенство есть следствие допущения взаимной независимости случайных величин $X_i$, $i = 1, 2, \ldots, n$.

При некоторых условиях на случайные величины $X_i$, $i = 1, 2, \ldots, n$, мы ограничим каждый из сомножителей, а затем минимизируем полученное выражение по отношению к параметру $h$.

Укажем на тот тривиальный факт, что если на случайные величины $X_i$, $i = 1, 2, \ldots, n$, не наложено никаких предположений ограниченности, то максимум $E \exp\{hX_i\}$ при фиксированном $h$ равен $+\infty$. Таким образом, подход Бернштейна представляет интерес только для случайных величин, которые «собственно ограничены». Процедура, набросанная выше, весьма слабо использует предположение о независимости и не гарантирует, что получаемая граница будет точной.

Приведем простейшую теорему, доказываемую описанным методом.

Теорема 1.1 (Хефдинг [1963]). *Пусть $X_1, X_2, \ldots, X_n$ — независимые случайные величины, подчиненные ограничениям $0 \leqq X_i \leqq 1$, $i = 1, 2, \ldots, n$, и $E\overline{X} = \mu$, где $n\overline{X} = X_1 + X_2 + \ldots + X_n$. Тогда при $0 \leqq t < 1 - \mu$ выполняется неравенство*

$$\Pr(\overline{X} - \mu \geqslant t) \leqq \left\{\left(\frac{\mu}{\mu + t}\right)^{\mu+t}\left(\frac{1-\mu}{1-\mu-t}\right)^{1-\mu-t}\right\}^n \leqq \tag{1.6}$$

$$\leqq \exp\{-nt^2 g(\mu)\} \leqq \tag{1.7}$$

$$\leqq \exp\{-2nt^2\}, \tag{1.8}$$

где

$$g(\mu) = \begin{cases} \dfrac{1}{1-2\mu} \log\left(\dfrac{1-\mu}{\mu}\right), & 0 < \mu < \dfrac{1}{2}, \\ \dfrac{1}{2\mu(1-\mu)}, & \dfrac{1}{2} \leqq \mu < 1. \end{cases}$$

Замечание 1.1. Если вместо $0 \leqq X_i \leqq 1$. $i = 1, 2, \ldots, n$, предполагается $a \leqq X_i \leqq b$, то $\mu$ и $t$ заменяется соответственно на $(\mu - a)/(b - a)$ и $t/(b - a)$.

Замечание 1.2. Граница (1.6) есть наилучшая из границ, которые могут быть выведены на основе (1.5). В самом деле, эта граница есть минимум правой части (1.5), оцениваемый по отношению к $h$ в том случае, когда каждое $X_i$, $i = 1, 2, \ldots, n$, имеет распределение $\Pr(X_i = 0) = 1 - \Pr(X_i = 1) = 1 - \mu$, $i = 1, 2, \ldots, n$.

Замечание 1.3. Как указано ранее, предположение типа $0 \leqq X_i \leqq 1$, $i = 1, 2, \ldots, n$, является существенным для данной теоремы. В самом деле, если случайные величины имеют конечные математические ожидания и предполагается только, что $X_i$ неотрицательны, то неравенство Маркова (1.4) является точным, так как равенство достигается при $\Pr(X_i = 0) = 1$, $i = 1, 2, \ldots, n$, и $\Pr(X_1 = 0) = 1 - \Pr(X_1 = m(\mu + t)) = \mu(\mu + t)^{-1}$.

Распространение теоремы 1.1 на случай, когда границы для каждой случайной величины различны, дается следующей теоремой.

Теорема 1.2 (Хефдинг [1963]). *Пусть* $X_1, X_2, \ldots, X_n$ — *независимые случайные величины, подчиненные ограничениям* $a_i \leqslant X_i \leqslant b_i, i = 1, 2, \ldots, n$. *Тогда при* $t > 0$

$$\Pr(\overline{X} - \mu \geqslant t) \leqslant \exp\left\{\frac{-2n^2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right\}. \tag{1.9}$$

Замечание 1.4. Доказательство (1.9) более прямое, чем доказательство (1.8), однако в теореме 1.2 мы не имеем параллельных неравенств, аналогичных (1.6) и (1.7).

В следующей теореме мы налагаем только односторонние границы на случайные величины $X_i$. Мы далее будем предполагать, что дисперсии этих величин конечны, так что $nE(\overline{X} - \mu)^2 = n^{-1}E(S_n - ES_n)^2 = \sigma^2$.

Теор ма 1.3. (Хефдинг [1963]). *Пусть* $X_1, X_2, \ldots, X_n$ — *независимые случайные величины, удовлетворяющие условиям* $EX_i = 0$ *и* $X_i \leqslant b$, $i = 1, 2, \ldots, n$. *Тогда при* $0 < t < b$

$$\Pr(\overline{X} \geqslant t) \leqslant$$

$$\leqslant \left\{\left(1 + \frac{bt}{\sigma^2}\right)^{-(1+[bt/\sigma^2])\sigma^2(b^2+\sigma^2)^{-1}} \left(1 - \frac{t}{b}\right)^{-(1-t/b)b^2(b^2+\sigma^2)^{-1}}\right\}^n \leqslant \tag{1.10}$$

$$\leqslant \exp\left\{-\frac{nt}{b}\left[\left(1 + \frac{\sigma^2}{bt}\right)\log\left(1 + \frac{bt}{\sigma^2}\right) - 1\right]\right\}. \tag{1.11}$$

Замечание 1.5. Как и в замечании 1.2, следует отметить, что граница (1.10) является наилучшей границей, которая может быть выведена с использованием (1.5). В самом деле, это есть минимум по $h$ правой части (1.5), когда $X_i$ имеют распределение

$$\Pr\left(X_i = -\frac{\sigma^2}{b}\right) = \frac{b^2}{b^2 + \sigma^2},$$

$$\Pr(X_i = b) = \frac{\sigma^2}{b^2 + \sigma^2}, \qquad i = 1, 2, \ldots, n.$$

Замечание 1.6. Предположение о том, что $X_i \leqslant b$ существенно в данной теореме, так как соотношение (1.3) дает точную границу при снятии этого условия, т. е. в этом случае равенство в (1.3) достигается при $X_2 = X_n = 0$ и $X_1$, имеющем распределение $\Pr(X_1 = nt) = 1 - \Pr(X_1 = -\sigma^2/t) = \sigma^2/[\sigma^2 + nt^2]$.

В следующей теореме мы получаем границу (1.11) при несколько иных предположениях.

Теорема 1.4. (Беннет [1962)]. *Пусть* $X_1, \ldots, X_n$ — *независимые случайные величины, подчиненные ограничениям* $EX_i = 0$ *и*

$E|X_i|^r = v_i^r \leq M^{r-2}\sigma_i^2$, $r = 2, 3, \ldots$ *Тогда*

$$\Pr(\bar{X} \geqslant t) \leq \exp\left\{-\frac{nt}{M}\left[\left(1 - \frac{\sigma^2}{Mt}\right)\log\left(1 + \frac{Mt}{\sigma^2}\right) - 1\right]\right\}, \quad (1.12)$$

*где* $n\sigma^2 = \sum_{i=1}^{n} \sigma_i^2 = \sum_{i=1}^{n} EX_i^2$.

## § 2. Доказательство теорем 1.1 — 1.4

Нам потребуется следующий общий результат, обслуживающий доказательства всех четырех теорем. Пусть $X$ — случайная величина, подчиненная ограничению $a \leqslant X \leqslant b$. При любом вещественном $h$ функция $e^{hx}$ выпукла, так что

$$Ee^{hX} \leq E\left(\frac{b-X}{b-a}e^{ha} + \frac{X-a}{b-a}e^{hb}\right) = \frac{b-EX}{b-a}e^{ha} + \frac{EX-a}{b-a}e^{hb}. \quad (2.1)$$

Отметим, что $(1, x, e^{hx})$ есть $T$-система и правая часть (2.1) есть просто значение $Ee^{hx}$ по отношению к верхнему главному представлению $(c_0, c_1) = (1, EX)$.

Доказательство теоремы 1.1. Соотношение (1.5) в сочетании с (2.1), где $a = 0$ и $b = 1$, приводит к неравенствам

$$\Pr(\bar{X} - \mu \geqslant t) \leq \exp[-hn(t+\mu)] \prod_{i=1}^{n} Ee^{hX_i} \leq$$

$$\leq \exp[-hn(t+\mu)] \prod_{i=1}^{n} (1 - \mu_i + \mu_i e^h),$$

где $\mu_i = EX_i$ и $n\mu = \sum_{i=1}^{n} \mu_i$. Применение неравенства между средним арифметическим и средним геометрическим дает

$$\Pr(\bar{X} - \mu \geqslant t) \leq \exp[-hn(t+\mu)](1 - \mu + \mu e^h)^n. \quad (2.2)$$

Простое вычисление теперь показывает, что правая часть этого неравенства минимизируется по $h$ при

$$h = \log\frac{(1-\mu)(\mu+t)}{(1-\mu-t)\mu},$$

и, подставляя значение в (2.2), получим (1.6).

Неравенство (1.7) получается, если записать правую часть (1.6) в виде $\exp(-nt^2G(t,\mu))$, где

$$G(t,\mu) = \frac{\mu+t}{t^2}\log\frac{\mu+t}{\mu} + \frac{1-\mu-t}{t^2}\log\frac{1-\mu-t}{1-\mu},$$

и проверить, что функция $g(\mu)$ есть минимум $G(t, \mu)$ по отношению к $t$, $t \in [0, 1-\mu)$. С этой целью запишем

$$t^2 \frac{\partial}{\partial t} G(t, \mu) = H\left(\frac{t}{1-\mu}\right) - H\left(\frac{t}{\mu+t}\right),$$

где $H(x) = (1-2x^{-1}) \log(1-x)$. При $0 < x < 1$, $H'(x) > 0$, откуда следует, что $H(x)$ возрастает. При условии $0 \leq t(1-\mu)^{-1} < 1$ мы теперь имеем $\partial/\partial t G(t, \mu) > 0$ в том и только в том случае, если $t(1-\mu)^{-1} > t/(\mu+t)$ или $t > 1-2\mu$. Следовательно, если $1-2\mu > 0$, то $G(t, \mu)$ имеет минимум при $t = 1-2\mu$, откуда следует справедливость доказываемого утверждения при $0 < \mu < 1/2$. Справедливость утверждения при $1/2 \leq \mu < 1$ следует из того факта, что если $1-2\mu < 0$, то $G(t, \mu)$ минимизируется при $t = 0$.

Неравенство (1.8) есть следствие соотношения $g(\mu) \geqslant g(1/2) = 2$.

Доказательство теоремы 1.2. Повторяя доказательство теоремы 1.1, получим

$$E \exp(h(X_i - \mu_i)) \leq e^{-h\mu_i}\left(\left(\frac{b_i-\mu_i}{b_i-a_i}\right)e^{ha_i} + \left(\frac{\mu_i-a_i}{b_i-a_i}\right)e^{hb_i}\right) = e^{L(h_i)},$$

где

$$L(h_i) = -h_i p_i + \log(1 - p_i + p_i e^{h_i}),$$

$$h_i = h(b_i - a_i), \quad p_i = \frac{\mu_i - a_i}{b_i - a_i}, \qquad i = 1, 2, \ldots, n.$$

Первые две производные от $L(h_i)$ есть

$$L'(h_i) = -p_i + \frac{p_i}{(1-p_i)e^{-h_i} + p_i}, \quad L''(h_i) = \frac{p_i(1-p_i)e^{-h_i}}{((1-p_i)e^{-h_i} + p_i)^2}.$$

Функция $L''(h_i)$ имеет вид $u(1-u)$, где

$$u = \frac{(1-p_i)e^{-h_i}}{(1-p_i)e^{-h_i} + p_i},$$

так что $L''(h_i) \leq 1/4$ и, следовательно,

$$L(h_i) \leq L(0) + L'(0) h_i + \frac{1}{8} h_i^2 = \frac{1}{8} h^2 (b_i - a_i)^2.$$

Следовательно,

$$\Pr(\bar{X} - \mu \geqslant t) \leq e^{-hnt} \prod_{i=1}^{n} E \exp[h(X_i - \mu_i)] \leq$$

$$\leq \exp\left(-hnt + \frac{1}{8} h^2 \sum_{i=1}^{n} (b_i - a_i)^2\right),$$

и так как последнее выражение имеет минимум при

$$h = 4nt\left(\sum_{i=1}^{n}(b_i - a_i)^2\right)^{-1},$$

то мы получаем, наконец, что

$$\Pr(\bar{X} - \mu \geqslant t) \leq \exp\left\{-\frac{2n^2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right\}.$$

Доказательство теоремы 1.3. Пусть $X$ — случайная величина такая, что $EX = 0$, $EX^2 = \sigma^2$ и $X \leqslant b$. Верхнее главное представление для $(1, 0, \sigma^2)$ по отношению к системе $(1, x, x^2)$ имеет массы $b^2(b^2 + \sigma^2)^{-1}$ и $\sigma^2(b^2 + \sigma^2)^{-1}$, сосредоточенные в $-\sigma^2 b^{-1}$ и $b$ соответственно. Так как функции $1, x, x^2, e^{hx}$ образуют $T$-систему, то заключаем, что

$$Ee^{hX} \leqslant \frac{b^2}{b^2 + \sigma^2}\exp\left(-\frac{\sigma^2 h}{b}\right) + \frac{\sigma^2}{b^2 + \sigma^2}e^{bh}.$$

Если $EX_i^2 = \sigma_i^2$ и $n\sigma^2 = \sum_{i=1}^{n}\sigma_i^2$, то

$$\Pr(\bar{X} \geqslant t) \leqslant e^{-hnt}\prod_{i=1}^{n}Ee^{hX_i} \leqslant$$

$$\leqslant e^{-hnt}\prod_{i=1}^{n}\left(\frac{b^2}{b^2 + \sigma_i^2}\exp\left(-\frac{\sigma_i^2 h}{b}\right) + \frac{\sigma_i^2}{b^2 + \sigma_i^2}e^{bh}\right) \leqslant$$

$$\leqslant \exp\left(-hnt + \sum_{i=1}^{n}f\left(\frac{\sigma_i^2}{b}\right)\right),$$

где

$$f(u) = \log\left(\frac{1}{1+u}e^{-cu} + \frac{u}{1+u}e^{c}\right) \qquad (c = bh).$$

Если мы покажем, что $f(u)$ вогнута при $u \geqslant 0$, то сможем заключить, что

$$\frac{1}{n}\sum_{i=1}^{n}f\left(\frac{\sigma_i^2}{b^2}\right) \leqslant f\left(\frac{1}{n}\frac{\sum_{i=1}^{n}\sigma_i^2}{b^2}\right) =$$

$$= f\left(\frac{\sigma^2}{b^2}\right) = \log\left(\frac{b^2}{b^2 + \sigma^2}\exp\left(-\frac{\sigma^2 h}{b}\right) + \frac{\sigma^2}{b^2 + a^2}e^{bh}\right).$$

Следовательно,

$$\Pr(\overline{X} \geqslant t) \leqslant \left(\frac{b^2}{b^2+\sigma^2}\exp(-(t+\sigma^2/b)h) + \frac{\sigma^2}{b^2+\sigma^2}e^{(b-t)h}\right)^n \quad (2.3)$$

и минимум правой части достигается при

$$h = \frac{b}{b^2+\sigma^2}\log\left[\frac{(\sigma^2+tb)b}{\sigma^2(b-t)}\right].$$

Подставляя это значение в (2.3), после некоторых простых преобразований получаем (1.10).

Для того чтобы проверить, что $f(u)$ вогнута при $u \geqslant 0$, мы запишем $f(u) = c + \log f_1(y)$, где $y = 1 + u$, а $f_1(y) = y^{-1}e^{-cy} - y^{-1} + 1$. Тогда

$$f_1^2(y) f''(u) = f_1(y) f_1''(y) - (f_1'(y))^2,$$

$$f_1''(y) = -2y^{-3}e^{-cy}\left(e^{cy} - 1 - cy - \frac{1}{2}c^2y^2\right),$$

так что $f_1''(y)$ отрицательна при $y > 0$. Так как $f_1(y) > 0$ при $y > 1$, то отсюда следует, что $f''(u) < 0$ при $u > 0$. Этим завершается доказательство (1.10).

Мы докажем теперь, что (1.10) есть наилучшая граница, которая может быть получена при использовании (1.5). Чтобы доказать это, мы припишем $X_i$ распределение, определенное следующим образом:

$$\Pr\left(X_i = -\frac{\sigma^2}{b}\right) = 1 - \Pr(X_i = b) = \frac{b^2}{b^2+\sigma^2}, \quad i = 1, 2, \ldots, n, \quad (2.4)$$

так что

$$\exp(-hnt)\prod_{i=1}^{n} E\exp(hX_i)$$

сводится к правой части (2.3).

Как упоминалось, минимум этого выражения есть правая часть (1.10).

Чтобы завершить доказательство теоремы 3.1, достаточно показать, что $\Pr(\overline{X} \geqslant t)$ ограничено правой частью (1.11). Так как $f''(u) < 0$ при $u > 0$, имеем $f(u) \leqslant f(0) + f'(0)u = (e^{bh} - 1 - bh)u$, так что из (1.5) мы выводим

$$\Pr(\overline{X} \geqslant t) \leqslant \exp\left[-hnt + n(e^{bh} - 1 - bh)\frac{\sigma^2}{b^2}\right]. \quad (2.5)$$

Окончательный результат получается минимизацией этого выражения по $h$.

Доказательство теоремы 1.4. Как и в предшествующих доказательствах, мы начинаем с неравенства

$$\Pr(\overline{X} \geqslant t) \leqslant e^{-hnt} \prod_{i=1}^{n} Ee^{hX_i}. \tag{2.6}$$

Имеем

$$Ee^{hX_i} = 1 + \frac{1}{2} h^2\sigma_i^2 \sum_{r=2}^{\infty} \frac{h^{r-2}EX_i^r}{\frac{1}{2} r!\sigma_i^2} \leq \exp\left(\frac{1}{2} h^2\sigma_i^2 F_i\right),$$

где

$$F_i = \sum_{r=2}^{\infty} \frac{h^{r-2}EX_i^r}{\frac{1}{2} r!\sigma_i^2}.$$

Подстановка в (2.6) дает

$$\Pr(\overline{X} \geqslant t) \leq e^{-hnt} \prod_{i=1}^{n} \exp\left[\frac{1}{2} h^2\sigma_i^2 F_i\right] \leq \exp\left[-hnt + \frac{1}{2} h^2 n\sigma^2 F\right],$$

где $F = \max F_i$.

Так как $r$-й абсолютный момент $X_i$ не превосходит $M^{r-2}\sigma_i^2$, мы получаем

$$F_i = \sum_{r=2}^{\infty} \frac{h^{r-2}EX_i^r}{\frac{1}{2} r!\sigma_i^2} \leq \sum_{r=2}^{\infty} \frac{(hM)^{r-2}}{\frac{1}{2} r!} = 2\,\frac{e^{hM} - 1 - hM}{(hM)^2}.$$

Следовательно,

$$\Pr(\overline{X} \geqslant t) \leq \exp\left(-hnt + n\sigma^2 \frac{(e^{hM} - 1 - hM)}{M^2}\right) =$$

$$= \left[\exp\left(-ht + \frac{\sigma^2}{M^2}(e^{hM} - 1 - hM)\right)\right]^n.$$

Минимизируя последнее выражение по $h$, получим неравенство

$$\Pr(\overline{X} \geqslant t) \leq \exp\left(-\frac{ht}{M}\left[\left(1 + \frac{\sigma^2}{Mt}\right)\log\left(1 + \frac{Mt}{\sigma^2}\right) - 1\right]\right).$$

## § 3. Гипотеза Самуэльса

Пусть $X_1, X_2, \ldots, X_n$ обозначают $n$ независимых величин с заданными средними $EX_i = \nu_i$, $i = 1, 2, \ldots, n$. Пусть $S$ обозначает класс случайных величин вида $S_n = \sum_{i=1}^{n} X_i$. Без потери общности для определенности положим с этого момента, что $0 \leqslant \nu_1 \leqslant \leqslant \nu_2 \leqslant \ldots \leqslant \nu_n$.

Требуется определить наименьшую верхнюю границу для

$$\Pr(S_n \geqslant \lambda) \tag{3.1}$$

при $S_n$, пробегающих класс $\mathcal{C}$. Можно предположить, что $\sum_{i=1}^{n} \nu_i < \lambda$, так как иначе граница равна единице и достигается, если взять $\Pr(X_i = \nu_i) = 1, \quad i = 1, 2, \ldots, n$.

Существенная трудность для попытки применения анализа гл. XII состоит в том, что множество $\mathcal{C}$, вообще говоря, не выпукло (см. Самуэльс [1966]).

Конечно, грубые границы для (3.1) можно получить, если рассматривать $S_n$ как отдельную величину, игнорируя ее специальную структуру. Например, применение к $S_n$ неравенства Маркова дает тот тривиальный результат, что $\sum_{i=1}^{n} \nu_i/\lambda$ есть верхняя граница для $\Pr(S_n \geqslant \lambda)$. Если по крайней мере два средних значения положительны, то при $n \geqslant 2$ $S_n$ не может сосредоточивать все массы в 0 и $\lambda$; следовательно, верхняя граница не может быть достигнута и, таким образом, $\sum_{i=1}^{n} \nu_i/\lambda$ не есть наименьшая верхняя граница.

В § 2 отмечалось, что подход Бернштейна уместен только в том случае, когда случайные величины $X_i$ ограничены в собственном смысле. Мы вскоре показали, что неотрицательные случайные величины $X_i$ можно считать ограниченными посредством $\lambda$. Однако даже с границей $\lambda$ на отдельные $X_i$ подход Бернштейна дает границу, превосходящую границу, получаемую из неравенства Маркова. В самом деле, применяя (1.5) и (2.1), мы получим границу

$$e^{-h\lambda} \prod_{i=1}^{n} \left\{1 + \frac{\nu_i}{\lambda}(e^{h\lambda} - 1)\right\} > e^{-h\lambda} \left\{1 + \frac{\sum_{i=1}^{n} \nu_i (e^{h\lambda} - 1)}{\lambda}\right\} =$$

$$= \frac{\sum_{i=1}^{n} \nu_i}{\lambda} + e^{-h\lambda} \left\{1 - \frac{\sum_{i=1}^{n} \nu_i}{\lambda}\right\} > \frac{\sum_{i=1}^{n} \nu_i}{\lambda}.$$

Точная граница для (3.1) предложена Самуэльсом [1966]. Предполагается для данной задачи, что

$$\sup_{\mathcal{C}} \Pr(S_n \geqslant \lambda) = \max\{P_0, P_1, \ldots, P_{n-1}\}, \tag{3.2}$$

где для $k = 0, 1, \dots, n-1$

$$P_k = 1 - \prod_{i=k+1}^{n} \left(1 - \frac{\nu_i}{\lambda - \sum_{j=1}^{k} \nu_j}\right) \tag{3.3}$$

$\left(\text{при } k = 0 \text{ определим } \sum_{j=1}^{k} \nu_j = 0\right)$. Значение $P_k$ достигается, если

$$\Pr(X_i = \nu_i) = 1, \qquad i = 1, 2, \dots, k, \tag{3.4}$$

$$\Pr(X_i = 0) = 1 - \Pr\left(X_i = \lambda - \sum_{j=1}^{k} \nu_j\right) =$$

$$= 1 - \frac{\nu_j}{\lambda - \sum_{j=1}^{k} \nu_j}, \quad i = k+1, \dots, n. \tag{3.5}$$

Мы представим здесь ряд замечаний и лемм, установленных Самуэльсом [1966], делающих очевидной обоснованность этой гипотезы. В частности, мы даем полное доказательство для случая $n = 2$. Гипотеза будет также проверена в случае $n = 3$. Если $n = 1$, то результат есть просто неравенство Маркова.

Мы начнем с утверждения о том, что можно ограничить внимание подклассом $\mathcal{C}(\lambda)$ класса $\mathcal{C}$, где отдельные случайные величины в сумме $S_n = \sum_{i=1}^{n} X_i$ ограничены $\lambda$. Чтобы проверить это утверждение, мы запишем

$$\Pr(S_n \geqslant \lambda) = \int_0^{\infty} \{1 - F^{(n-1)}(\lambda - x)\}\, dF_n(x), \tag{3.6}$$

где $F^{(n-1)}(x)$ и $F_n(x)$ есть функции распределения $\sum_{i=1}^{n-1} X_i$ и $X_n$ соответственно. Из (3.6) ясно, что мы можем переместить в $\lambda$ любую массу $X_n$, расположенную выше $\lambda$ без изменения значения интеграла, а затем переместить массу, расположенную ниже $\nu_n$, вверх так, чтобы восстановить математическое ожидание, равное $\nu_n$. Последнее перемещение может только увеличить значение интеграла.

Мы можем, далее, ограничить рассмотрение подклассом $\mathcal{C}_2(\lambda)$ класса $\mathcal{C}(\lambda)$, где распределения случайных величин $X_i$, $i = 1, 2, \dots, n$, дискретны и состоят самое большое из двух скачков. Для этого мы и используем теорему 6.4 гл. III, которая утверждает, что если $g(x)$ — любая непрерывная функция на $[0, \lambda]$, то

$$\sup_{F \in \mathcal{F}_\nu} \int_0^\lambda g(x)\, dF(x) = \sup_{F \in \mathcal{E}_\nu} \int_0^\lambda g(x)\, dF(x), \tag{3.7}$$

где $\mathcal{F}_\nu$ — выпуклое множество функций распределения на $[0, \lambda]$ со средним $\nu$, и $\mathcal{E}_\nu$ обозначает подкласс $\mathcal{F}_\nu$, состоящий из всех ступенчатых функций $F \in \mathcal{F}_\nu$, которые имеют не более двух скачков. Без потери общности мы можем предположить, что функция $1 - F^{(n-1)}(\lambda - x)$ непрерывна справа. Неубывающая ограниченная функция $1 - F^{(n-1)}(\lambda - x)$ может быть поточечно аппроксимирована непрерывными функциями. Из (3.7) следует, что распределение случайной величины $X_n$ можно считать сосредоточенным в двух точках. По симметрии такие же выводы применимы к $X_1, \ldots, X_{n-1}$.

Дальнейшая редукция класса $\mathcal{C}_2(\lambda)$ осуществляется в следствии к приводимой ниже лемме.

Лемма 3.1. *Пусть* $X_i$, $i = 1, 2, \ldots, n$, *имеют массы* $(\nu_i - a_i)/(b_i - a_i)$ *и* $(b_i - \nu_i)/(b_i - a_i)$ *в точках* $b_i$ *и* $a_i$ *соответственно и* $0 \leqslant a_i \leqslant \nu_i \leqslant b_i \leqslant \lambda$. *Если для некоторого* $i$

$$\Pr(S_n = \lambda \mid X_i = b_i) = 0,$$

*где* $S_n \in \mathcal{C}_2(\lambda)$ *и* $a_i < \nu_i < b_i$, *то существует случайная величина* $S_n' \in \mathcal{C}_2(\lambda)$ *такая, что* $\Pr(S_n' \geqslant \lambda) > \Pr(S_n \geqslant \lambda)$.

Доказательство. Определим

$$f(b) = \frac{b - \nu_i}{b - a_i} \Pr(S_n \geqslant \lambda \mid X_i = a_i) + \frac{\nu_i - a_i}{b - a_i} \Pr(S_n \geqslant \lambda \mid X_i = b);$$

тогда $f(b_i) = \Pr(S_n \geqslant \lambda)$. По предположению, существу т $\delta > 0$ такое, что $\Pr(S_n \geqslant \lambda \mid X_i = b)$ постоянно при $b_i - \delta < b \leq b_i$. В этом интервале

$$f'(b) = \frac{\nu_i - a_i}{(b - a_i)^2} [\Pr(S_n \geqslant \lambda \mid X_i = a_i) - \Pr(S_n \geqslant \lambda \mid \lambda_i = b_i)].$$

Это выражение строго отрицательно, если только не выполняется равенство

$$\Pr(S_n \geqslant \lambda \mid X_i = a_i) = \Pr(S_n \geqslant \lambda \mid X_i = b_i). \tag{3.8}$$

Следовательно, лемма справедлива, если (3.8) не выполняется. Если (3.8) выполняется, то мы можем, очевидно, положить $X_i \equiv \nu_i$ без изменения значения $\Pr(S_n \geqslant \lambda)$, так как

$$\Pr(S_n \geqslant \lambda \mid X_i = b) \geqslant \Pr(S_n \geqslant \lambda \mid X_i = \nu_i) \geqslant \Pr(S_n \geqslant \lambda \mid X_i = a_i).$$

Более того, $a_i < \nu_i < b_i$ влечет

$$\Pr(S_n \geqslant \lambda \mid X_i = b_i) - \Pr(S_n \geqslant \lambda \mid X_i = a_i) \geqslant \\ \geqslant \Pr(S_n = \lambda \mid X_i = \nu_i) \geqslant 0,$$

откуда $\Pr(S_n = \lambda \mid X_i = \nu_i) = 0$.

Следовательно, если $X_i \equiv \nu_i$, то мы имеем $\Pr(S_n = \lambda \mid X_i = \nu_i, X_j = b_j) = 0$ для всех $j \neq i$. Возникают две возможности: либо

$$\Pr(S_n \geqslant \lambda \mid X_i = \nu_i, X_j = b_j) = \Pr(S_n \geqslant \lambda \mid X_i = \nu_i, X_j = a_j)$$

для всех $j \neq i$, либо строгое неравенство имеет место для некоторого $j$. В последнем случае, как мы показали, мы можем увеличить $\Pr(S_n \geqslant \lambda)$, уменьшая $b_j$, а в первом случае можно положить все $X_j$ равными их средним без изменения $\Pr(S_n \geqslant \lambda)$. Но, по предположению, $\sum\limits_1^n \nu_i < \lambda$, так что в этом случае $\Pr(S_n \geqslant \lambda) = 0$, что, очевидно, не является верхней границей. Доказательство леммы закончено.

Пусть $\mathcal{C}_2^*(\lambda)$ обозначает подкласс $\mathcal{C}_2(\lambda)$, в котором $\Pr(S_n = \lambda \mid X_i = b_i) > 0$ для всех $i$ (если $X_i \equiv \nu_i$, мы можем условиться, что $b_i = \nu_i$).

Следствие 3.1. *Верхняя граница для* $\Pr(S_n \geqslant \lambda)$ *по отношению к* $\mathcal{C}_2^*(\lambda)$ *является верхней границей по отношению к* $\mathcal{C}$. *Более того, если верхняя граница достигается для единственного* $S_n$ *в классе* $\mathcal{C}_2^*(\lambda)$, *то* $S_n$, *принадлежащее* $\mathcal{C}$, *на котором достигается верхняя граница, также единственно.*

Докажем теперь справедливость гипотезы в случае $n = 2$.

Теорема 3.1.

$$\max_{S_2 \in \mathcal{C}} \Pr(S_2 \geqslant \lambda) = \begin{cases} \dfrac{\nu_2}{\lambda - \nu_1}, & \nu_1 + \nu_2 \leqslant \lambda \leqslant \lambda_0, \\[2ex] 1 - \left(1 - \dfrac{\nu_1}{\lambda}\right)\left(1 - \dfrac{\nu_2}{\lambda}\right), & \lambda \geqslant \lambda_0, \end{cases}$$

где $\lambda_0 = \frac{1}{2}\left[\nu_1 + 2\nu_2 + \sqrt{\nu_1^2 + 4\nu_2^2}\right]$.

Доказательство. Из обсуждения, проведенного выше, мы знаем, что можно ограничить внимание случайными величинами $X_1$ и $X_2$, каждая из которых имеет самое большее два различных значения. Пусть $X_1$ имеет возможные значения $a_1$ и $b_1$, а $X_2$ — значения $a_2$ и $b_2$. Тогда возможные значения $S_2 = X_1 + X_2$ есть $a_1 + a_2$, $a_1 + b_2$, $b_1 + a_2$ и $b_1 + b_2$. Можно различить следующие четыре возможности:

(1) $a_1 + b_2 < \lambda, \quad b_1 + a_2 < \lambda, \quad b_1 + b_2 \geqslant \lambda$,
(2) $a_1 + b_2 \geqslant \lambda, \quad b_1 + a_2 < \lambda, \quad b_1 + b_2 > \lambda$,
(3) $a_1 + b_2 < \lambda, \quad b_1 + a_2 \geqslant \lambda, \quad b_1 + b_2 > \lambda$,
(4) $a_1 + b_2 \geqslant \lambda, \quad b_1 + a_2 \geqslant \lambda, \quad b_1 + b_2 > \lambda$.

По лемме 3.1 можно исключить из рассмотрения случаи (2) и (3).

Для случая (1) мы должны иметь $b_1 + b_2 = \lambda$ и

$$\Pr(S_n \geqslant \lambda) = \frac{\nu_1 - a_1}{b_1 - a_1} \frac{\nu_2 - a_2}{b_2 - a_2}.$$

Легко видеть, что это выражение возрастает, когда $a_1$ и $a_2$ убывают, так что можно положить $a_1 = a_2 = 0$. Остается минимизировать $b_1 b_2 = b_1(\lambda - b_1)$ при $\nu_1 \leqslant b_1 \leqslant \lambda - \nu_2$. Если $\nu_1 < \nu_2$, то минимум достигается только при $b_1 = \nu_1$ и $b_2 = \lambda - \nu_1$, что дает первое значение в утверждении теоремы. Если $\nu_1 = \nu_2$, то минимум достигается только в точках на концах.

В случае (4) лемма 3.1 дает нам, что $a_1 + b_2 = \lambda$ и $b_1 + a_2 = \lambda$. Тогда

$$\Pr(S_n \geqslant \lambda) = 1 - \left(\frac{b_1 - \nu_1}{b_1 - a_1}\right)\left(\frac{b_2 - \nu_2}{b_2 - a_2}\right).$$

Если мы максимизируем это выражение по $a_1, a_2, b_1$ и $b_2$ при $a_1 \leqslant \nu_1 \leqslant b_1$, $a_2 \leqslant \nu_2 \leqslant b_2$, то получим, что максимум достигается при $b_1 = b_2 = \lambda$ и $a_1 = a_2 = 0$, что дает второе значение.

Сравнение двух возможных значений заканчивает доказательство.

## § 4. Неравенство для сумм случайных величин без предположения независимости

В этом параграфе мы применим теорему 2.1 гл. XII для получения границ некоторых вероятностей для сумм случайных величин, которые не обязательно являются независимыми. Следующая теорема обобщает результат Дубинса и Севиджа [1965]. Мы используем прямые методы в противоположность упомянутым авторам, которые вывели неравенство как следствие некоторых обширных исследований в теории игровых систем.

Теорема 4.1. *Пусть $\mathcal{M}$ обозначает класс последовательностей $\{X_n\}_1^\infty$ случайных величин, которые образуют мартингал, т. е. $EX_1=0$, $E(X_n \mid X_{n+1}, \ldots, X_1) = 0$, $n \geqslant 2$. Пусть $\alpha$ и $\beta$ — положительные числа.*

*Определим*

$$\varphi(n) = \inf_{\mathcal{M}} \Pr\{X_1 + \ldots + X_k < \alpha(V_1 + \ldots + V_k\} + \beta, \quad 1 \leqslant k \leqslant n\},$$

*где $V_1 = EX_1^2$, $V_k = E(X_k^2 \mid X_{k-1}, \ldots, X_1)$, $k \geqslant 2$. Тогда*

$$\varphi(1) = \frac{4\alpha\beta}{4\alpha\beta + 1}, \tag{4.1}$$

$$\varphi(n) = \frac{\alpha\beta f_{n-1}(4)}{\alpha\beta f_{n-1}(4) + 1}, \qquad n \geqslant 2,$$

*где*

$$f_1(s) = \left(\frac{2s}{s+1}\right)^2, \quad f_n(s) = f_1(f_{n-1}(s)), \qquad n \geqslant 2.$$

Простым следствием этой теоремы является следующий результат Дубинса и Фридмана [1965].

Следствие 4.1. *При условиях теоремы* 4.1

$$\Pr(X_1+\ldots+X_n \geqslant \alpha(V_1+\ldots+V_n)+\beta \text{ для некоторого } n \geqslant 1) \leqslant \leqslant \frac{1}{1+\alpha\beta}. \quad (4.2)$$

Доказательство. Вероятность в (4.2) может быть записана как

$$1-\Pr(X_1+\ldots+X_n<\alpha(V_1+\ldots+V_n)+\beta \text{ для всех } n\geqslant 1) = = 1-\lim_{n\to\infty}\Pr(X_1+\ldots+X_k<\alpha(V_1+\ldots+V_k)+\beta,\ 1\leqslant k\leqslant n) \leqslant \leqslant 1-\lim_{n\to\infty}\frac{\alpha\beta f_{n-1}(4)}{\alpha\beta f_{n-1}(4)+1}.$$

Функция $f_1(s)=\left(\frac{2s}{s+1}\right)^2$ строго возрастает и ее единственная положительная неподвижная точка ($f_1(s)=s$) есть $s=1$. Отсюда следует, что $\lim_{n\to\infty} f_{n-1}(s)=1$ для всех $s>$ и, следовательно,

$$\Pr(X_1+\ldots+X_n\geqslant\alpha(V_1+\ldots+V_n)+\beta \text{ для некоторого } n\geqslant 1)\leqslant \leqslant 1-\frac{\alpha\beta}{1+\alpha\beta}=\frac{1}{1+\alpha\beta}.$$

Доказательство теоремы 4.1. Достаточно доказать соотношения (4.1) в случае $\beta=1$. В противном случае мы делаем замену $Y_k=\beta^{-1}X_k$, $k\geqslant 1$. Тогда

$$E(Y_n\,|\,Y_{n-1},\ldots,Y_1)=0,$$

$$W_n=E(Y_n^2\,|\,Y_{n-1},\ldots,Y_1)=\beta^{-2}V_n,$$

и, следовательно,

$$\Pr(X_1+\ldots+X_k<\alpha(V_1+\ldots+V_k)+\beta,\ 1\leqslant k\leqslant n)=$$

$$=\Pr(Y_1+\ldots+Y_k<\alpha'(W_1+\ldots+W_k)+1,\ 1\leqslant k\leqslant n)$$

$$(\alpha'=\alpha\beta).$$

Доказательство осуществим индукцией по $n$. При $n=1$ мы сначала найдем наибольшую нижнюю границу $L(\nu)$ для $\Pr(X< <\alpha\nu+1)$, где $X$ — случайная величина, подчиненная условиям $EX=0$, $EX^2=\nu$.

Пусть $I_A$ обозначает характеристическую функцию множества $A=\{x\,|\,x<\alpha\nu+1\}$, и пусть $\mathcal{P}$ обозначает класс квадратичных многочленов $P(t)=a+bt+ct^2$, удовлетворяющих неравенству

$P(t) \leqslant I_A(t)$ для всех $t$. Метод теоремы 2.1 гл. XII дает нам, что $L(\nu) = \sup_{\mathcal{P}} (a + c\nu)$. Эта верхняя грань легко вычисляется и дает $L(\nu) = (a\nu + 1)^2/[(\alpha\nu + 1)^2 + \nu]$. Минимум функции $L(\nu)$ при $\nu > 0$ достигается при $\nu = \alpha^{-1}$. Отсюда мы заключаем, что $\varphi(1) = L(\alpha^{-1}) = 4\alpha/(4\alpha + 1)$, что завершает доказательство в случае $n = 1$.

Чтобы осуществить индукцию, положим, что функция $\varphi(n)$ имеет вид $\varphi(n) = s\alpha/(s\alpha + 1)$ при некотором $s$. Доказательство будет завершено как только мы докажем, что $\varphi(n+1) = \alpha f_1(s)/(\alpha f_1(s) + 1)$. С этой целью запишем

$$\varphi(n+1) =$$
$$= \inf \Pr (X_1 + \ldots + X_k < \alpha(V_1 + \ldots + V_k) + 1, \quad 1 \leqslant k \leqslant n+1) =$$
$$= \inf \int_{-\infty}^{\infty} \Pr (X_1 + \ldots + X_k <$$
$$< \alpha(V_1 + \ldots + V_k) + 1, \ 1 \leqslant k \leqslant n+1 \mid X_1 = \xi)\, dF_1(\xi),$$

где нижняя грань берется по всем мартингалам из $n+1$ случайных величин, описанным в формулировке теоремы, а $F_1$ обозначает функцию распределения случайной величины $X_1$. Теперь значение $\varphi(n+1)$ сводится к

$$\varphi(n+1) = \inf \int_{-\infty}^{\infty} \Pr (X_2 + \ldots + X_k < \alpha(V_2(\xi) + \ldots + V_k(\xi)) +$$
$$+ \alpha\nu + 1 - \xi, \ 2 \leqslant k \leqslant n+1 \mid X_1 = \xi)\, dF_1(\xi),$$

где $\nu = V_1 = \int_{-\infty}^{\infty} \xi^2 dF_1(\xi)$,

$$V_k(\xi) = E(X_k^2 \mid X_{k-1}, \ldots, X_2, X_1 = \xi), \qquad 2 \leqslant k \leqslant n+1.$$

Применяя предположение индукции для произвольного $\beta > 0$ к случайным величинам $X_2, \ldots, X_{n+1}$ и функции вероятности $\Pr(\cdot \mid X_1 = \xi)$, выводим, что

$$\varphi(n+1) = \inf \int_{-\infty}^{\alpha\nu+1} \frac{s(\alpha\nu + 1 - \xi)\alpha}{s(\alpha\nu + 1 - \xi)\alpha + 1}\, dF_1(\xi), \tag{4.3}$$

где нижняя грань берется над классом функций распределения $F_1$, удовлетворяющих условиям $\int_{-\infty}^{\infty} \xi dF_1(\xi) = 0$ и $\int_{-\infty}^{\infty} \xi^2 dF_1(\xi) = \nu$ при произвольном $\nu$.

Применение теоремы 2.1 гл. XII показывает, что нижняя грань в (4.3) для фиксированного $\nu$ есть $\alpha s(\alpha\nu+1)/[(s+1)\alpha\nu+1+\alpha s(\alpha\nu+1)^2]$. Минимум этого выражения достигается при $\nu=(s-1)/\alpha(s+1)$. Отсюда следует, что $\varphi(n+1)=\alpha f_1(s)/[\alpha f_1(s)+1]$. Доказательство закончено.

## § 5. Нелинейные задачи

Теория и примеры в гл. XII и XIII были частично посвящены задаче определения экстремальных значений интеграла

$$\int_a^b \Omega(t)\, d\sigma(t), \tag{5.1}$$

где меры $\sigma$ принадлежат множеству

$$V(c)=V(c_0, c_1, \ldots, c_n)=\left\{\sigma \mid c_i=\int_a^b u_i d\sigma,\ i=0, 1, \ldots, n\right\}, \tag{5.2}$$

а $\{u_i\}_0^n$ есть $T$-система на $[a, b]$.

В гл. VIII мы изучали задачу определения экстремальных значений

$$\int_a^b \Omega(t)\, \varphi(t)\, d\mu(t), \tag{5.3}$$

где $\varphi$ таково, что $0\leqslant\varphi(t)\leqslant 1$ и удовлетворяет моментным условиям вида $c_i=\int_a^b u_i(t)\,\varphi(t)\, d\mu(t)$, $i=0, 1, \ldots, n$. В обоих случаях интегралы (5.1) и (5.3) линейны по переменным $\sigma$ и $\varphi$ соответственно.

Рассмотрим теперь задачу максимизации некоторых интегральных выражений, которые являются нелинейными функционалами от $\varphi$. Например, мы изучим задачу характеризации функции $\varphi$, на которой достигается минимум функционала

$$\int_0^1 [\varphi(x)-\lambda(x)]^2\, dx$$

при некоторых ограничениях на гладкость и моменты.

Обычный метод решения таких задач состоит в формировании уравнений Эйлера и попытке их решить. Однако инструменты этой дисциплины не могут быть непосредственно применены из-за неравенств-ограничений, наложенных на функцию $\varphi$. Было бы необходимо построить решение, составляя его отчасти из решения, даваемого вариационным исчислением при игнорировании неравенств — ограничений, а затем подгоняя его под ограничения. Эта процедура была формализована в принципе максимума Понтрягина (см. Понтрягин и др. [1969]).

Мы следуем более элементарному пути, использующему результаты § 8 гл. VIII. Для удобства ссылок мы вновь сформулируем главный результат, который используется в нашем дальнейшем обсуждении.

Предположим, что $\{u_i\}_0^n$ есть $T$-система $\mu$-интегрируемых функций на открытом интервале $(a, b)$ и что $\Omega$ есть еще одна интегрируемая функция на $(a, b)$. Мера $\mu$ предполагается безатомной (ср. гл. VIII, § 12) и такой, что мера $(a, b)$ конечна. Определим

$$\Phi_{n+1} = \{\mathbf{c} = (c_0, c_1, \dots, c_n) \mid c_i = \int u_i \varphi \, d\mu, \ \ i = 0, 1, \dots n\}, \tag{5.4}$$

где $0 \leq \varphi \leq 1$ и

$$V(\mathbf{c}) = \{\varphi \mid c_i = \int u_i \varphi \, d\mu, \ \ i = 0, 1, \dots, n, \ \ 0 \leq \varphi \leq 1\}. \tag{5.5}$$

Теорема 5.1. *Если* $\mathbf{c} \in \operatorname{Int} \Phi_{n+1}$, *то максимум*

$$\max_{\varphi \in V(\mathbf{c})} \int \Omega \varphi \, d\mu \tag{5.6}$$

*достигается на функции* $\bar{\varphi} \in V(\mathbf{c})$, *характеризуемой в терминах некоторого подходящего многочлена* $\bar{u}(t)$ *соотношениями*

$$\bar{\varphi}(t) = \begin{cases} 1, & \Omega(t) > \bar{u}(t), \\ 0, & \Omega(t) < \bar{u}(t), \end{cases} \tag{5.7}$$

*и если* $\Omega(t) = \bar{u}(t)$, то $\bar{\varphi}(t)$ *неопределено.*

На эту теорему обычно ссылаются как на обобщенную лемму Неймана—Пирсона.

Замечание 5.1. При замене max на min в (5.6) теорема остается верной, если обратить оба неравенства в (5.7).

Замечание 5.2. Теорема также остается верной, если мы предположим только, что функции $u_i(t)$, $i = 0, 1, \dots, n$, интегрируемы на $(a, b)$ и опустим условие, что функции $\{u_i\}_{i=0}^n$ образуют $T$-систему.

Замечание 5.3. Если ограничения $0 \leq \varphi \leq 1$ заменены обобщенными условиями

$$\lambda_1(t) \leq \varphi(t) \leq \lambda_2(t),$$

где $\lambda_1$ и $\lambda_2$ — непрерывные на $[a, b]$ функции, то (5.7) заменяется на

$$\varphi(t) = \begin{cases} \lambda_2(t), & \Omega(t) > \bar{u}(t), \\ \lambda_1(t), & \Omega(t) < \bar{u}(t), \end{cases}$$

и теорема остается верной.

Область применения метода, которым мы будем пользоваться, включает обширное множество выпуклых и вогнутых функционалов с аргументом, удовлетворяющим подходящим моментным условиям в добавление к условиям ограниченности.

Мы резюмируем идею этого метода следующим образом. Задача состоит в отыскании функции $\varphi_0$, которая максимизирует (минимизирует) интегральное выражение, включающее вогнутую (выпуклую) функцию $\varphi$, где $\varphi$ подчинена линейным ограничениям. Преобразуя интегральный функционал от $\varphi$, мы сводим данную задачу к задаче, для которой применяется решение Неймана—Пирсона с одной оговоркой: функция $\Omega(t)$ в (5.6) теперь обычно включает в себя решение $\varphi_0(t)$. Таким образом, решение определено неявно и нет уверенности, что оно свободно от противоречия. Однако с помощью тщательного анализа и искусственных приемов часто можно достичь корректного решения, которое удовлетворяет наложенным ограничениям. Это осуществляется введением подходящих параметров, определяемых природой задачи, и дальнейшим их отысканием таким образом, чтобы удовлетворялись ограничения. Наконец, должна быть проверена корректность полученного решения.

Использование леммы Неймана — Пирсона для решения вариационных задач встречается у многих авторов; особенно у Беллмана, Гликсберга и Гросса [1954], Данскина [1955], Коупманса [1956], Эрроу и Карлина [1958] (гл. 4—7), Карлина, Пруитта и Мадоу [1963] и др.

В §§ 6 и 7 мы представим детальную трактовку двух примеров, что даст достаточно полную иллюстрацию метода.

## § 6. Пример I. Максимум и размах выборки

Пусть $X_1, X_2, \ldots, X_n$ — независимые наблюдения с функций распределения $F(x)$ на $(-\infty, +\infty)$. Распределение величины $U = \max(X_1, X_2, \ldots, X_n)$ есть $G(u) = [F(u)]^n$, и, таким образом, $\int_{-\infty}^{+\infty} u\,dG(u)$ есть среднее значение максимального наблюдения в выборке. Наша задача — найти вид распределения, которое максимизирует $\int_{-\infty}^{+\infty} u\,dG(u)$ при моментных условиях

$$\int_{-\infty}^{+\infty} x\,dF(x) = 0, \tag{6.1}$$

$$\int_{-\infty}^{+\infty} x^2 dF(x) = 1. \tag{6.2}$$

Другой вариант задачи такого типа: максимизировать среднее значение размаха выборки из $n$ независимых наблюдений, т. е. максимизировать $E(U - V)$, где $V = \min(X_1, X_2, \ldots, X_n)$ при условиях (6.1) и (6.2). Мы покажем, как можно решить целый класс задач такого типа, используя обобщенную лемму Неймана — Пирсона (теорема 5.1).

Общая задача может быть поставлена следующим образом: найти функцию распределения $F(x)$, которая максимизирует

$$\int_{-\infty}^{+\infty} x\, d\Phi(F(x)) \tag{6.3}$$

при условиях (6.1) и (6.2), где $\Phi(u)$ строго выпуклая возрастающая функция своего аргумента. Если бы мы искали минимум, то потребовали бы от $\Phi(u)$ строгой вогнутости. Специальные случаи этой общей задачи изучались Рустажи [1957] в случае, когда $\Phi(u)$ выпуклое, и Хартли и Дэвидом [1954] в случае, когда эта функция вогнута (см. также Моригьюти [1951]). В первой задаче $\Phi(F(x))$ соответствует $[F(x)]^n$, а во второй—$[F(x)]^n + \{1 - F(x)\}^n$.

Мы будем предполагать для простоты, что $\Phi'(x)$ существует и конечна всюду на $[0, 1]$.

Наш метод будет состоять в преобразовании первоначальной задачи в задачу, которая может быть решена с помощью обобщенной леммы Неймана—Пирсона. С этой целью мы сначала ограничимся распределениями $F$, определенными на интервале $-I < x < I$, где $I > \max\{\Phi'(0), \Phi'(1)\}$, т. е. $F(-I) = 0$ и $F(I) = 1$, причем функция $F$ предполагается непрерывной в $-I$ и в $I$.

Во-первых, мы преобразуем соотношения (6.1) и (6.2) к более удобному виду. Интегрирование (6.1) и (6.2) по частям дает

$$\int_{-\infty}^{\infty} F(x)\, dx = I, \tag{6.4}$$

$$\int_{-\infty}^{\infty} xF(x)\, dx = \frac{I^2 - 1}{2}. \tag{6.5}$$

Такое же интегрирование по частям дает

$$\int_{-I}^{I} x\, d\Phi(F(x)) = x\Phi(F(x)) \Big|_{-I}^{I} - \int_{-I}^{I} \Phi(F(x))\, dx. \tag{6.6}$$

Таким образом, максимизация $\int_{-I}^{I} x\, d\Phi(F(x))$ при ограничениях (6.1) и (6.2) эквивалентна минимизации $\int_{-I}^{I} \Phi(F(x))\, dx$ при ограничениях (6.4) и (6.5).

Предположим далее, что $F_0$ — распределение, на котором достигается минимум, и $F$ — любое другое распределение, удовлетворяющее ограничениям (6.4) и (6.5). Пусть $F_\lambda = \lambda F_0 + (1-\lambda) F, 0 \leqslant \lambda \leqslant 1$. Тогда $F_\lambda$, очевидно, также удовлетворяют этим ограничениям. Далее определим $H(\lambda)$ как

$$H(\lambda) = \int_{-l}^{l} \Phi(\lambda F_0(x) + (1-\lambda) F(x))\, dx.$$

Ясно, что $H(\lambda)$ выпукло, так как $\Phi$ выпукло. Кроме того, по определению $F_0$, $H(\lambda)$ достигает своего минимума при $\lambda = 1$. Таким образом, $H'(1) \leqslant 0$.

Обратное также верно. Так как, если $H'(1) \leqslant 0$ для всех $F$, удовлетворяющих (6.4) и (6.5), то $H(\lambda)$, являясь выпуклой функцией, должна достигать своего минимума при $\lambda = 1$. Следовательно, $F_0$ — распределение, на котором достигается минимум.

Мы показали, что нахождение $F_0$, минимизирующего $\int_{-l}^{l} \Phi(F(x))$, эквивалентно нахождению $F_0$, для которого $H'(1) \leqslant 0$ среди всех допустимых $F$. Из $H'(1) \leqslant 0$ следует, что $\int_{-l}^{l} \Phi'(F_0(x))(F_0 - F)dx \leqslant 0$.

Ищем, таким образом, распределение $F_0$, которое удовлетворяет

$$\int_{-l}^{l} \Phi'(F_0) F_0\, dx = \min_F \int_{-l}^{l} \Phi'(F_0) F\, dx, \tag{6.7}$$

где минимум оценивается по отношению к классу всех функций распределения, удовлетворяющих (6.4) и (6.5).

Наша задача теперь подобна задаче, рассмотренной в теореме 5.1, за исключением того, что ограничения на $F$ более жесткие, и требуется, чтобы $F$ возрастала в добавление к условиям (6.4) и (6.5). Так или иначе, мы расширяем класс допустимых функций, рассматривая вместо функций распределения $F$ функции $\varphi$, для которых $0 \leqslant \varphi \leqslant 1$. В терминах $\varphi$ мы ищем функцию $\varphi_0$, $0 \leqslant \varphi_0 \leqslant 1$, которая удовлетворяет соотношению

$$\int_{-l}^{l} \Phi'(\varphi_0)\, \varphi_0 dx = \min_{\varphi} \int_{-l}^{l} \Phi'(\varphi_0)\, \varphi\, dx \tag{6.8}$$

при ограничениях

$$\int_{-l}^{l} \varphi(x)\, dx = l, \tag{6.9}$$

$$\int_{-l}^{l} x\varphi(x)\, dx = \frac{l^2 - 1}{2}. \tag{6.10}$$

Таким образом, наша задача свелась к задаче типа тех, которые могут быть решены с использованием обобщенной леммы Неймана—Пирсона. Используя характеризацию теоремы 5.1, получим, что существуют константы $\eta$ и $\mu$, для которых

$$\varphi_0(x) = \begin{cases} 1, \text{ если } \Phi'(\varphi_0(x)) < \eta + \mu x, \\ 0, \text{ если } \Phi'(\varphi_0(x)) > \eta + \mu x, \\ \text{соответствующая функция между} \\ 0 \text{ и } 1, \text{ если } \Phi'(\varphi_0(x)) = \eta + \mu x. \end{cases} \tag{6.11}$$

Это решение определено неявным образом, так как $\varphi_0$ встречается в правой части.

Наша очередная задача состоит в определении $\eta$, $\mu$ и конструировании подходящей функции при $\Phi'(\varphi_0(x)) = \eta + \mu x$ так, чтобы убедиться в выполнении ограничений (6.9) и (6.10) и в том, что в (6.11) не содержится противоречия.

Сначала мы изучим природу множества тех $x$, для которых $\Phi'(\varphi_0) = \eta + \mu x$. Переход к обратной функции дает

$$\varphi_0(x) = \Phi'^{-1}(\eta + \mu x). \tag{6.12}$$

Далее $\Phi'$ — строго возрастающая функция (так как $\Phi$, по предположению, строго выпукла); следовательно, $\Phi'^{-1}$ есть однозначная строго возрастающая функция. Если $\mu > 0$, то $\Phi'^{-1}(\eta + \mu x)$ — строго возрастающая функция $x$. Мы предположим далее, что $\mu > 0$.

Теперь убедимся, что $0 \leqslant \varphi_0 \leqslant 1$. Определим $a_1, a_2$ соотношениями $\Phi'^{-1}(a_1) = 0$ и $\Phi'^{-1}(a_2) = 1$. Наконец, положим

$$\varphi_0(x) = \begin{cases} 0, \text{ если } \eta + \mu x \leqslant a_1, \\ \Phi'^{-1}(\eta + \mu x), \text{ если } a_1 < \eta + \mu x < a_2, \\ 1, \text{ если } \eta + \mu x \geqslant a_2 \end{cases} \tag{6.13}$$

Отметим, что (6.13) согласуется с (6.11), так как:

(a) при $\eta + \mu x \leqslant a_1$ и (6.11) и (6.13) задают значение $\varphi_0(x) = 0$;

(b) при $a_1 < \eta + \mu x < a_1$ и (6.11) и (6.13) задают $\varphi_0$, удовлетворяющее соотношению $\varphi_0(x) = \Phi'^{-1}(\eta + \mu x)$;

(c) при $\eta + \mu x \geqslant a_2$ и (6.11) и (6.13) задают значение $\varphi_0(x) = 1$.

Теперь мы используем ограничения (6.9) и (6.10) для того, чтобы получить информацию относительно $\eta$ и $\mu$. Уравнение (6.9) для функции (6.13) имеет вид

$$\int\limits_{(a_1-\eta)/\mu}^{(a_2-\eta)/\mu} \Phi'^{-1}(\eta + \mu x)\, dx + \int\limits_{(a_2-\eta)/\mu}^{l} 1 dx = l.$$

Производя замену переменных $y=\eta+\mu(x)$ в первом интеграле и упрощая, получим

$$\eta=a_2-\int_{a_1}^{a_2}\Phi'^{-1}(y)\,dy. \tag{6.14}$$

Подобным образом (6.10) эквивалентно условию

$$\frac{1}{\mu_2}\int_{a_1}^{a_2}(y-\eta)\,\Phi'^{-1}(y)\,dy=\frac{1}{2}\left(\frac{a_2-\eta}{\mu}\right)^2-\frac{1}{2}. \tag{6.15}$$

Подставляя $\eta$ из (6.14) в (6.15), получим

$$\mu^2=\left(2a_2-\int_{a_1}^{a_2}\Phi'^{-1}(y)\,dy\right)\int_{a_1}^{a_2}\Phi'^{-1}(y)\,dy-2\int_{a_1}^{a_2}y\,\Phi'^{-1}(y)\,dy. \tag{6.16}$$

Ниже покажем, что (6.16) всегда будет удовлетворяться (т. е. правая часть всегда будет положительной). Удобно записать правую часть (6.16) в виде

$$\mu^2=2\int_{a_1}^{a_2}(a_2-y)\,\Phi'^{-1}(y)\,dy-\left\{\int_{a_1}^{a_2}\Phi'^{-1}(y)\,dy\right\}^2, \tag{6.17}$$

где $\Phi'^{-1}(y)$ — строго возрастающая функция и $\Phi'^{-1}(a_1)=0$, $\Phi'^{-1}(a_2)=1$.

Найдем минимум правой части (6.16) при постоянном значении $\int_{a_1}^{a_2}\Phi'^{-1}(y)\,dy$, скажем, $\alpha$, т. е. минимизируем $2\int_{a_1}^{a_2}(a_2-y)\,\Phi'^{-1}(y)\,dy$ при ограничении $\int_{a_1}^{a_2}\Phi'^{-1}(y)\,dy=\alpha$ с дальнейшей оговоркой, что $\Phi'^{-1}(y)$ есть строго возрастающая функция от $y$. Если не принимать во внимание требование, что $\Phi'^{-1}$ строго возрастает, то наша задача есть в точности задача, решаемая леммой Неймана—Пирсона. Решение таково:

$$\Phi_0'^{-1}(y)=\begin{cases}1, \text{ если } c+(a_2-y)<0 \text{ (или эквивалентно, если } y>k, \text{ где } k \text{ — некоторая константа)},\\ 0, \text{ если } c+(a_2-y)>0 \text{ (или эквивалентно, если } y<k).\end{cases} \tag{6.18}$$

Соответствующее минимальное значение $\mu^2$ есть

$$2\int_{k}^{a_2}(a_2-y)\,dy-(a_2-k)^2=0.$$

Отметим, что этот минимум не зависит от $\alpha$, так что для всех $\Phi'^{-1}(y)$ будет $\mu^2\geqslant 0$. Следовательно, мы можем определить $\mu$ с помощью (6.17) таким, что $\mu > 0$. Значение $\eta$ находится из (6.14). Решение найдено полностью. Наконец, заметим, что решение (6.13) не зависит от $l$ при условии, что $l$ выбрано достаточно большим. Итак, мы получаем следующий результат.

Теорема 6.1. *Максимум* $\int_{-\infty}^{+\infty} x\,d\Phi(F(x))$ *достигается для распределения* $F_0$, *представленного в* (6.13), *где* $\eta$ *и* $\mu$ *задаются соотношениями* (6.14) *и* (6.16) *соответственно и* $a_1 = \Phi'(0)$, $a_2 = \Phi'(1)$.

Следствие 6.1a. *Если* $X_1, X_2, \ldots, X_n$ $(n \geqslant 2)$ — *независимые наблюдения с функцией распределения* $F$, *а* $U = \max\{X_1, \ldots, X_n\}$, *то*

$$EU = \int x\,dF^n(x) \leq \frac{n-1}{\sqrt{2n-1}} \tag{6.19}$$

*и* (6.19) *превращается в равенство при*

$$F_0(x) = \begin{cases} 0, & 1 + \mu_0 x \leq 1, \\ \left(\dfrac{1+\mu_0 x}{n}\right)^{1/(n-1)}, & 0 < 1 + \mu_0 x \leq n, \\ 1, & n \leq 1 + \mu_0 x, \end{cases}$$

*где* $\mu_0 = (n-1)/\sqrt{2n-1}$.

Замечание 6.1. Неравенство (6.19) может быть получено более прямым способом на основе неравенства Шварца.

Следствие 6.1b. *Если* $V = \min\{X_1, \ldots, X_n\}$, *то максимум среднего значения* $E(U - V)$ *достигается при*

$$F_1(x) = \Phi'^{-1}(\eta_1 + \mu_1 x), \quad -n < \eta_1 + \mu x \leq n,$$

*где*

$$\Phi'(x) = n(x^{n-1} - (1-x)^{n-1}), \qquad \eta_1 = 0,$$

$$\mu_1 = n\left[\frac{2}{2n-1}\left(1 - \frac{1}{\binom{2n-2}{n-1}}\right)\right]^{1/2}.$$

Будем теперь минимизировать интеграл

$$\int_{-l}^{l} x\,d\Phi(F(x)). \tag{6.20}$$

Из (6.6) следует, что это эквивалентно максимизации интеграла

$$\int_{-l}^{l} \Phi(F(x))\,dx. \tag{6.21}$$

Эта максимизация снова осуществляется при ограничениях (6.4) и (6.5).

Так как функция $\Phi$ предполагается выпуклой, максимум должен достигаться в крайней точке. Следовательно, максимизирующее распределение $F$ должно состоять самое большее из трех скачков.

Пусть $F$ имеет скачки $\lambda, \mu$ и $1-\lambda-\mu$ в точках $x, y$ и $z$ $(x<y<z)$, тогда интеграл в (6.21) сводится к

$$\Phi(0)(x+l)+\Phi(\lambda)(y-x)+\Phi(\lambda+\mu)(z-y)+\Phi(1)(l-z). \tag{6.22}$$

Ограничения (6.1) и (6.2) или (6.4) и (6.5) принимают вид

$$\begin{gathered} x\lambda+y\mu+z(1-\lambda-\mu)=0, \\ x^2\lambda+y^2\mu+z^2(1-\lambda-\mu)=1. \end{gathered} \tag{6.23}$$

Мы теперь покажем, что выражение в (6.22) выпукло по переменной $y$ на интервале $(x, z)$. Решая (6.23) относительно $\lambda$ и $\mu$, получим

$$\lambda=\frac{1+yz}{(z-x)(y-x)}, \quad \mu=\frac{-1-xz}{(z-y)(y-x)},$$

$$\lambda+\mu=\frac{z^2-1-z(y+x)}{(z-x)(z-y)}.$$

Обозначая выражение в (6.22) через $g(y)$ с помощью непосредственного дифференцирования, получим формулу

$$\begin{aligned} g''(y) = {} & \Phi''(\lambda)\left(\frac{\partial\lambda}{\partial y}\right)^2(y-x)+\Phi'(\lambda)\frac{\partial^2\lambda}{\partial y^2}(y-x)+2\Phi'(\lambda)\frac{\partial\lambda}{\partial y}+ \\ & +\Phi''(\lambda+\mu)\left(\frac{\partial(\lambda+\mu)}{\partial y}\right)^2(z-y)+ \\ & +\Phi'(\lambda+\mu)\frac{\partial^2(\lambda+\mu)}{\partial y^2}(z-y)-2\Phi'(\lambda+\mu)\frac{\partial(\lambda+\mu)}{\partial y}. \end{aligned}$$

Теперь простое вычисление дает

$$\frac{\partial^2\lambda}{\partial y^2}=\frac{2(1+xz)}{(z-x)(y-x)^3}, \quad \frac{\partial^2(\lambda+\mu)}{\partial y^2}=\frac{-2(1+zx)}{(z-x)(z-y)^3}.$$

После подстановки и упрощения находим

$$g''(y)=\Phi''(\lambda)\left(\frac{\partial\lambda}{\partial y}\right)^2(y-x)+\Phi''(\lambda+\mu)\left(\frac{\partial(\lambda+\mu)}{\partial y}\right)^2(z-y).$$

Это выражение, очевидно, неотрицательно, так как $\Phi''$ неотрицательно по предположению, а $y$ находится внутри интервала $[x, z]$. Итак, функция $g(y)$ выпукла, как и утверждалось. Максимум (6.22) или, что эквивалентно, минимум (6.20) достигается, следовательно, на двухточечном распределении.

Для двухточечного распределения $F$ с массами $p$ и $1-p$ в точках $x$ и $y$ соответственно интеграл в (6.20) может быть представлен в следующем виде:

$$\int_{-l}^{l} t\,d\Phi(F(t)) = y\Phi(1) - x\Phi(0) - (y-x)\Phi(p).$$

Ограничения сводятся к соотношениям

$$xp + y(1-p) = 0, \quad x^2p + y^2(1-p) = 1,$$

так что $y = \sqrt{\frac{p}{1-p}}$ или $p = \frac{y^2}{1+y^2}$ и $x = -\sqrt{\frac{1-p}{p}}$. Используя симметрию между $x$ и $y$, мы можем считать, что $1 \leq y \leq l$ или, что эквивалентно, $1/2 \leq p \leq l^2/(1+l^2)$.

Суммируя сказанное, мы получаем следующий результат.

Теорема 6.2. *Пусть $\Phi$ — выпуклая на $[0, 1]$ функция. Для любого распределения на $[-l, l]$, удовлетворяющего условиям*

$$\int_{-l}^{l} t\,dF(t) = 0, \quad \int_{-l}^{l} t^2 dF(t) = 1,$$

*имеем*

$$\int_{-l}^{l} t\,d\Phi(F(t)) \geqslant$$

$$\geqslant \min_{\frac{1}{2} \leqslant p \leqslant a} \left\{ \sqrt{\frac{p}{1-p}}\,\Phi(1) + \sqrt{\frac{1-p}{p}}\,\Phi(0) - \frac{\Phi(p)}{\sqrt{p(1-p)}} \right\},$$

*где $a = l^2/(1+l^2)$.*

Для специального случая, когда $\Phi(u) = u^n + (1-u)^n$, имеем

$$\int_{-l}^{l} t\,d\Phi(F(t)) \geqslant \min_{\frac{1}{2} \leqslant p \leqslant a} \frac{1-p^n-(1-p)^n}{\sqrt{p(1-p)}}.$$

Отсюда следует, что для выборки размера $n$ из распределения на $[-l, l]$ ожидаемый размах выборки ограничен снизу выражением

$$E(U-V) \geqslant \inf_{\frac{1}{2} \leqslant p \leqslant a} \frac{1-p^n-(1-p)^n}{\sqrt{p(1-p)}}, \tag{6.24}$$

где $a = l^2/(1+l^2)$.

Мы покажем теперь, что нижняя грань в (6.24) достигается, только когда $p = 1/2$ или $p = a$.

Для этого достаточно проверить, что функция

$$g(p) = \frac{1-p^n-(1-p)^n}{\sqrt{p(1-p)}}$$

не имеет локального минимума при $1/2 < p < 1$. Исследование производной

$$g'(p) = \frac{(1-p)^n(1+2(n-1)p) - p^n(1+2(n-1)(1-p)) - (1-2p)}{2(p(1-p))^{3/2}}$$

показывает, что $g'(p) \neq 0$ при $p \in (1/2, 1/2+\delta)$, где $\delta$ положительно и мало и $g'(p) < 0$ для $p$, близких к единице. Следовательно, достаточно показать, что $g'(p)$ или, что эквивалентно, числитель этой производной $l(p)$ обращается в нуль самое большее один раз при $1/2 < p < 1$. Так как $l(p) = 0$ при $p = 1/2$ и 1, то функция $l(p)$ будет, очевидно, обращаться в нуль самое большее один раз в $(1/2, 1)$ при условии, что $l''(p)$ имеет не более одного нуля в $(1/2, 1)$. Непосредственно проверяется, что

$$l''(p) = 2n(n^2-1)[p^{n-1} - (1-p)^{n-1}] - \\ - n(n-1)(2n-1)[p^{n-2} - (1-p)^{n-2}].$$

Для завершения анализа мы теперь покажем, что функция

$$h(p) = \frac{p^{n-1} - (1-p)^{n-1}}{p^{n-2} - (1-p)^{n-2}}, \qquad n \geqslant 3,$$

строго возрастает на $(1/2, 1)$ (при $n = 2$ нетрудно отдельно проверить, что $l''(p)$ строго возрастает). Функция $(p^{n-2} - (1-p)^{n-2})\,h'(p)$ может быть записана в виде

$$p^{2n-4} - (1-p)^{2n-4} - (n-2)(2p-1)\,p^{n-3}(1-p)^{n-3} = \\ = (2p-1)\sum_{i=0}^{n-3}\{p^i(1-p)^i[p^{2n-5-2i} + (1-p)^{2n-5-2i} - (p(1-p))^{n-3-i}]\}.$$

Член $p^{2n-5-2i} + (1-p)^{2n-5-2i}$ является выпуклой функцией на $(0, 1)$ и симметричен относительно $1/2$ и, следовательно, имеет минимум при $p = 1/2$, а член $(p(1-p))^{n-3-i}$ имеет максимум при $p = 1/2$.

Но в точке $p = 1/2$ обе эти функции совпадают. Поэтому функция $h'(p)$ положительна на $(1/2, 1)$, и, следовательно, $h(p)$ строго возрастает. Отсюда следует, что $l''(p)$ обращается в нуль не более одного раза, что и требовалось доказать.

Суммируя полученные результаты, имеем следующую теорему.

Теорема 6.3. *Если $X_1, X_2, \ldots, X_n$ — выборка размера $n$ из распределения $F$ на $[-l, l]$, то*

$$E(U-V) \geqslant \min\begin{cases} 2\left(1-\left(\frac{1}{2}\right)^{n-1}\right), \\ \dfrac{1-p_0^n-(1-p_0)^n}{\sqrt{p_0(1-p_0)}}, \end{cases}$$

*где $U = \max\{X_1, \ldots, X_n\}$, $V = \min\{X_1, \ldots, X_n\}$ и $p_0 = l^2/(1+l^2)$.*

Заметим, что минимум зависит от $l$ в противоположность максимуму в следствии 1b, который не зависит от $l$ при условии, что $l$ выбирается достаточно большим.

## § 7. Пример II

Некоторые специальные аспекты этого примера трактовались Рустажи [1957] и Бирнбаумом и Клосе [1955].

(A) Мы хотим минимизировать

$$J(F)=\int_0^1 (F(x)-\lambda x)^2\,dx \tag{7.1}$$

над классом $\mathcal{F}_p$ распределений $F$ на $[0, 1]$, удовлетворяющих

$$\int_0^1 x\,dF=p \quad \text{или} \quad \int_0^1 F(x)\,dx=1-p, \tag{7.2}$$

где $0<p<1$ и $\lambda>0$ заданы

Если $F_0$ — минимизирующее распределение и $I(\alpha)$ — функционал, определяемый соотношением

$$I(\alpha)=J(\alpha F_0+(1-\alpha)F), \quad 0\leqq\alpha\leqq 1, \quad F\in\mathcal{F}_p,$$

то, повторяя анализ, ведущий к (6.7), мы получим, что функция распределения $F_0$, минимизирующая (7.1), минимизирует также интеграл

$$\int_0^1 (F_0(x)-\lambda x)\,F(x)\,dx. \tag{7.3}$$

Отсюда по теореме 5.1 следует, что $F_0$ должно иметь вид

$$F_0(x)=\begin{cases} 0, & \text{если } F_0(x)-\lambda x>k, \\ 1, & \text{если } F_0(x)-\lambda x<k, \\ \lambda x+k, & \text{если } F_0(x)-\lambda x=k. \end{cases} \tag{7.4}$$

Уравнение (7.4) означает, что $F_0$ состоит из линейного отрезка $\lambda x+k$ и двух крышеобразных кусков, присоединенных к концам отрезка $\lambda x+k$.

Теперь, если $c_1$ и $c_2$ определены посредством равенств

$$\lambda c_1+k=0, \quad \lambda c_2+k=1 \tag{7.5}$$

и $b_1=\max\{0, c_1\}$, $b_2=\min\{1, c_2\}$, то из (7.2) следует

$$\frac{\lambda}{2}(b_2^2-b_1^2)+k(b_2-b_1)+1-b_2=1-p. \tag{7.6}$$

Требуется исследовать три возможности:

(1) $b_1 = 0, \quad b_2 = c_2$;

(2) $b_1 = c_1, \quad b_2 = 1$;

(3) $b_1 = c_1, \quad b_2 = c_2$, или $b_1 = 0$, $b_2 = 1$.

Для случая (1) мы заключаем из (7.5), что $c_2 = \lambda^{-1}(1-k)$, и (7.6) дает

$$\frac{\lambda}{2} c_2^2 + kc_2 + 1 - c_2 = 1 - p.$$

Отсюда следует, что

$$\frac{\lambda}{2}\left(\frac{1-k}{\lambda}\right)^2 + k\left(\frac{1-k}{\lambda}\right) + 1 - \left(\frac{1-k}{\lambda}\right) = 1 - p, \text{ или } k = 1 \pm \sqrt{2\lambda p}.$$

Условие, что $c_2 = \lambda^{-1}(1-k) = \mp\sqrt{2p/\lambda}$ неотрицательно, влечет

$$c_2 = \sqrt{2p/\lambda}, \quad k = 1 - \sqrt{2\lambda p}. \tag{7.7}$$

Кроме того, так как $k \geqslant 0$ и $c_2 \leq 1$, мы имеем, что $2\lambda p \leq 1$ и $2p \leq \lambda$ или

$$p \leq \min\left\{\frac{1}{2\lambda}, \frac{\lambda}{2}\right\}.$$

Теперь, если (a) $\lambda \geqslant 1$ и $p \leqslant 1/2\lambda$ или (b) $\lambda \leq 1$ и $p \leq \lambda/2$, то мы можем легко проверить, что функция

$$F_0(x) = \begin{cases} \lambda x + 1 - \sqrt{2\lambda p}, & 0 \leq x \leq \sqrt{2p/\lambda}, \\ 1, & x \geqslant \sqrt{2p/\lambda} \end{cases}$$

удовлетворяет (7.2) и (7.4) и, следовательно, является решением.

Случаи (2) и (3) трактуются подобным образом.

(**B**) В этой части мы снова хотим минимизировать функционал

$$\int_0^1 (F(x) - \lambda x)^2\, dx, \quad \lambda > 0, \tag{7.8}$$

при ограничениях

$$\int_0^1 F(x)\, dx = 1 - p, \quad 0 < p < 1, \tag{7.9}$$

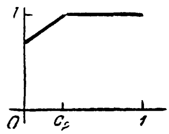
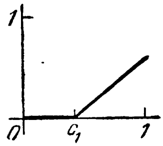
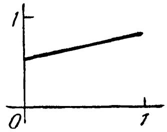
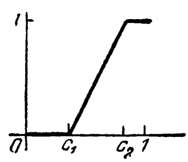
с дополнительным ограничением

$$F(x) \geqslant x. \tag{7.10}$$

В случае (1) части A, представленном на первой диаграмме, ясно, что решение остается тем же самым, так как уже $F_0(x) \geqslant x$.

В случае (2) условия $\lambda \geqslant 1$ и $1 - 1/2\lambda \leq p$ или $\lambda \leq 1$ и $1 - \lambda/2 \leq p$ оба влекут, что $p > 1/2$, если только $\lambda \neq 1$, когда $p \geqslant 1/2$. Если

ТАБЛИЦА 7

| Условие | График $F_0$ | Значения параметров |
|---|---|---|
| $\lambda \geqslant 1$ и $p \leqslant \frac{1}{2\lambda}$ или $\lambda \leqslant 1$ и $p \leqslant \frac{\lambda}{2}$ |   | $k = 1 - \sqrt{2\lambda p}$ $c_2 = \sqrt{\frac{2p}{\lambda}}$ |
| $\lambda \geqslant 1$ и $1 - \frac{1}{2\lambda} \leqslant p$ или $\lambda \leqslant 1$ и $1 - \frac{1}{2\lambda} \leqslant p$ |   | $k = \sqrt{2\lambda(1-p)} - \lambda$ $c_1 = 1 - \sqrt{\frac{2(1-p)}{\lambda}}$ |
| $\lambda \leqslant 1$ и $\frac{\lambda}{2} \leqslant p \leqslant 1 - \frac{\lambda}{2}$ |   | $k = 1 - p - \frac{\lambda}{2}$ |
| $\lambda \geqslant 1$ и $\frac{1}{2\lambda} \leqslant p \leqslant 1 - \frac{1}{2\lambda}$ |   | $k = \frac{1}{2} - p$ $c_1 = p - \frac{1}{2\lambda}$ $c_2 = p + \frac{1}{2\lambda}$ |

$p > 1/2$, то не существует ни одного распределения $F$, удовлетворяющего (7.9) и (7.10) а если $p = 1/2$ и $\lambda = 1$, то $F(x) = x$ есть единственное распределение, удовлетворяющее (7.9) и (7.10), и оно, таким образом, является тривиальным решением.

В случае (3), который соответствует последним двум диаграммам, можно предполагать, что решение будет получено заменой всякой части $F_0$, для которой $F_0(x) < x$, на $F_0(x) = x$ с соответствующим изменением параметров.

Поэтому, если $\lambda > 1$ и $1/2\lambda \leq p \leq 1/2$, мы рассмотрим решение в виде

$$F(x) = \begin{cases} x, & 0 \leq x \leq c_1, \\ \lambda x + k, & c_1 \leq x \leq c_2, \\ 1, & c_2 \leq x \leq 1. \end{cases} \tag{7.11}$$

Если предположить, что функция $F(x)$ непрерывна, то $c_1 = k(1-\lambda)^{-1}$ и $c_2 = \lambda^{-1}(1-k)$. Ограничение (7.9) выражается соотношением

$$1-p = \int_0^1 F(x)\,dx = \frac{c_1^2}{2} + \frac{\lambda}{2}(c_2^2 - c_1^2) + k(c_2 - c_1) + 1 - c_2.$$

Подстановка значений $c_1$ и $c_2$ и упрощение приводят к уравнению $k^2 + 2(\lambda-1)k + (\lambda-1)(2\lambda p - 1) = 0$, решения которого суть $k = 1 - \lambda \pm \sqrt{\lambda(\lambda-1)(1-2p)}$. Из условия $c_2 \leq 1$ мы заключаем, что знак минус не является допустимым, и, следовательно,

$$\begin{aligned} k &= 1 - \lambda + \sqrt{\lambda(\lambda-1)(1-2p)}, \\ c_1 &= 1 - \sqrt{\frac{\lambda(1-2p)}{\lambda-1}}, \\ c_2 &= 1 - \sqrt{\lambda^{-1}(\lambda-1)(1-2p)}. \end{aligned} \tag{7.12}$$

Легко проверить, что $0 \leq c_1 \leq c_2 \leq 1$ и что распределение $F$, определенное посредством (7.11) с этими значениями $k_1$, $c_1$ и $c_2$, имеет вид

$$F_0(x) = \begin{cases} x, & F_0(x) - \lambda x > k, \\ 1, & F_0(x) - \lambda x < k, \\ \lambda x + k, & F_0(x) - \lambda x = k. \end{cases}$$

Поэтому на основе замечания 5.3 мы заключаем, что $F_0$ является решением.

Если $\lambda < 1$ и $\lambda/2 \leq p \leq 1/2$, то рассмотрим решение в виде

$$F(x) = \begin{cases} \lambda x + k, & 0 \leq x \leq c_1, \\ x, & c_1 \leq x \leq 1. \end{cases} \tag{7.13}$$

В этом случае находим

$$k = \sqrt{(1-2p)(1-\lambda)}, \quad c_1 = \sqrt{\frac{1-2p}{1-\lambda}}, \tag{7.14}$$

и можно проверить, что распределение $F_0$, определенное из (7.13) с $k$ и $c_1$, заданными в (7.14), есть решение.

Суммируя изложенное, мы имеем: если $\lambda > 1$ и $(2\lambda)^{-1} \leq p \leq 1/2$, то решение $F_0$ имеет вид (7.11) и (7.12); если $\lambda < 1$ и $2^{-1}\lambda \leq p \leq 1/2$, то решение выражается (7.13) и (7.14). Во всех остальных случаях решение либо не существует, либо совпадает с тем, что дано в части А.

# БИБЛИОГРАФИЯ


Алберг, Нильсон (Ahlberg J. H., Nilson E. N.) [1963]. Convergence properties of the spline fit, J. Soc. Ind. Appl. Math. **11**, 95—104.

Алберг, Нильсон, Уолш (Ahlberg J. H., Nilson E. N., Walsh J. L.) [1964]. Fundamental properties of generalized splines, Proc. Natl. Acad. Sci. **52**, 1412—1419.

Алберг, Нильсон, Уолш (Ahlberg J. H., Nilson E. N., Walsh J. L.) [1965]. Best approximation and convergence properties of highervorder spline approximations, J. Math. Mech. **14**, 231—243.

Апостол (Apostol T. M.) [1957]. Mathematical analysis, Addison-Wesley, Reading, Massachusetts.

Ахиезер Н. И. [1931]. Über ein Tschebyscheffsches Extremumproblem, Math. Ann. **104**, 739—744.

Ахиезер Н. И. [1932]. Об асимптотическом значении наилучшего приближения некоторых рациональных функций посредством полиномов, Сообщ. Харьк. матем. о-ва.

Ахиезер Н. И., Крейн М. Г. [1938]. О некоторых вопросах теории моментов, Харьков, ГОНТИ.

Ахиезер Н. И. [1961]. Классическая проблема моментов и некоторые вопросы анализа, связанные с нею, М., Физматгиз.

Ахиезер Н. И. [1965]. Лекции по теории аппроксимации, М., «Наука».

Бак (Buck C. R.) [1959]. Linear spaces and approximation theory, Ersch. in: On numerical approximation, Proc. Sympos. Madison, 11—23.

Варлоу, Маршалл (Barlow R. E., Marshall A. W.) [1964]. Bounds for distributions with monotone hazard rate, I and II, Ann. Math. Stat. **35**, 1234—1274.

Беккенбах (Beckenbach E. F.) [1937]. Generalized convex functions, Bull. Am. Math. Soc. **43**, 363—371.

Веккенбах, Беллман (Beckenbach E. F., Bellman R.) [1961]. Inequalities, Ergebnisse der Math. und ihrer Grenzgebiete, Neue Folge **30**, Springer-Verlag, Berlin. (Русский перевод: Неравенства, М., «Мир», 1965.)

Беллман (Bellman R.) [1953]. On an inequality due to Weinberger, Am. Math. Monthly **60**, 402.

Веллман, Гликсберг, Гросс (Bellman R., Glicksberg I., Gross O.) [1954]. Some variational problems occurring in the theory of dynamic programming, Rendiconti del Circolo Mathematico di Polermo (2), **3**, 1—35.

Беллман, Гликсберг, Гросс (Bellman R., Glicksberg I., Gross O.) [1958]. Some aspects of the mathematical theory of Control process. The Rand Corporation, Report R-313, Santa Monica, California. (Русский перевод: Некоторые вопросы математической теории процессов регулирования, М., ИЛ, 1962.)

Беннет (Bennet G.) [1962]. Probability inequalities for the sum of independent random variables, J. Am. Stat. Assoc. **57**, 33—45.

Б е р в а л ь д (Berwald L.) [1947]. Verallgemeinerung eines Mittelwartsatzes von J. Favard für Positive Konkave Funktionen, Acta Math. **79**, 17—37.
Б е р г е (Berge P. O.) [1937]. A note on a form of Tchebysheff's Theorem for two variables, Biometrika **29**, 405—406.
Б е р н ш т е й н С. Н. [1912]. О наилучшем приближении непрерывных функций посредством многочленов данной степени, Сообщ. Харьк. матем. о-ва, сер. 2, т. 13, 49—194.
Б е р н ш т е й н С. Н. [1923]. Sur une propriéte des functions entiéres, Compt. Rend. **176**, 1603—1605.
Б е р н ш т е й н С. Н. [1924]. Об одном видоизменении неравенства Чебышева и о погрешности формулы Лапласа, Харьк. ученые зап. НИ кафедр Украины, Отд. матем., вып. 1, 38—49.
Б е р н ш т е й н С. Н. [1926]. Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réele, Gauthier-Villars, Paris. (См. примечание в конце библиографии.)
Б е р н ш т е й н С. Н. [1927]. Sur une propriété des polynômes de Tchebycheff, ДАН СССР (А), 405—407.
Б е р н ш т е й н С. Н. [1928]. Sur les fonctions absolument monotones, Acta Math. **52**, 1—66.
Б е р н ш т е й н С. Н. [1937a]. Sur un théoréme de M. Szegö, Prace Matem. Fyzyezne **44**, 173—177.
Б е р н ш т е й н С. Н. [1937 b]. О некоторых видоизменениях неравенства Чебышева, ДАН СССР **17**, 275—277.
Б е р н ш т е й н С. Н. [1938]. О наилучшем приближении $|x|^p$ при помощи многочленов весьма высокой степени, Изв. АН СССР, сер. матем., № 2, 169—190.
Б е р н ш т е й н С. Н. [1950]. Избранные труды, I, М., Изд-во АН СССР.
Б и р н б а у м, Р а й м о н д, Ц у к е р м а н (Birnbaum Z. W., Raymond J., Zuckerman H. S.) [1947]. A generalization of Tshebyshev's inequality to two dimensions, Ann. Math. Stat. **18**, 70—79.
Б и р н б а у м, К л о с е (Birnbaum Z. W., Klose O. M.) [1955]. Bounds for the variance of the $U$-statistic in terms of probability $Y < X$, Abstract, Bull. Am. Math. Soc.
Б и р н б а у м, М а р ш а л л (Birnbaum Z. W., Marshall A. W.) [1961]. Some multivariate Tchebyshev inequalities with extensions to continuous parameter process, Ann. Math. Stat. **32**, 687—703.
Б л е к у э л л (Blackwell D.) [1950]. Comparison of experiments, Proceedings Second Berkeley symposium of mathematical statistics and probability, University of California Press, Berkeley, California, 1951, 93—102.
Б л е к у э л л (Blackwell D.) [1951]. The range of certain vector integrals, Proc. Am. Math. Soc. **2**, 390—395.
Б о а с (Boas R. P.) [1954]. Entire functions, Academic Press, New York.
Б о н с а л л (Bonsall F. F.) [1950]. The characterization of generalized convex functions, Quart J. Math. **1**, 100—111.
Б р и к м а н (Brickman L.) [1960]. A new generalization of a problem of F. Lukacs, Compositio Mathematica **14**, Fasc. 3, 195—227.
Б р у н к, И в и н г, Ю т ц (Brunk H. D., Ewing G. M., Utz W. R.) [1954]. Minimizing integrals in certain classes of monotone functions with applications, University of Missouri, Department of Math.
Б р у н к (Brunk H. D.) [1956]. On an inequality for convex functions, Proc. Am. Math. Soc. **7**, 817—824.
Б р у н к (Brunk H. D.) [1964]. Integral inequalities for functions with nondecreasing increments, Pacific J. Math. **14**, 783—793.
д е Б у р (de Boor C.) [1963]. Best approximation properties of spline functions of odd degree, J. Math. Mech. **12**, 747—749.
д е Б у р, Л и н ч (de Boor C., Lynch R. E.) [1966]. On splines and their minimum properties, J. Math. Mech. **15**, 953—969.
В а л л е - П у с с е н (de la Vallée Poussin) [1911a]. Sur la Méthode de l'approximation minimum, Soc. Sci. Bruxelles, Annales, Second Partie, Mémoires **35**, 1—16.

Валле-Пуссен (de la Vallée Poussin) [1911b]. Sur les Polynômies d'approximation à une Variable Complexe, Bull. Acad. Roy. Belgique, Cl. Sci., 199—211.
Валле-Пуссен (de la Vallée Poussin) [1919]. Leçons sur l'approximation des fonctions d'une Variable Réelle, Gauthier-Villars, Paris.
Виддер (Widder D. V.) [1941]. The Laplace transform, Princeton University Press, Princeton, New Jersey.
Вейнбергер (Weinberger H. F.) [1961]. Optimal approximation for functions prescribed at equally spaced points, J. Res. Nat. Bur. Stand.— B, Math. and Math. Phys. 65B, 99—104.
Гантмахер Ф. Р., Крейн М. Г. [1950]. Осцилляционные матрицы и ядра и малые колебания механических систем, М., Гостехиздат.
Гамбел (Gumbel E. J.) [1954]. The maxima of the mean largrest walue and the range, Ann. Math. Stat. 25, 76—84.
Герглотц (Herglotz G.) [1911]. Über Potenrgreihen mit Positivem, reelem Teil im Einheitskreis, Ber. Verk. Sächs, Ges. Wiss. Leipzig Math. Phys. Kl. 63, 501—511.
Геронимус Я. Л. [1958]. Многочлены, ортогональные на окружности и на отрезке. Оценки, асимптотические формулы, ортогональные ряды. М., Физматгиз.
Гест (Guest P. G.) [1958]. The spacing of observations in polynomial regression, Ann. Math. Stat. 29, 294—299.
Глассер (Glasser G. J.) [1961]. Tchebycheff-type inequalities in terms of the mean deviation, Sankhya Ser. A, 23, 397—400.
Голомб, Вейнбергер (Golomb M., Weinberger H. F.) [1959]. Optimal approximation and error bounds, Proc. Symp. Num. Approx., R. E. Sanger, ed., pp. 117—190, Univ. Wisc. Press, Madison Wisc.
Гревиль (Greville T. N. E.) [1964a]. Numerical procedures for interpolation by spline functions, J. SIAM, Numer. Anal. Ser. B, 1, 53—68.
Гревиль (Greville T. N. E.) [1964b]. Interpolation by generalized spline functions, MRC Tech. Summary Rep. № 476, Univ. of Wisc., Madison, Wisconsin.
Григорьева И. А. [1962]. Об одном обобщении теоремы С. Н. Бернштейна о форме неотрицательных тригонометрических полиномов на случай любого количества соотношений между коэффициентами полиномов, ДАН СССР 147, № 2, 283—286.
Гудвин (Godwin H. J.) [1955]. On generalizations of Tchebycheff's inequality, Am. Stat. Assoc. J. 50, 923—945.
Гудвин (Godwin H. J.) [1964]. Inequalities on distribution functions, Griffin's Statistical Monographs and Courses, Hafner, New York.
Гутман (Guttman L.) [1948]. An inequality for kurtosis, Ann. Math. Stat. 19, 277—278.
Данскин (Danskin J.) [1955]. Mathematical treatment of a stock-piling problem, Naval Res. Logistics Quart. 2, 99—109.
Данциг, Вальд (Dantzig G. B., Wald A.) [1951]. On the fundamental lemma of Neyman and Pearson, Ann. Math. Stat. 22, 87—93.
Дворецкий, Вальд, Вольфовиц (Dvoretsky A., Wald A., Wolfowitz J.) [1951]. Relations among certain ranges of vector measures, Pacific J. Math. 1, 59—74.
Джексон (Jackson D.) [1911]. Uber die Genanigkeit der Annäperung Stetiger Funktionen durch ganze rationale Funktionen Gegebenen Grades und Trigonometrische Summen Gegebener Ordnung, Dissertation, Göttingen.
Джексон (Jackson D.) [1924]. A general class of problems in approximation, Am. J. Math. 46, 215—234.
Джексон (Jackson D.) [1930]. The theory of approximation, Am. Math. Soc. Coll. Publ., XI
Джонсон (Johnson R. S.) [1960]. On monosplines of least deviation, Trans. Am. Math. Soc. 96, 458—477.
Дочев К. [1962]. Об одной теореме С. Н. Бернштейна, ДАН СССР 146, № 1, 17—19.

Д р е ш е р (Dresher M.) [1953]. Moment spaces and inequalities, Duke Math. J. 20, 261—271.
Д у б и н с, С п а н ь е (Dubins L. E., Spanier E. H.) [1961]. How to cut a cake fairly, Am. Math. Monthly 68, 1—17.
Д у б и н с, Ф р и д м а н (Dubins L. E., Freedman D. A.), [1965]. A sharper form of the Borel—Cantelli lemma and the strong law, Ann. Math. Stat. 36, 800—807.
Д у б и н с, С е в и д ж (Dubins L. E., Savage L. J.) [1965a]. A Tshebycheff-like inequality for stochastic process, Natl. Acad. Sci. 53, 274—275.
Д у б и н с, С е в и д ж (Dubins, L. E., Savage L. J.) [1965b]. How to gamble if you must, Mc Craw-Hill, New York.
Е г е р в а р и, Т у р а н (Egerváry E., Turán P.) [1958]. Notes on interpolation V, Acta Math. Acad. Sci. Hung. IX, 259—267.
Е г е р в а р и, Т у р а н (Egerváry E., Turán P.) [1959]. Notes on interpolation VI, Acta Math. Acad. Sci. Hung. X, 55—62.
З и г м у н д (Zygmund A.) [1959]. Trigonometric series, vol. 2, Cambridge University Press, Cambridge, England. (Р у с с к и й п е р е в о д: Тригонометрические ряды, т. 2, М., «Мир», 1965.)
З о л о т а р е в Е. И. [1932]. Приложение эллиптических функций к вопросам о функциях, наименее и наиболее отклоняющихся от нуля, Собр. соч., т. II.
И с и и (Isii K.) [1957]. Some investigations of the relation between distribution functions and their moments, Ann. Inst. Stat. Math. 9, 1—11.
И с и и (Isii K.) [1959a]. On Tchebycheff type inequalities, Proc. Inst. Stat. Math. 7, 123—143.
И с и и (Issi K.) [1959b]. On a method for generalizations of Tchebycheff's inequality, Ann. Inst. Stat. Math. 10, 65—88.
И с и и (Issi K.) [1959c]. Bounds on probability for non-negative random variables, Ann. Inst. Stat. 11, 89—99.
И с и и (Issi K.) [1960]. The extrema of probability determined by generalized moments (I) bunded random variables, Ann. Inst. Stat. Math. 12, 119—133.
И с и и (Issi K.) [1963]. On sharpness of Tchebycheff-type inequalities, Ann. Inst. Stat. Math. 14, 185—197.
К а н и э л ь (Kaniel S.) [1966]. Estimates for some computational techniques in linear algebra, Math. Comput. 20, 369—378.
К а р а м а т а (Karamata J.) [1932]. Sur une inégalité relative aux fonctions convexes, Publ. Math. Univ. Belgrade 1, 145—148.
К а р а т е о д о р и (Carathéodory C.) [1907). Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die Gegebene Werte Nicht Annehmen, Math. Ann. 64, 95—115.
К а р а т е о д о р и (Carathéodory C.) [1911]. Über den Variabilitätsbereich der Fourierschen Konstanten von Positiven Harmonischen Funktionen, Rend. Circ. Mat. Palermo 32, 193—217.
К а р л и н (Karlin S.) [1953a]. Extreme points of vector functions, Proc. Am. Math. Soc. 4, 603—610.
К а р л и н (Karlin S.) [1953b]. The theory of infinite games, Ann. Math. 58, 371—401.
К а р л и н, Ш е п л и (Karlin S., Shapley L. S.) [1953] Geometry of moment spaces, Mem. Am. Math. Soc., № 12.
К а р л и н (Karlin S.) [1956]. Decision theory for Pólya distributions, case of two actions, I, Proceedings Third Berkeley Sumposium on Probability and Statistics, vol. 1, University of California Press, Berkeley, California, pp. 115—129.
К а р л и н, Р у б и н (Karlin S., Rubin H.) [1956]. The theory of decision procedures for distributions with monotone likelihood ratio, Ann. Math. Stat. 27, 272—299.
К а р л и н (Karlin S.) [1957]. Pólya type distributions II, Ann. Math. Stat. 28, 281—308.
К а р л и н (Karlin S.) [1958]. Pólya type distributions IV, Some principles of selecting a single procedure from a complete class, Ann. Math. Stat. 29, 1—21.

К а р л и н (Karlin S.) [1959]. Mathematical methods and theory in games, programming and economics, vol. 1, 2, Addison-Wesley Publishing Co., Inc., Reading, Mass. (Р у с с к и й п е р е в о д: Математические методы в теории игр, программировании и экономике, М., «Мир», 1964.)
К а р л и н, П р о с ч е н, Б а р л о у (Karlin S., Proschan F., Barlow R. E.) [1961]. Moment inequalities of Pólya frequency functions, Pacific J. Math. II, 1023—1033.
К а р л и н (Karlin S.) [1963]. Representation theorems for positive functions, J. Math. Mech. 12, 599—618.
К а р л и н, Н о в и к о в (Karlin S., Novikoff A.) [1963]. Generalized convex inequalities, Pacific J. Math. 13, 1251—1279.
К а р л и н, П р у и т т, М а д о у (Karlin S., Pruitt W. E., Madow W. G.) [1963]. On choosing combinations of Weapons, Naval Res. Logistics Quart. 10, № 2, 95—119.
К а р л и н (Karlin S.) [1964a]. Generalized Bernstein inequalities, Tech. Report № 41, Contract Nonr-225 (28), Stanford University, Stanford, California.
К а р л и н (Karlin S.) [1964b]. Total positivity, absorption probabilities and applications, Trans. Am. Math. Soc. 111, 33—107.
К а р л и н (Karlin S.) [1964c]. The existence of Eigenvalues for integral operators, Trans. Am. Math. Soc. 113, 1—17.
К а р л и н, С т а д д е н (Karlin S., Studden W. J.) [1966]. Optimal experimental desings, Ann. Math. Stat. 37, № 4, 783—815.
К а р л и н, Ц и г л е р (Karlin S., Zeigler Z.) [1967]. Tchebycheffian spline functions, Inegualities, Proc. Sympos. Wright—Patterson Air Forse Base, Ohio, 1965, Academic Press, New York, 137—149.
К а р л и н, Ш у м е й к е р (Karlin S., Shumaker L.) [1967a]. Characterization of moment points in terms of Christoffel numbers, J. d'Anal. Math. 20, 213—231.
К а р л и н, Ш у м е й к е р (Karlin S., Schumaker L.) [1967b]. The fundamental theorem of algebra for Tchebycheffian splines, J. d'Anal. Math. 20, 233—270.
К а р л и н (Karlin S.) [1968]. Total positivity and applications, Stanford, California, Stanford Unir. press.
К е м п е р м а н (Kemperman J. H. B.) [1965]. On the sharpness of Tchebycheff type inequalities, Indag. Math. 27, 554—601.
К и н г м а н (Kingman J. F. C.) [1963]. On inequalitics of the Tshebyshev type, Proc. Cembridge Phil. Soc. 59, 135—146.
К и ф е р (Kiefer J.) [1959]. Optimum experimental designs, JRSS (Ser. B) 21, 272—319.
К и ф е р, В о л ь ф о в и ц (Kiefer J., Wolfowitz J.) [1959]. Optimum designs in regression problems, Ann. Math. Stat. 30, 271—294.
К и ф е р (Kiefer J.) [1960]. Optimal experimental designs V, with applications to systematic and rotatable designs, Proceedings Fourth Berkeley Symposium, vol. I, University of California Press, Berkeley, California, pp. 381—405.
К и ф е р, В о л ь ф о в и ц (Kiefer J., Wolfowitz J.) [1960]. The equivalence of two extremum problems, Can. J. Math. 12, 363—366.
К и ф е р (Kiefer. J.) [1961]. Optimum desings in regression problems, II, Ann. Math. Stat. 32, 298—325.
К и ф е р (Kiefer J.) [1962a]. An extremum result, Can. J. Math. 14, 597—601.
К и ф е р (Kiefer J.) [1962b]. Two more criteria equivalent to $D$-optimality of designs, Ann. Math. Stat. 33, 792—796.
К и ф е р, В о л ь ф о в и ц (Kiefer J., Wolfowitz J.) [1965]. On a theorem of Hoel and Levine on extrapolation desings, Ann. Math. Stat. 36, 1627—1665.
К о у п м а н с (Koopmans B. O.) [1956]. The theory of search, I. Kinematics bases, Operations Res. 4, 324—346.
К р е й н М. Г. [1934]. Об одном обобщении исследований акад Маркова о предельных величинах интегралов, Труды II Всесоюзного матем. съезда, т. II, 152—154.
К р е й н М. Г. [1951]. Идеи П. Л. Чебышева и А. А. Маркова в теории предельных величин интегралов и их дальнейшее развитие, УФН VI, вып. 4, 3—120.

Крейн М. Г., Рехтман П. Г. [1955]. Развитие теории Чебышева — Маркова предельных величин интегралов в одном новом направлении, УФН X, вып. 1, 67—78.
Лал (Lal D. N.) [1955]. A note on a form of Tchebycheff's inequality for two or more variables, Sankhya, **15**, 317—320.
Левин В. И., Стечкин С. Б. [1948]. Неравенства, Дополнение к переводу книги Харди, Литтлвуд, Полиа, Неравенства, М., ИЛ.
Левитан Б. Об одном обобщении неравенств С. Н. Бернштейна и Н. Boh'a, ДАН **15**, № 4, 169—172.
Лоев (Loeve M.) [1963]. Probability theory, third ed., Van Nostrand, New York.
Ляпунов А. А. [1940]. О вполне аддитивных вектор-функциях, Изв. АН СССР **4**, № 6, 465—478.
Маданский (Madansky A.) [1959]. Bounds on the expectation of a convex function of a multivariate random variable, Ann. Math. Stat. **30**, 743—746.
Малголланд, Роджерс (Mulholland H. P., Rogers C. A.) [1958]. Representation theorems fordistribution functions, Proc. London Math. Soc. (3) **8**, 177—223.
Маллоу (Mallows C. L.) [1956]. Generalizations of Tchebycheff's inequalities, J. Roy. Stat. Soc., Ser. B, **18**, 139—176.
Маллоу (Mallows C. L.) [1963]. A generalization of Chebyshev inequalities, Proc. London Math. Soc., Third. Ser., **13**, 385—412.
Марков А. А. [1886]. Sur les racines de certaines equations (second note), A math. Ann. **27**, 177—182.
Марков А. А. [1889]. Об одном вопросе Д. И. Менделеева, С.-Пб., ДАН, **62**, 1—24.
Марков А. А. [1898]. О предельных величинах интегралов в связи с интерполированием, Зап. АН, VIII серия, **6**, № 5, 146—230.
Марков В. А. [1916]. Über Polynome, die in einem gegebenem Instervalle möglichst weinig vol Null adwechen, Math. Ann. **77**, 213—258.
Маршалл (Marshall A. W.) [1960]. A one-sided analog of Kolmogoroff's inequality, Ann. Math. Stat. **31**, 483—487.
Маршалл, Олкин (Marshall A. W., Olkin I.) [1960a]. Multivariate Chebyshev inequalities, Ann. Math. Stat. **31**, 1001—1014.
Маршалл, Олкин (Marshall A. W., Olkin I.) [1960b]. A one-sided inequality of the Tshebyshev type, Ann. Math. Stat. **31**, 488—491.
Маршалл, Олкин (Marshall A. W., Olkin I.) [1961]. Game theoretic proof that Chebyshev inequalities are sharp, Pacific J. Math. **11**, 1421—1429.
Мейер (Meyer P. A.) [1964]. Progrés récent dans la théorie des cônes sonvexe a base compacte, Seminaire Brelot—Choquet—Deny **8**, № 1.
Мейнардус (Meinardus G.) [1962]. Über Tschebyscheffsche approximationen, Arch. Rational Mech. Anal. **9**, 329—351.
Мейнардус (Meinardus G.) [1964]. Approximation von funktionen und ihre numerische behandlung, Springer-Verlag, Berlin.
фон Мизес (von Mises R.) [1939]. The limits of distribution function if two expected values are given, Ann. Math. Stat. **10**, 99—104.
Моригьюти (Moriqjuti S.) [1951]. Extremal properties of extreme value distributions, Ann. Math. Stat. **22**, 523—536.
Натансон И. П. [1949]. Конструктивная теория функций, М., Гостехиздат.
Начбин (Nachbin L.) [1965]. Weighted approximation for algebras and modules of continuous functions: real and self-adjoint complex cases, Ann. Math. **81**, 289—302.
Ньюман, Шапиро (Newman D. J., Shapiro H. S.) [1963]. Some theorems on Čebyšev approximation, Duke Math. J. **30**, 673—681.
Олкин, Пратт (Olkin I., Pratt J. W.) [1958]. A multivariate Tchebycheff inequality, Ann. Math. Stat. **29**, 226—234.
Олкин (Olkin I.) [1959]. On inequalities of Szegö and Bellman, Proc. Natl. Acad. Sci. **45**, 230—231.
Пик (Pick G.) [1915]. Über die Beschränkungen Analytischer Funktionen, Welche Durch Vorgegebene Funktionswerte Berwirkt Sind, Math. Ann. **77**, 7—23.

П и к (Pick G.) [1920]. Über Beschränkte Funktionen mit Vorgeschriebenen Wertznordmungen, Ann. Acad. Sci. Fenn. Ser. A, **15**, № **3**.
П и р с о н (Pearson K.) [1919]. On generalized Tchebycheff Theorems in the mathematical theory of statistics, Biometrika **12**, 284—296.
П л э к е т т (Plackett R. L.) [1947]. Limits of the ratio of mean range to standard deviation, Biometrika **34**, 120—122.
П о л и а (Pôlya G.) [1922]. On the mean value theorem corresponding to a given linear homogeneosus differential equation, Trans. Am. Math. Soc. **24**, 312—324.
П о л и а, С е г е (Pôlya G., Szegö G.) [1925]. Aufgaben und Lehrsätze aus der Analysis, Band I, II, Springer, Berlin.
П о н т р я г и н Л. С. и др. [1962]. Математическая теория оптимальных процессов, М., «Наука».
П о п о в и ч и у (Popoviciu T.) [1934]. Sur quelques proporietes des fonctions d'une ou de deux variables reeles, Mathematica **8**, 1—85.
П о п о в и ч и у (Popoviciu T.) [1936]. Notes sur les functions convexes d'ordre supérieur (I), Mathematica **12**, 81—92.
П о п о в и ч и у (Popoviciu T.) [1945]. Les fonctions convexes, Herman and Cie, Paris.
П о п о в и ч и у (Popoviciu T.) [1959]. Sur le reste dans certaines formules linéaires d'approximation de l'analyse, Mathematica **1** (24); **1**, 95—142.
П о п о в и ч и у (Popoviciu T.) [1961]. Sur la conservation de l'allure de convexité d'une fonction par ses polunomes d'interpolation, Mathematica **3** (26); **2**, 311—329.
П о п о в и ч и у (Popoviciu T.) [1962]. Sur la conservation, par la polynôme d'interpolation de L. Fejer, du signe ou de la monotonie de la fonction, Analeles. Stiintifice **8**, 65—84.
П о с с е (Posee C.) [1886]. Sur quelques applications des fractions continues algebriques, St.-Petersbourg, 1886, 1—175.
П о с с е (Posse K. A.) [1885]. К вопросу о предельных значениях интегралов или сумм, Сообщения Харьк. матем. о-ва за 1885 г., 35—58.
П э л и, З и г м у н д (Paley R. E. A. C., Zygmund A.) [1930]. On some series of functions, Proc. Cambridge Phil. Soc. **26**, 337—357.
Р а д а у (Radau R.) [1880]. Etude sur les formules d'approximation qui servent à calculer la valeur numérique d'une intégrale definie, J. Math. Pures Appl., 3rd Series, **6**, 283—336.
Р а д о н (Radon J.) [1935]. Restansdrücke bei Interpolation und Quadratur-formeln durch Bestimmte integrale, Monatsh Math. Phys. **42**, 389—396.
Р а й с (Rice J. R.) [1959]. On the convergence of an algorithm for best Tchebycheff approximations, J. Soc. Ind. Math. **7**, 133—142.
Р а й с (Rice J. R.) [1961]. Tchebycheff approximation by functions unicolvent of variable degree, Trans. Am. Math. Soc. **99**, 298—302.
Р а й с (Rice J. R.) [1963]. Tchebucheff approximation in several variables, Trans Am. Math. Soc. **109**, 444—466.
Р е н ь и (Rènyi A.) [1962]. Wahrscheinlichkeitsrechnung, Mit einen Anhang über informations-theorie, veb Deutscher Verlag der Wissenschaften, Berlin.
Р и в л и н, Ш а п и р о (Rivlin T. J., Shapiro H. S.) [1960]. Someuniqueness problems in approximation theory, Comm. Pure Appl. Math. **13**, 35—47.
Р и в л и н, Ш а п и р о (Rivlin T. J., Shapiro H. S.) [1961]. A unified approach to certain problems of approximation and minimization, J. Soc. Ink. Appl. Math. **9**, 670—699.
Р и в л и н (Rivlin T. J.) [1962]. Polynomials of best uniform approximation to certain rational functions, Numer. Math. **4**, 345—349.
Р и с с (Riesz M.) [1914a]. Eine trigonometrische Interpolations-formel und einige Ungleichungen für Polynome, Jahresber. Deuts. Math. Vereinigung **23**, 354—368.
Р и с с (Riesz M.) [1914b]. Formule d'interpolation pour la Dérivée d'un Polynôme Trigonometrique, Compt. Rend. **158** (1914); Eine trigonometrische Interpolations formel und Einige Ungleichungen für Polynôme, Jahresber. Deuts. Math. Verinigung.

Рихтер (Richter H.) [1957]. Paremeterfreie Abschätzung und Realisierung von Erwartungswerten, Blatter der Deutschen Gecellschaft für versicherungsmathematik (Konrad Triltsch Verlag, Wuneburg) 3, 147—161.
Роуден (Rouden H. L.) [1953]. Bounds on a distribution function when its first $n$ moments are given, Ann. Math. Stat. 24, 361—376.
Роббинс (Robbins H.) [1948]. Some remarks on the inequality of Tchebycheff, in Studies and Essays, presented to R. Courant on his 60th Birthday, Interscience, New York, pp. 345—350.
Рогозинский (Rogosinski W. W.) [1954]. Extremum problems for polynomials and trigonometric polynomials, J. London Math. Soc. 29, 259—275.
Рогозинский (Rogosinski W. W.) [1958]. Moments of non-negative mass, Proc. Roy. Soc. (London), A, 245, 1—27.
Рогозинский (Rogosinski W. W.) [1962]. In Gabor Szegö et al., Eds. Non-negative linear functionals, moment problems and extremum problems in polynomial space, Studies in Mathematical Analysis and Related Topics; Essays in Honor of George Pólya, Contributors; Lars V. Ahlfors, and others, Stanford University Press, Stanford, California.
Розенблюм (Rosenbloom P. C.) [1951]. Quelques classes de problémes extremaux, Bull. Soc. Math. Fr. 79, 1—58; 80, 183—215.
Рустажи (Rustagi J. S.) [1957]. On minimizing and maximizing a cetrain integral with statistical applications, Ann. Math. Stat. 28, 309—328.
Самуэльс (Samuels S. M.) [1965]. On the number of successes in independent trials, Ann. Math. Stat. 36, 1272—1278.
Самуэльс (Samuels S. M.) [1966]. On a Chebyshev-type inequality for sums of independent random variables, Ann. Math. Stat. 3 , 1, 248—259.
Сард (Sard A.) [1949]. Best approximate integration formulas; best approximation formulas, Am. J. Math. 71, 80—91.
Сард (Sard A.) [1963]. Linear approximations, Math. Surveys № 9, Am. Math. Soc., Providence R. I.
Севидж (Savage I. R.) [1961]. Probability inequalities of the Tchebycheff type, J. Research Natl. Bureau Standards-B, Math. Mathematical Phys., 65B, № 3, 211—222. (Русский перевод в сб. Математика, 6 : 4, 1962, 71.)
Сеге (Szegö G.) [1922]. Bemerkungen zu einem Satz von J. H. Grace über die Wurzeln Algebraischer Gleichungen; Math. Z. 13, 28—55.
Сеге (Szegö G.) [1925]. Über einem Satz von A. Markoff, Math. Z. 23, 45—61.
Сеге (Szegö G.) [1928]. Über einem Satz des Herrn Serge Bernstein, Schriften Konigsberg Gel. Ges. 5, 59—70.
Сеге (Szegö G.) [1950]. Über eine Yerallgemeinerung des Dirichletschen Integrals, Math. Z. 52, 676—685.
Сеге (Szegö G.) [1959]. Orthogonal polynomials, revised edition, Am. Math. Soc. Coll. Pub., vol. 23, New York. (Русский перевод:. Ортогональные многочлены, М., Физматгиз. 1962).
Сеге, Туран (Szegö G., Turan P.) [1961]. On the monotone convergence of certain riemann sums, Publ. Math., Debrecen, 8, 326—335.
Сеге (Szegö G.) [1962]. On positive harmonic polynomials, Illinois J. Math. 6, 181—186.
Сеге (Szegö G.) [1964a]. On a problem of the best approximation, Abhandlungen aus dem Mathematischen Seminar der Universitat Hamburg 27, 193—198.
Сеге (Szegö G.) [1964b]. On some problems of approximations, Pub. Math. Inst. Hung. Acad. 9, 3—9.
Сельберг (Selberg H. L.) [1942]. On an inequality in mathematical statistics, Norsk. Mat. Tidsskr. 24, 1—12.
Сельберг (Selberg H. L.) [1944]. Bemerkninger om et multipelt integral, Norsk Mat. Tidsskr. 26, 71—78.
Сизельски (Siesielski Z. [1957]. A note on some inequalities of Jensen's type, Annales Polonici Mathematici 4, 269—274.
Стеффенсен (Steffensen J. F.) [1947]. Bounds on certain trigonometric integrals, Tenth Scandinavian Math. Congr. 181—186 (1946); Copenhagen J. Gjellerups Forlag.

Т и м а н А. Ф. [1963]. Теория приближения функций действительного переменного, М., Физматгиз.
У и т т л (Whittle P.) [1958a]. Continuous generalizations of Tchebichev's inequality; Теория вероятностей и ее применения, т. III (1958), № 4, 386—396.
У и т т л (Whittle P.) [1958b]. A multivariate generalization of Tchebichev's inequality, Quart. J. Math., Oxford, ser. 3, 9, 232—240.
У о л ш, А л б е р г, Н и л с о н (Walsh J. L., Ahlberg J. H., Nislon E. N.) [1962]. Best approximation properties of the spline fit, J. Math. Mech. **11**, 225—234.
У о л ш, А л б е р г, Н и л с о н (Walsh J. L., Ahlberg J. H., Nilson E. N.) [1963]. Best approximation and convergence properties of higherorder spline fits, Am. Math. Soc. Notices **10**; Abstract 63t—103, p. 202.
Ф а в а р (Favard J ) [1933]. Sur les Valeurs Moyennes; Bull. Sci. Math. (2), **57**, 54—64.
Ф а н ь, Л о р е н т ц (Fan K., Lorentz G. G.) [1954]. An integral inequality, Am. Math. Monthly **61**, 626—631.
Ф е д ж е р (Féjer L.) [1932]. Bestimmung derjenigen Adszissen eines Intervalles, für velche die Quadratsumme der Grundfunktionen der Lagrangeschen Interpolation im Intervalle ein möglechst kleines Maximum besitzt, Annali della R. Scuola Normale Superisore de Pisa, Serie II, **1**, 263—276.
Х а а р (Haar A.) [1918]. Die Minkowskische Geometric und die Annäherung an stetige Funktionen, Math. Ann. 78, 294—311.
Х а д ж е к, Р е н ь и (Hajek J., Renyi A.) [1955]. Generalization of an inequality of Kolmogorov, Acta Math. Acad. Sci. Hungar. **6**, 281—283.
Х а л к и н (Halkin H.) [1962]. Liapounov's theorem on the range of a vector measure and Pontryagin's maximum principle, Arch, Rational Mech. Anal. **10**, 296—304.
Х а л м о ш (Halmos P.) [1948]. The range of a vector measure, Bull. Am. Math. Soc. **54**, 416—421.
Х а н т (Hunt G. A.) [1955] An inequality in probability theory, Proc. Am. Math. Soc. **6**, 506—510.
Х а р д и, Л и т т л в у д, П о л и а (Hardy G. H., Littlewood J. E., Polya G.) [1952]. Inequalities, Gambridge University Press, Cambridge, England. (Р у с с к и й п е р е в о д п е р в о г о и з д а н и я: Неравенства, М., ИЛ, 1948.)
Х а р т л и, Д э в и д (Hartley H. O., David H. A.) [1954]. Universal bounds for mean range and extreme observations, Ann. Math. Stat. **25**, 85—99.
Х ё ф д и н г (Hoeffding W.) [1955]. The extrema of the expected value of a function of independent random veriable,, Ann. Math. Stat. **26**, 268—275.
Х ё ф д и н г, Ш р и к х а н д (Hoeffding W., Shrikhande S. S.) [1955]. Bounds for the distribution function of a sum of independent identically distributed random variables, Am. Math. Stat. **26**, 439—449.
Х ё ф д и н г (Hoeffding W.) [1956]. On the distribution of the number of successes in independ trials, Ann. Math. Stat. **27**, 713—721.
Х ё ф д и н г (Hoeffding W.) [1963] Probability inequalities for sums of bounded random variables, Am. Stat. Assoc. 58, 13—29.
Х и н ч и н А. [1923]. Über Dyadische Brüche, Math. Z. **18**, 109—116.
Х и р с м а н, В и д д е р (Hirschman I. I., Widder D. V.) [1955]. The convolution transform, Princeton University Press, Princeton, New Jersey.
Х о п ф (Hopf E.) [1926]. Über die Zusammenhänge Zwischen Gewissen Höheren Differen zenquotienten Reeler Funktionen Eines Reelen Variablen und Deren Differenzierbarkeitseigenshaften, Dissertation Berlin, 30 pp.
Х о р м а н д е р (Hormander L.) [1954]. On a theorem of grace, Math. Scand. **2**, 55—64.
Х о э л (Hoel P. G.) [1958]. Efficiency problems in polynomial estimation, Ann. Math. Stat. **29**, 1134—1146.
Х о э л, Л е в и н (Hoel P. G., Levine A.) [1964]. Optimal spacing and wieghing in polynomial prediction, Ann. Math. Stat. **35**, 1553—1560.
Ц и г л е р (Ziegler Z.) [1966]. Generalized convexity cones, Pacific J. Math. **17**, № 3, 561—580.

Ч е б ы ш е в П. Л. [1853]. Теория механизмов, известных под названием параллелограммов, Полное собр. соч., т. I.
Ч е б ы ш е в П. Л. [1857]. Вопросы о наименьших величинах, связанные с приближенным представлением функций, Собр. соч., т. I.
Ч е б ы ш е в П. Л. [1874]. Sur les valeurs limites des integrales, J. Math. Pures Appl. (2), **19**, 157—160.
Ч е б ы ш е в П. Л. [1948]. Полное собрание сочинений, т. III, Математический анализ, М., Изд-во АН СССР.
Ч е н и, Г о л ь д ш т е й н (Cheney E. W., Goldstein A. A.) [1959]. Newton's method for convex programming and Tshebysheff approximation, Numer. Math. **1**, 253—268.
Ч е н и, Л о е б (Cheney E. W., Loeb H. L.) [1961]. Two new algorithms for rational approximation, Numer. Math. **3**, 72—75.
Ч е н и, Г о л ь д ш т е й н (Cheney E. W., Goldstein A. A.) [1962]. Tchebycheff approximation in locally convex spaces, Bull. Am. Math. Soc. **68**, 449—450.
Ч о к е т (Choquet G.) [1956]. Existense des representations integrales on moyen deposits extremaux dans les cônvex convexes. C. R. Acad. Sci. Paris, **258**, 699—702.
Ш а п и р о (Shapiro H. S.) [1955]. In W. Kaplan et al. Eds., Applications of normal linear spaces to function-theoretic extremal problems. Lectures on functions of a complex variable, University of Michigan Press, Ann. Arbor, Michigan.
Ш а п и р о (Shapiro H. S.) [1961]. On a class of extermal problems for polynomials in the unit circle, Portugal Math. **20**, 67—93.
Ш е ф ф е р (Schaeffer A. C.) [1941]. Inequalities of A. Markoff and S. Bernstein for polynomials and related functions, Bull. Am. Math. Soc. **47**, 565—579.
Ш е н б е р г (Schoenberg I. J.) [1946]. Contributions to the problem of approximation of equidistand data by analytic functions, Part A, Quart. Appl. Math. **4**, 45—99; Part B, ibid. **4**, 112—141.
Ш е н б е р г (Shoenberg I. J.) [1948]. Some analytical aspects of the problem of smoothing, in Studies and Essays presented to R. Courant on his 60th Birthday, Interscience, New York, pp. 152—156.
Ш е н б е р г (Schoenberg I. J.) [1951]. On Pólya frequency functions, I: the totally positive functions and their Laplace transforms, J. d'Anal. Math. **1**, 331—374.
Ш е н б е р г (Schoenberg I. J.) [1953]. On smoothing operations and their generating functions, Bull. Am. Math. Soc. **59**, 199—230.
Ш е н б е р г, У и т н и (Schoenberg I. J., Whitney A.) [1953]. On Polya frequency functions, III: the positivity of translation determinants with an application to the interpolation problem by spline curves, Trans. Am. Math. Soc **74**, 246—259.
Ш е н б е р г (Shoenberg I. J.) [1954a]. On multiple positive sequences and functions, Bull. Am. Math. Soc. **60**, 160.
Ш е н б е р г (Schoenberg I. J.) [1954b]. An isoperimetric inequality for closed curves convex in even-dimensional Euclidean Spaces, Acta Math. **91**, 143—164.
Ш е н б е р г (Schoenberg I. J.) [1958]. Spline functions, convex curves and mechanical quadrature, Bull. Am. Math. Soc. **54**, 352—357.
Ш е н б е р г (Schoenberg I. J.) [1959]. On the maxima of certain Hankel determinants and the zeros of the classical orthogonal polynomials, Indag. Math. **21**. 282—290.
Ш е н б е р г, С е г е (Schoenberg I. J., Szegö G.) [1960]. An extremum problem for polynomials, Compositio Mathematica **14**, Fasc. **3**, 260—268.
Ш е н б е р г (Schoenberg I. J.) [1963]. On interpolation by spline functions and its minimal properties, to appear in the Proceedings of the Conference on Approximation Theory, Oberwolfach, Germany.
Ш е н б е р г (Schoenberg I. J.) [1964a]. On interpolation by spline functions and its minimal properties, International Series Numerical Mathematics, **5**, On Approximation Theory, Birhauser Verlag Basel Stuttgart.
Ш е н б е р г (Schoenberg I. J.) [1964b]. On best approximations of linear operators, Indag. Math. **26**, № 2, 155—163.

Ш е н б е р г (Schoenberg I. J.) [1964c]. Spline interpolation and the highr derivatives, Proc. Natl. Acad. Sc,. USA, **51**, 24—28.
Ш е н б е р г (Schoenberg I. J.) [1964d]. Spline interpolation and best duaqrature formulae, Bull. Am. Math. Soc. **70**, 143—148.
Ш е н б е р г (Schoenberg I. J.) [1964e]. On trigonometric spline interpolation, J. Math. Mech. **13**, 795—826.
Ш е н б е р г (Shoenberg I. J.) [1965]. On monosplines of least deviation and best quadrature formulae, SIAM, Num. Anal., Series B, **2**, 144—170.
Ш о х а т, Т а м а р к и н (Shohat J. A., Tamarkin J. D.) [1943]. The problem of moments, Mathematical surveys, № 1, Am. Math. Soc., New York.
Э л в и н г (Elfving G.) [1952]. Optimum allocation in linear regression theory, Ann. Math. Stat. **23**, 255—262.
Э л в и н г (Elfving G.) [1959]. Design of linear experiments, Probability and Statistics (Cramér Volume), Viley, New York, pp. 58—74.
Э р р о у, К а р л и н (Arrow K., Karlin S.) [1958]. In K. Arrow, S. Karlin, H. Scarf, Eds., Studies in the mathematical theory of inventory and production, Stanford University Press, Stanford, California, Chap. 4.

П р и м е ч а н и е р е д а к т о р а. См. также монографию С. Н. Бернштейна «Экстремальные свойства полиномов и наилучшее приближение непрерывных функций одной вещественной переменной», Л.— М., ОНТИ, 1937, в которой материал, касающийся систем Чебышева, значительно развит и расширен в сравнении с тем, как он дан во французской монографии 1926 года.

## ЛИТЕРАТУРА, ДОБАВЛЕННАЯ ПРИ ПЕРЕВОДЕ

А д а м я н В. М., А р о в Д. З., К р е й н М. Г. [1968]
Бесконечные ганкелевы матрицы и обобщенные задачи Карагеодори — Файера и И. Шура, Функц. анализ и его прил. **2**, 4, 1—17.
А д а м я н В. М., А р о в Д. З., К р е й н М. Г. [1968]
О бесконечных ганкелевых матрицах и обобщенных задачах Каратеодори — Файера и Ф. Рисса, Функц. анализ и его прил. **2**, 81, 1—19.
А д а м я н В. М., А р о в Д. З., К р е й н М. Г. [1971]
Аналитические свойства пар Шмидта, ганкелева оператора и обобщенная задача Шура — Такачи, Матем. сб. **86**, 1, 34—75.
В и д е н с к и й В. С. [1972]
К двум работам С. Н. Бернштейна о декартовом базисе системы функций Чебышева. 1. Уч. записки ЛГПИ **541**, 36—52.
В и д е н с к и й В. С. [1972]
К двум работам С. Н. Бернштейна о декартовом базисе системы функций Чебышева. 2. «Современный анализ и геометрия», Сб. научн. трудов ЛГПИ, 31—44.
В и д е н с к и й В. С. [1974]
Об одном квазианалитическом классе функций, ДАН СССР **219**, 2, 282—285.
В и д е н с к и й В. С. [1974]
О системах Чебышева, которые обладают декартовым базисом, Сб. исследований по теории функций и функциональному анализу. Владимирский пединститут, 10—18.
В и д е н с к и й В. С. [1975]
О $V$-разрешимых семействах многочленов по системе Чебышева, Докл. АН Арм. ССР **60**, 2, 65—68.
Д е в и с Ч. [1967]
Свойства отображений некоторых систем Чебышева, ДАН СССР **175**, 2, 280—283.
З и н г е р И. (Singer I.) [1967]
Cea mai bună aproximare în spatii vectoriale normate prin elemente din subspatii vectoriale, Ed. Acad. Rep. Soc. Romania, Bucuresti.

З и н г е р (Singer I.) [1970]
Best approximation in normed linear spaces by elements of linear subspaces, Springer-Verlag.
И о ф ф е А. Д., Т и х о м и р о в В. М. [1968]
Двойственность выпуклых функций и экстремальные задачи, УМН **23**, 6 (144), 51—116.
И о ф ф е А. Д., Т и х о м и р о в В. М. [1969]
О минимизации интегральных функционалов, Функц. анализ и его прил. **3**, 3, 61—70.
К а р л и н (K a r l i n S.) [1968]
Representation theorems for positive functions, J. Math. Mech. **12**, 559—618.
К а р л и н (Karlin S.) [1968]
Total positivity, Standard University Press.
К и ф е р (Kiefer J.) [1972]
General optimality equivalence theory for approximative designs (abstract), Bull. Inst. Math. Statist. **1**, 258.
К и ф е р (Kiefer J.) [1974]
General equivalence theory for optimum designs (approximate theory), Ann. Statist. **2**, 849—879.
К р е й н М. Г., Н у д е л ь м а н А. А. [1973]
Проблема моментов Маркова и экстремальные задачи, «Наука», М., 1—551.
К р е й н М. Г., Н у д е л ь м а н А. А. [1975]
Аппроксимация функций из $L_2(\omega_1, \omega_2)$ передаточными функциями линейных систем с минимальной энергией, Пробл. передачи информ. **11**, 2, 37—60.
Л о р а н П. [1975]
Аппроксимация и оптимизация, «Мир», М.
Н у д е л ь м а н А. А., [1967]
Канонические решения проблемы моментов на нескольких интервалах, Матем. заметки **1**, 4, 435—442.
Н у д е л ь м а н А. А. [1970]
Геометрическая характеристика специальных систем Чебышева, Матем. исследования, Кишинев, **5**, 1, 160—165.
Н у д е л ь м а н А. А. [1974]
Обобщенная проблема Неванлинны — Пика и частичная теорема, сб. «Динамика систем», вып. 3, Горький 49—57.
Р е м е з Е. Я. [1969]
Основы численных методов чебышевского приближения, «Наукова думка», Киев, 1—624.
Р у б и н ш т е й н Г. Ш. [1970]
Двойственность в математическом программировании и некоторые вопросы выпуклого анализа, УМН XXV, 5, 171—201.
С м и р н о в Г. С. [1970]
Некоторые свойства наилучшего приближения абстрактных функций полиномами специального вида, сб. «Теория приближения функций и ее прил.», Киев, 129—134.
С т е ч к и н С. Б. [1968]
О некоторых экстремальных свойствах положительных тригонометрических полиномов, Матем. заметки **7**, 4, 411—422.
Т и х о м и р о в В. М. [1969]
Наилучшие методы приближения и интегрирования дифференцируемых функций в пространстве C [—1, 1], Матем. сб. 80(10), 290--304.
У и н н (Wynn H. P.) [1970]
The sequential generation of *D*-optimum experimental designs, Ann. Math. Statist. **41**, 1655—1664.
У и н н (Wynn H. P.) [1972]
Results in theory and construction of *D*-optimum experimental designs, J. Roy. Statist. Soc., Ser. B, **34**, 133—147.
У с Г. Ф. [1974]
Усеченная обобщенная степенная систематическая проблема моментов, Укр. матем. ж. **26**, 3, 348—358.

# ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ