

ЭКОНОМИКО-
МАТЕМАТИЧЕСКАЯ
БИБЛИОТЕКА

И. И. ЕРЕМИН
В. Д. МАЗУРОВ

НЕСТАЦИОНАРНЫЕ
ПРОЦЕССЫ
МАТЕМАТИЧЕСКОГО
ПРОГРАММИРОВАНИЯ

И. И. ЕРЕМИН, В. Д. МАЗУРОВ

НЕСТАЦИОНАРНЫЕ
ПРОЦЕССЫ
МАТЕМАТИЧЕСКОГО
ПРОГРАММИРОВАНИЯ



МОСКВА «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
1979

22.18
Е 70
УДК 519.6

Нестационарные процессы математического программирования.
Еремин И. И., Мазуров В. Д. Наука. Главная редакция физико-математической литературы, М., 1979.

Книга посвящена разработке аппарата моделирования и оптимизации сложных эволюционирующих систем средствами итерационных процессов математического программирования, приспособленных к учету нестационарности тех или иных элементов модели, а также учету их плохой формализуемости. Последнее осуществляется с помощью математических процедур распознавания образов.

Рассматриваемые в книге методы могут быть применены, с одной стороны, к решению обычных задач математического программирования, а с другой — к решению плохо определенных задач оптимизации. Нестационарные процессы математического программирования могут выступать как средство моделирования и оптимизации конкретных развивающихся систем: экономических, технических, биологических и др.

Иван Иванович Еремин, Владимир Данилович Мазуров

**НЕСТАЦИОНАРНЫЕ ПРОЦЕССЫ
МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ**

(Серия: «Экономико-математическая библиотека»)

М., 1979 г., 288 стр. с илл.

Редакторы М. Н. Мушников, Е. Ю. Ходан.
Техн. редактор С. Я. Шкляр. Корректор М. Л. Медведская.

ИБ № 11298

Сдано в набор 20.02.79. Подписано к печати 22.06.79. Т-13104. Бумага 84×108^{1/32}, тип. № 1. Литературная гарнитура. Высокая печать. Условн. печ. л. 15,12. Уч.-изд. л. 15,07. Тираж 6500 экз. Заказ № 3847. Цена книги 1 р. 20 к.

Издательство «Наука»

Главная редакция физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

Ордена Октябрьской Революции и ордена Трудового Красного Знамени
Первая Образцовая типография имени А. А. Жданова Союзполиграфпрома при
Государственном комитете СССР по делам издательств, полиграфии и книжной
торговли. Москва, М-54, Валовая, 28.

Отпечатано во 2-й тип. изд-ва «Наука», Москва, Шубинский пер., 10. Зак. 2118.

Е $\frac{20204-111}{053(02)-79}$ 78-79. 1702070000

© Главная редакция
физико-математической
литературы издательства «Наука»,
1979.

Введение	13
Глава I. Нестационарные системы и методы их моделирования	13
§ 1. Базовые модели	14
§ 2. Математическое программирование, его возможности и границы применимости	20
§ 3. Нестационарные процессы математического программирования	26
§ 4. Процесс обмена информацией в некоторых моделях иерархии управления	36
Глава II. Итерационные процессы фейеровского типа	46
§ 1. Фейеровские отображения и порождаемые ими последовательности	47
§ 2. Применение к решению систем выпуклых неравенств	55
§ 3. Декомпозиционный метод построения фейеровских отображений	64
§ 4. Анализ скорости сходимости	68
§ 5. Рандомизация фейеровских процессов	87
§ 6. Фейеровские процессы для задач выпуклого программирования с негладкими ограничениями	92
§ 7. Несобственные задачи линейного программирования	95
Глава III. Фейеровские методы для нестационарных задач выпуклого программирования с негладкими ограничениями	101
§ 1. Фейеровские отображения, зависящие от параметра, и порождаемые ими процессы	101
§ 2. Применение общих теорем о сходимости к нестационарным системам выпуклых неравенств	106
§ 3. Сходимость основного стандартного процесса	115
§ 4. Применения к нестационарным задачам выпуклого программирования: общие замечания	125

Глава IV. Математические методы распознавания образов	127
§ 1. Математические модели распознавания образов . . .	129
§ 2. Методы аффинного дискриминантного анализа . . .	137
§ 3. Комитетные методы распознавания образов	151
§ 4. Комитетное решение задачи многокритериальной оптимизации	171
Глава V. Нестационарные процессы математического программирования, включающие алгоритмы распознавания образов	182
§ 1. Синтез нестационарных процессов, включающих распознавание образов	182
§ 2. Учет плохо определенных ограничений методами дискриминантного анализа и таксономии	190
§ 3. Автоматизация управления параметрами итерационного процесса для задач математического программирования	213
Глава VI. Приложения в задачах планирования и управления технико-экономическими и природными системами	218
§ 1. Техничко-экономические приложения	218
§ 2. Приложения в биологии и медицине	230
§ 3. Программное обеспечение решения задач производственного планирования и распознавания образов . . .	233
§ 4. Задача объемно-календарного планирования машиностроительного производства	239
Обозначения	278
Библиографические замечания	279
Литература	281
Предметный указатель	288

Предлагаемая книга посвящена итерационным методам математического программирования, вбирающим в себя одновременный процесс идентификации модели по ее отдельным составляющим: исходным данным, ограничениям, описывающим связи между элементами моделируемой системы, целевой функции. Такие методы допускают широкий диапазон интерпретаций и, с одной стороны, расширяют вычислительные возможности в математическом программировании, а с другой — могут служить средством моделирования и оптимизации эволюционирующих систем, адекватное математическое описание которых затруднено или невозможно. Это отвечает, как правило, ситуации, когда точного аналитического выражения тех или иных ограничений (или целевой функции) получить не удастся, однако имеется косвенная информация, характеризующая эти ограничения (целевую функцию). Например, может быть предъявлено множество точек пространства состояний системы, удовлетворяющих конкретному ограничению, и множество точек, ему не удовлетворяющих. В качестве аппарата приближенного моделирования такого ограничения здесь естественным образом выступают математические методы классификации. Синтез итерационных методов оптимизации, порождаемых итерационными операторами (отображениями) с нестационарной системой параметров, и алгоритмов доопределения плохо формализуемых элементов модели (или соответствующего ей оператора) приводит к процессам, названным *нестационарными*. В этом случае структура итерационного оператора зависит от структуры моделируемого объекта и изменяется в соответствии с эволюцией объекта или наших знаний о нем.

Ряд узловых соображений книги навеян наблюдениями за особенностями решения прикладных задач (оптимального производственного планирования, медицинской и

технической диагностики, прогнозирования и др.), плохо поддающихся адекватному моделированию из-за сложности объекта (процесса), недостаточности или неточности сведений о нем, неудовлетворительной его структуризованности, *плохой формализуемости* (что следует отличать от трудной формализуемости) его некоторых связей, наконец, из-за подвижности, быстрой и не всегда предсказуемой изменчивости. В этой ситуации просчет варианта по «застывшему», статическому состоянию системы дает мало практической информации либо вообще таковой не дает. Решение подобных задач превращается в довольно сложный вычислительный процесс итерационного типа, в определенном смысле адекватный динамике обчисляемого объекта. При этом непрерывно учитывается новая информация об объекте. Источником этой информации может быть исследовательский анализ возникающих в системе ситуаций либо информация может поступать извне, либо то и другое. Следовательно, вычисления ведутся в режиме нестационарного процесса, т. е. итерационного процесса для нестационарной модели.

При синтезе нестационарных процессов мы отталкиваемся от понятия итерационного метода, порождаемого итерационным оператором $\varphi \in \{\mathbf{R}^n \rightarrow \mathbf{R}^n\}$ и характеризующегося такими свойствами, как вычислительная простота итерации (относительная), помехоустойчивость, независимость логической структуры оператора φ от конкретных реализаций исходных данных и некоторые другие.

В книге подробно изучается класс *фейеровских* отображений. Этот класс достаточно богат и конструктивен. Способы построения фейеровских отображений для задач выпуклого программирования хорошо приспособлены к условиям негладкости их определяющих функций. В этом смысле итерационные процедуры, основанные на использовании фейеровских отображений, выступают в качестве методов недифференцируемой оптимизации, правда, в рамках предположений выпуклости.

Перечисленные выше свойства итерационных методов определяют возможность введения в итерационный оператор φ той или иной системы параметров, характеризующих объект и среду, и организации процесса

$$\{x_{t+1} = \varphi[\Omega_t](x_t)\}_{\forall t},$$

где $\Omega_t = [y_t, \omega_t, \dots]$ — названная система параметров (векторных, предикатных) со значениями на момент t реализации процесса *).

В этом случае итерационный шаг с точки зрения объема вычислений может усложняться, при этом иногда существенно, за счет использования алгоритмов, доопределяющих и уточняющих значения параметров итерационного оператора и его структурных частей.

Важными с точки зрения основной идеи книги процедурами доопределения являются алгоритмы *распознавания образов*. Коснемся работы этих алгоритмов. Пусть в задаче математического программирования

$$\sup \{f(x) : g(x) \leq 0, x \in M\}, \quad (1)$$

в которой все ее элементы: целевая функция $f \in \{\mathbb{R}^n \rightarrow \mathbb{R}\}$, вектор левых частей системы ограничений $g \in \{\mathbb{R}^n \rightarrow \mathbb{R}^m\}$, условие в форме включения $x \in M$ — плохо формализуемы. Это означает, что конструктивного задания элементов f , g и M нет, однако для ограничений $g(x) \leq 0$ и $x \in M$ заданы (как результат некоторых наблюдений, экспериментов, экспертных оценок и т. д.) соответственно пары множеств G' и G'' , M' и M'' такие, что $g(x) \leq 0$ ($\forall x \in G'$), $g(y) > 0$ ($\forall y \in G''$), $M' \subset M$, $M'' \subset \mathbb{R}^n \setminus M$. Что касается функции f , то информация о ней может иметь, например, следующее представление. Имеется множество $Q \subset \mathbb{R}^n$ (обычно оно является конечным) такое, что для всякого его элемента \bar{x} известны либо значение $f(\bar{x})$, либо некоторые характеристики этого значения, как то: $\text{sgn } f(\bar{x})$, принадлежность числа $f(\bar{x})$ тому или иному интервалу и т. д., т. е. вообще отображение $x \rightarrow \Delta(x) \subset \mathbb{R}$ для всех $x \in Q$. Если Q конечно, то конечно и множество всех $\Delta(x)$ для $x \in Q$; пусть $\Delta(Q) = \{\Delta_1, \dots, \Delta_k\}$ и $\Delta_i = \Delta(x_i)$ ($\forall i \in \mathbb{N}_k$).

Доопределение элементов модели (1) состоит в решении следующих задач (*дискриминантного анализа*). Найти:

1) отображение \bar{g} , разделяющее множества G' и G'' , т. е. решение системы

$$g(x) \leq 0 \quad (\forall x \in G'), \quad g(y) > 0 \quad (\forall y \in G''), \quad g \in \mathcal{G};$$

* Здесь символ $\forall t$ в индексе означает, что последовательность $\{x_t\}$ построена по соответствующей рекурсивной формуле.

2) функцию $\tilde{d}(\mathbf{x})$, разделяющую множества M' и M'' , т. е. решение системы

$$d(\mathbf{x}) \leq 0 \quad (\forall \mathbf{x} \in M'), \quad d(\mathbf{y}) > 0 \quad (\forall \mathbf{y} \in M''), \quad d \in \mathcal{D}; \quad (2)$$

3) разделяющие функции $\tilde{f}_1, \dots, \tilde{f}_{k-1}$ такие, что j -я функция удовлетворяет соотношениям

$$f_j(\Delta_j) \subset R_+, \quad f_j\left(\bigcup_{i>j} \Delta_i\right) \subset R_-, \quad f_j \in \mathcal{F} \\ (\forall j \in \mathbf{N}_{k-1}).$$

Положим $\tilde{M} = \{\tilde{d}(\mathbf{x}) \leq 0\}$,

$$\tilde{f}(\mathbf{x}) = \begin{cases} f(\mathbf{x}_1), & \text{если } \tilde{f}_1(\mathbf{x}) > 0, & \text{иначе} \\ f(\mathbf{x}_2), & \text{если } \tilde{f}_2(\mathbf{x}) > 0, & \text{иначе} \\ \dots & \dots & \dots \\ f(\mathbf{x}_{k-1}), & \text{если } \tilde{f}_{k-1}(\mathbf{x}) > 0, & \text{иначе} \\ f(\mathbf{x}_k). & & \end{cases}$$

Тогда задача

$$\sup \{\tilde{f}(\mathbf{x}) : \tilde{g}(\mathbf{x}) \leq 0, \mathbf{x} \in \tilde{M}\} \quad (3)$$

является конструктивным описанием символической модели (1) и допускает аналитические исследования, а также численные расчеты.

В связи с построением модели (3) интересно проследить идею *обучения*, связанную с процессом последовательного пополнения *материала наблюдений* G', G'', M', M'' и $\{f(\mathbf{x}) : \mathbf{x} \in Q\}$. Это дает возможность все более точной идентификации функции f , ограничения $g(\mathbf{x}) \leq 0$ и множества M (если и не в целом, то в окрестности оптимального множества). В основе этого процесса лежит предположение о существовании *эксперта*, который безошибочно диагностирует любой предъявляемый элемент $\mathbf{x} \in \mathbf{R}^n$ на предмет оценки значения $f(\mathbf{x})$, выполнимости неравенства $g(\mathbf{x}) \leq 0$ и включения $\mathbf{x} \in M$. Если обозначить через Σ_t весь *материал обучения* (наблюдений) на момент t , а через $\Phi[\Sigma_t](\mathbf{x})$ — итерационный оператор для модели (3) при $\tilde{f}(\mathbf{x}) = \tilde{f}_t(\mathbf{x})$, $\tilde{g}(\mathbf{x}) = \tilde{g}_t(\mathbf{x})$, $\tilde{M} = \tilde{M}_t$, т. е. для

$$\sup \{\tilde{f}_t(\mathbf{x}) : \tilde{g}_t(\mathbf{x}) \leq 0, \mathbf{x} \in \tilde{M}_t\},$$

то нестационарный процесс

$$\{\mathbf{x}_{t+1} = \varphi [\Sigma_t] (\mathbf{x}_t)\}_{\forall t} \quad (4)$$

и реализует идею обучения. Пополнение материала обучения может происходить, в частности, следующим образом. Пусть уже построен отрезок $\mathbf{x}_0, \dots, \mathbf{x}_N$ процесса (4) и $\mathbf{x}_{N+1} = \varphi [\Sigma_N] (\mathbf{x}_N)$. Тогда элемент \mathbf{x}_{N+1} диагностируется экспертом, в соответствии с чем пополняется одно из множеств каждой пары G' и G'' , M' и M'' , а множество Q_N переходит в $Q_N \cup \{\mathbf{x}_{N+1}\}$.

Отметим, что некоторые из рассмотренных выше задач дискриминантного анализа, решаемых с целью нахождения разделяющих функций $d(\mathbf{x})$, $f_j(\mathbf{x})$ ($\forall j \in N_{k-1}$) и отображения $g(\mathbf{x})$, могут быть неразрешимыми, в частности за счет узости выбранных классов \mathcal{G} , \mathcal{D} и \mathcal{F} . Тогда обычное понятие решения можно заменить на понятие *дискретного приближения*, примером чего может служить понятие *комитета* (или *комитетного решения*). Например, комитетом системы (2) в классе \mathcal{D} называется такое конечное множество $K = \{d_1(\cdot), \dots, d_q(\cdot)\} \subset \mathcal{D}$, что всякому неравенству системы (2) удовлетворяют более половины элементов множества K . При этом под множеством, допустимым по ограничению $\mathbf{x} \in M$, понимается множество $\{\mathbf{x}: d_i(\mathbf{x}) \leq 0 \text{ более чем для половины индексов } i = 1, \dots, q\}$.

Комитетные конструкции могут рассматриваться в более широком контексте теории принятия решений: вместо единственного решения (в виде точки или множества — как элемента) вводится «размытое» решение в форме системы решений (точек или множеств), каждому из которых может быть приписана вероятность его использования.

Коснемся некоторых идейных источников книги. Системы неравенств можно считать центральным объектом в используемом в книге математическом аппарате, а задачи математического программирования и дискриминантного анализа — производными. Наиболее важным и хорошо изученным классом систем неравенств является класс систем линейных неравенств (конечных и бесконечных). Всесторонний анализ таких систем дан в монографии [104]. Связь теории линейных неравенств с линейным программированием является особенно ощутимой. Так, например, двойственность в линейном программировании есть не что

иное, как более или менее очевидная переформулировка некоторых фактов теории линейных неравенств.

Значительной проблемой для систем неравенств является разработка методов их решения; последние составляют важную часть методов вычислительной математики. Исследования по итерационным методам решения систем линейных и выпуклых неравенств, излагаемые в данной книге, иницированы работой Т. С. Моцкина и И. И. Шенберга [118]. В последней, исходя из высказанного в 1922 г. Л. Фейером замечания о том, что для $A \subset \mathbb{R}^n$

$$\overline{x} \in \overline{\text{co } A} \Leftrightarrow \{x: |x - y| < |\overline{x} - y|, \forall y \in A\} = \emptyset,$$

рассмотрены свойства *поточечно-монотонных* относительно A последовательностей $\{x_t\}$ (*монотонных по Фейеру*), т. е. таких, что

$$|x_{t+1} - y| < |x_t - y|, \forall y \in A, \forall t.$$

Развитию фейеровских методов, использованию их в линейном и выпуклом программировании, а также в решении задач распознавания образов, посвящены работы [18—30, 32, 34, 79]. Исследованию таких методов (с ориентацией на решение задач линейного программирования большой размерности) посвящены работы [6—9].

Фейеровские методы привлекли внимание и специалистов в области распознавания образов [39, 79, 101]. Обнаружилось, что многие процедуры обучения в дискриминантном анализе принимают форму некоторых разновидностей фейеровских методов.

Понятие комитета системы однородных линейных неравенств было введено в [112]. Это понятие в дальнейшем было подробно исследовано и обобщено в работах одного из авторов [73—75, 78, 80, 84]. Новые комитетные конструкции нашли применение в задачах классификации, диагностике, в оптимизации при несовместных ограничениях, в многокритериальных задачах математического программирования. Отметим, что одним из интересных и важных с теоретической точки зрения методов построения комитета (в частности, минимального) является метод, основанный на выделении максимальных совместных подсистем систем неравенств с помощью сверток С. Н. Черникова [105].

Механизм моделирования плохо формализуемых ограничений в задачах оптимизации с помощью процедур распознавания образов впервые в явном виде был сформулирован, по-видимому, в [76, 77]. Этому вопросу посвящены также работы [40, 74, 79, 80]. Отметим большое значение исследований Ю. И. Журавлева, посвященных алгебрам алгоритмов для плохо определенных задач, в частности для задач распознавания образов [45—50].

В основе материала по нестационарным процессам лежат работы авторов [35—38, 39].

Сделаем одно замечание к содержанию гл. VI. В ней, помимо перечисления областей, направлений и отдельных прикладных задач в экономике, технике, биологии и медицине, содержится схематический разбор конкретных постановок, по некоторым из которых накоплен значительный опыт счета. Степень подробности разбора отвечает тому уровню, который позволяет, опираясь на результаты книги, составить модель конкретной задачи и решить ее.

В книгу включено овольно подробное описание решения задачи объемно-календарного планирования крупного машиностроительного предприятия; посвященный этому § 4 главы VI написан основными исполнителями по этой теме—В. М. Кисляком и А. С. Садчиковым. В разработке этой темы принимали также участие Г. Ф. Корнилова, Н. Н. Глезер, Л. Д. Попов и Л. Т. Коротаева.

Содержание книги не является замкнутым в том смысле, что в ней нет доказательств всех фактов, которые оказываются необходимыми по ходу дела. Например, на свойства субдифференциала выпуклых функций, теорему Какутани о существовании неподвижной точки точечно-множественного отображения делаются ссылки.

Авторы не ставили перед собой цель дать библиографический обзор всех работ, которые хотя и родственны в том или ином смысле основной идее книги, но тем не менее не имеют непосредственного отношения к рассматриваемым здесь вопросам.

Значительная часть работ, приведенных в библиографии, в тексте не упоминается.

Для книги принят следующий порядок нумерации параграфов и пунктов, формул, а также теорем, лемм

и определений. Параграфы нумеруются одной цифрой, в случае ссылки на него пишется номер главы (например, § 1 гл. II). Внутри параграфов пунктам, которым в книге даны свои названия, присваивается двух-индексный номер, первый индекс означает номер параграфа. Наконец, теоремы, леммы, определения, а также формулы нумеруются тремя индексами: первый из них — это номер главы, второй — номер параграфа, а третий — номер самой теоремы (леммы, определения, формулы).

ГЛАВА I

**НЕСТАЦИОНАРНЫЕ СИСТЕМЫ И МЕТОДЫ
ИХ МОДЕЛИРОВАНИЯ**

Математическое программирование является мощным средством моделирования и оптимизации технико-экономических и природных систем. Однако современная ситуация в области приложений математического программирования характеризуется расширением класса решаемых задач, их сложностью, ростом числа неопределенных и плохо формализуемых факторов, требующих учета. Это приводит к уменьшению относительного объема применяемых формальных методов, однако математическая основа играет здесь по крайней мере упорядочивающую роль. На этой основе создается общий подход, сочетающий количественный и качественный анализ.

К качеству работ, посвященных моделированию сложных систем (например, средствами математического программирования), предъявляются высокие требования, в частности каждый этап анализа моделируемой системы и ее модели должен быть основан на объективных процедурах и по возможности формализован; процедура моделирования должна предусматривать возможность использования новой текущей информации о системе для целей совершенствования модели от шага к шагу, ее уточнения и идентификации.

Методы математического программирования представляются одним из эффективных средств решения новых проблем, но в то же время возникла необходимость в дальнейшем развитии этих методов, в их адаптации к моделированию нестационарных и плохо определенных ограничений, процессов и систем, к использованию неформальных решений, в частности экспертиз. Естественным и в то же время эффективным методом учета плохо формализуемых ограничений является математический аппарат распознавания образов.

Построение и анализ нестационарных процессов оптимизации осуществляется в настоящей работе на основе синтеза идей математического программирования и распознавания образов.

§ 1. Базовые модели

1.1. Системы неравенств. Системы неравенств и уравнений являются фундаментальным объектом математики, широко используемым как в самой математике, так и в приложениях.

Пусть X — некоторое множество (чаще, из некоторого линейного пространства), $X^* \subset \{X \rightarrow \mathbf{R}\}$.

Систему неравенств обычного вида запишем в форме

$$\begin{aligned} x_\alpha^*(x) \leq 0, \quad x_\beta^*(x) < 0, \quad x_\gamma^*(x) = 0 \\ (\forall \alpha \in J_{\leq}, \quad \forall \beta \in J_{<}, \quad \forall \gamma \in J_{=}); \end{aligned} \quad (1.1.1)$$

здесь $x \in X$, $x_\tau^* \in X^*$ ($\forall \tau \in J = J_{\leq} \cup J_{<} \cup J_{=}$). Решить эту систему — значит найти x , удовлетворяющий каждому из ее соотношений.

Систему

$$\begin{aligned} x^*(y_\alpha) \leq 0, \quad x^*(y_\beta) < 0, \quad x^*(y_\gamma) = 0 \\ (\forall \alpha \in J_{\leq}^*, \quad \forall \beta \in J_{<}^*, \quad \forall \gamma \in J_{=}^*), \end{aligned} \quad (1.1.2)$$

в которой $y_\tau \in X$ ($\forall \tau \in J^* = J_{\leq}^* \cup J_{<}^* \cup J_{=}^*$), $x^* \in X^*$ — искомый элемент, назовем системой сопряженного вида.

На решения систем (1.1.1) и (1.1.2) могут быть наложены свои связи, которые мы запишем в виде

$$\begin{aligned} g_\alpha(z) \leq 0, \quad g_\beta(z) < 0, \quad g_\gamma(z) = 0 \\ (\forall \alpha \in I_{\leq}, \quad \forall \beta \in I_{<}, \quad \forall \gamma \in I_{=}), \end{aligned} \quad (1.1.3)$$

где $z = [x, x^*] \in X \times X^*$, $g_\tau \in \{X \times X^* \rightarrow \mathbf{R}\}$ ($\forall \tau \in I = I_{\leq} \cup I_{<} \cup I_{=}$).

Систему, состоящую из объединения (1.1.1) — (1.1.3), назовем системой неравенств смешанного вида.

Примеры.

1. Конечная система линейных неравенств и уравнений над пространством \mathbf{R}^n : $(a_j, x) \leq b_j \quad \forall j \in \mathbf{N}_m$, здесь $a_j, x \in \mathbf{R}^n$, $b_j \in \mathbf{R}$ ($\forall j \in \mathbf{N}_m$).

В матричном виде эту систему можно записать так:

$$Ax \leq b, \quad (1.1.4)$$

где

$$A = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}, \quad b = [b_1, \dots, b_m]^T.$$

Заметим, что задача линейного программирования (ЛП)

$$\max \{(c, x): Ax \leq b, x \geq 0\} \quad (1.1.5)$$

сводится к нахождению хотя бы одного решения конечной системы линейных неравенств над пространством \mathbf{R}^{n+m} :

$$\begin{aligned} Ax \leq b, \quad A^T u \geq c, \quad x \geq 0, \quad u \geq 0, \\ (c, x) \geq (b, u). \end{aligned} \quad (1.1.6)$$

Более точно: если $[\tilde{x}, \tilde{u}]$ —решение системы (1.1.6), то \tilde{x} —оптимальный вектор задачи (1.1.5), а \tilde{u} —оптимальный вектор двойственной задачи

$$\min \{(b, u): A^T u \geq c, u \geq 0\}.$$

2. Пусть M и N —выпуклые замкнутые множества из \mathbf{R}^n , одно из которых ограничено, и $M \cap N = \emptyset$. Задача их строгой разделимости аффинной функцией $(c, x) + b$ сводится к решению системы линейных (относительно $z = [c, b] \in \mathbf{R}^{n+1}$) неравенств сопряженного вида

$$\begin{aligned} (c, x) + b < 0, \quad (c, y) + b > 0 \\ (\forall x \in M, \quad \forall y \in N). \end{aligned}$$

3. Пусть $\varphi \in \{\mathbf{R}^n \rightarrow \mathbf{R}^n\}$, $M \subset \mathbf{R}^n$. С неравенством $(\varphi(x), y - x) \leq 0$, $y \in M$, свяжем задачу нахождения $x \in M$ такого, что

$$(\varphi(\bar{x}), y - \bar{x}) \leq 0 \quad (\forall y \in M) \quad (1.1.7)$$

(решение вариационного неравенства). Эта постановка эквивалентна решению (относительно \bar{x} и \bar{c}) следующей системы неравенств смешанного вида:

$$(\bar{c}, y - \bar{x}) \leq 0 \quad (\forall y \in M), \quad \bar{c} = \varphi(\bar{x}).$$

С решением вариационного неравенства (1.1.7) связана, например, задача

$$\max \{f(\mathbf{x}), \mathbf{x} \in M\}, \quad (1.1.8)$$

где $f(\mathbf{x})$ — вогнутая дифференцируемая функция. Эта связь состоит в том, что при $\varphi(\mathbf{x}) = \nabla f(\mathbf{x})$ вектор $\bar{\mathbf{x}}$, удовлетворяющий (1.1.7), является оптимальным вектором задачи (1.1.8).

4. Пусть некоторая задача планирования описывается с помощью модели

$$\sup \{f_0(\mathbf{x}): f_j(\mathbf{x}) \leq 0 \ (\forall j \in N_m), x^*(\mathbf{x}) \leq 0, \\ x^*(\mathbf{y}) \leq 0 \ (\forall \mathbf{y} \in Y_1), x^*(\mathbf{y}) > 0 \ (\forall \mathbf{y} \in Y_2)\},$$

в которой функционалы $f_0(\mathbf{x}), f_j(\mathbf{x})$ ($\forall j \in N_m$) являются известными, а x^* требует идентификации. Идентификация может быть осуществлена за счет сведений о том, что планы $\mathbf{y} \in Y_1$ удовлетворяют идентифицируемому неравенству, а планы $\mathbf{y} \in Y_2$ не удовлетворяют. Эти условия как раз и фигурируют в системе ограничений нашей модели. Очевидно, что допустимое множество в этой модели задается системой неравенств смешанного вида.

1.2. Модель математического программирования. Под моделью математического программирования (в узком смысле) будем понимать задачу

$$\sup \{f_0(\mathbf{x}): f_j(\mathbf{x}) \leq 0 \ (\forall j \in N_m), \mathbf{x} \in M \subset X\}, \quad (1.1.9)$$

в которой $\{f_i(\mathbf{x}): \forall i \in N_m\}$ — функции, определенные на некотором множестве X . Чаще всего в роли X выступает линейное пространство с теми или иными ограничениями.

Модель (1.1.9) является чрезвычайно общей, она допускает даже в рамках чистой математики большое количество интерпретаций (в вариационном исчислении, теории некорректных задач, математической экономике, теории оптимального управления, теории приближений функций и т. д.). Именно потому эта модель находит широкие приложения.

Для случая, когда X — вещественное линейное пространство, а (1.1.9) — задача выпуклого программирования (т. е. функции $\{-f_0(\mathbf{x}), f_j(\mathbf{x}): \forall j \in N_m\}$ и множество M выпуклы), приведем необходимые и достаточные условия оптимальности [33].

Назовем *K.—T. условием* любое условие на задачу (1.1.9), гарантирующее для нее справедливость теоремы Куна—Таккера. Последняя состоит в том, что для $\bar{x} \in \text{Arg}$ (1.1.9) существует неотрицательный вектор $\bar{u} \in \mathbb{R}^m$ такой, что

$$F(x, \bar{u}) \leq F(\bar{x}, \bar{u}) \leq F(\bar{x}, u)$$

для всех $x \in M$, $u \geq 0$; здесь

$$F(x, u) = f_0(x) - \sum_{j=1}^m u_j f_j(x)$$

— функция Лагранжа, отвечающая задаче (1.1.9).

Координаты вектора \bar{u} называются *оптимальными множителями Лагранжа* (или *оптимальными двойственными оценками ограничений*), а вектор $[\bar{x}, \bar{u}]$ — *седловой точкой* функции $F(x, u)$.

Примером *K.—T. условия* может служить следующее: существует вектор $p \in M$ такой, что $f_j(p) \leq 0$ ($\forall j \in N_m$) и $f_j(p) < 0$, если функция $f_j(x)$ не является линейной.

Если задача (1.1.9) является выпуклой, разрешимой (пусть при $x = \bar{x}$) и удовлетворяет *K.—T. условию*, то при некоторых $h_0 \in \partial f_0(\bar{x})$, $h_j \in \partial f_j(\bar{x})$ ($\forall j \in N_m$) выполнены соотношения

$$h_0 = \sum_{j=1}^m \bar{u}_j h_j + h, \quad \bar{u}_j f_j(\bar{x}) = 0 \quad (\forall j \in N_m), \quad (1.1.10)$$

где $\bar{u} = [\bar{u}_1, \dots, \bar{u}_m]$ — вектор двойственных оценок ограничений задачи (1.1.9), а вектор h такой, что $(h, x - \bar{x}) \geq 0$, $\forall x \in M$.

Обратно, если вектор $[\bar{x}, \bar{u}]$ подчинен соотношениям (1.1.10), то он является седловой точкой функции Лагранжа $F(x, u)$, что влечет $\bar{x} \in \text{Arg}$ (1.1.9) [33].

1.3. Модели распознавания образов. Пусть X — некоторое множество, $X^* \subset \{X \rightarrow \mathbb{R}\}$, $X_1 \subset X$, $X_2 \subset X$, $Y \subset X$.

В математическом содержании распознавания образов выделим следующие три задачи.

1. Задача дискриминантного анализа (для двух образов). Найти решения максимальных совместных подсистем системы сопряженного вида

$$x^*(x) > 0 \quad (\forall x \in X_1), \quad x^*(x) < 0 \quad (\forall x \in X_2), \quad x^* \in X^*.$$

Если данная система совместна, то задача дискриминантного анализа состоит в отыскании элемента x^* , удовлетворяющего всем соотношениям системы.

2. Задача таксономии. Найти максимальные совместные подсистемы системы сопряженного вида относительно x^* :

$$x^*(x) < 0 \quad (\forall x \in Y), \quad x^* \in X^*.$$

Если подсистема

$$x^*(x) < 0 \quad (\forall x \in V \subset Y)$$

является максимально совместной, то множество V называется *таксоном*.

3. Задача оценки информативности. Пусть задан класс преобразований $\Phi \subset \{X \rightarrow Z\}$. Требуется найти $\varphi \in \Phi$ такое, что решение задачи дискриминантного анализа для множеств $\varphi(X_1)$ и $\varphi(X_2)$ удовлетворяет некоторому критерию качества. Если, например, в Z^2 введена некоторая метрика ρ , то этим критерием может быть

$$\sup_{\varphi \in \Phi} \rho(\varphi(X_1), \varphi(X_2)).$$

1.4. Взаимосвязи моделей. Задача решения систем неравенств является, очевидно, частным случаем задачи математического программирования. С другой стороны, задача математического программирования при определенных условиях, обеспечивающих соотношения двойственности, сводится к решению некоторой системы неравенств. Для задачи ЛП (1.1.5) соответствующей системой является (1.1.6). В более общем случае, а именно, для задачи выпуклого программирования (1.1.9) сводимость к системе неравенств содержится в соотношении (1.1.10).

Коснемся вопроса о взаимосвязи классов задач математического программирования и распознавания образов. Задачи распознавания могут быть интерпретированы как некоторые специальные задачи оптимизации, а эти последние могут быть наделены смыслом задач распознавания образов. Это будет видно из того, что как те, так и другие укладываются в следующую модель классификации объектов.

Пусть задано некоторое множество X . Элементы $x \in X$ суть модели некоторых объектов, которые функциони-

руют как элементы некоторой реальной системы. Нужно найти функционал $x^* \in X^* \subset \{X \rightarrow \mathbf{R}\}$, классифицирующий элементы множества X . Цель классификации: установить *специализацию* элементов x такую, чтобы функционирование их в системе в соответствии с этой специализацией было допустимым с точки зрения всей системы и чтобы удовлетворялся тот или иной критерий оптимальности. Задачу специализации можно сформулировать как задачу нахождения x^* из условий

$$\varphi_j(\{x^*(x) : x \in X\}) \leq 0 \quad (\forall j \in N_m), \quad (1.1.11)$$

где φ_j — функции, диктующие ограничения на допустимость специализации. При наличии функции $f(x^*)$, характеризующей качество специализации, возникает задача

$$\sup \{f(x^*) : (1.1.11)\}. \quad (1.1.12)$$

Если исходить из обычной экономико-технологической интерпретации задачи математического программирования, то, можно считать, что в (1.1.12) x — вектор-модель *технологического способа*, $x^*(x)$ — интенсивность его использования. Кроме того, в предположении, что X — конечное множество в \mathbf{R}^m : $X = \{p_1, \dots, p_n\} \subset \mathbf{R}^m$, а для используемых ингредиентов выполняется гипотеза однородности и аддитивности (это определяет и вид функций φ_j), задача (1.1.12) (с точностью до простейших математических преобразований) принимает вид

$$\max \left\{ \sum_{i=1}^n c_i x^*(p_i) : \sum_{i=1}^n p_i x^*(p_i) \leq p_0, \right. \\ \left. x^*(p_i) \geq 0 \quad (\forall i \in N_n) \right\},$$

где c_i — оценка результата функционирования i -го технологического способа с единичной интенсивностью, p_0 — вектор запасов расходуемых ингредиентов. Таким образом, в частных предположениях мы пришли к задаче ЛП в форме (1.1.5).

Наконец, заметим, что в основу изучения нового объекта — плохо формализуемого ограничения, включаемого в формулировку задач оптимизации (независимо от их природы), могут быть положены, как будет показано позднее, те же методы математического программирования и распознавания образов.

§ 2. Математическое программирование, его возможности и границы применимости

2.1. Необходимость развития комплексных методов принятия решений. Методы математического программирования завоевывают все новые области приложений. Наряду с традиционными задачами анализа экономических ситуаций успешно решаются задачи в области физики, химии, технического проектирования, оптимального конструирования и многих других областях.

Это связано, с одной стороны, с эффективностью методов математического программирования, а с другой — с чрезвычайной общностью моделей, допускающей богатый спектр интерпретаций. Упомянем, например, экономическую интерпретацию (задача об оптимальных интенсивностях использования технологических способов), термодинамическую интерпретацию. Кроме того, сама идея оптимизации, получившая в теории математического программирования эффективные средства реализаций, отвечает современным требованиям к качеству научных исследований и их приложений.

Возникшие в последнее время новые области теории и приложений математического программирования отвечают новым проблемам научно-технического развития, задачам его интегральной оценки. В качестве примера можно привести комплексную проблему рационального природопользования. Один из главных объектов исследования в этой проблеме — производство — характеризуется непрерывным развитием и усложнением, ростом масштаба и динамичности. Таким образом, новые задачи математического программирования связаны со стохастической природой моделируемых объектов, с недостаточностью информации о них. При этом требования к качеству решения задач планирования и управления повышаются: все этапы моделирования и анализа должны быть основаны на объективных процедурах, алгоритмы должны оперативно реагировать на изменения информации; система факторов, влияющих на функционирование объектов, должна учитываться как можно более полно.

Итак, при анализе сложных объектов приходится учитывать большое число взаимосвязанных факторов, включая плохо формализуемые. При этом реально доступная

информация может иметь любую степень детерминированности и поступать в произвольной форме. Кроме того, информация о моделируемом объекте, отвечая его нестационарности, может эволюционировать во времени.

Известно, что современный этап разработки АСУ связан с проблемой стыковки человеческого и машинного интеллектов, соединения формального аппарата математических моделей и методов с реальными процессами принятия решений.

Вообще, в процессе принятия практических решений, кроме сферы формальных методов и моделей, участвует сфера дополнительных средств неформального (в частности, постоптимизационного) анализа. Последняя включает суждения экспертов для учета так называемых «вне-модельных» факторов. При решении достаточно сложных практических задач удельный вес методов этой второй сферы может быть настолько большим, что использование точных методов вообще может потерять смысл. В связи с этим возникает необходимость точного анализа возможностей комплексных методов принятия решений, включающих и процедуры математического программирования, и средства имитации, экспертизы и т. д.

Задача адаптации методов оптимизации к анализу сложных объектов (систем), на наш взгляд, может решаться с помощью таких математических и вычислительных средств, как итерационные методы, обогащенные возможностями учета от шага к шагу эволюции всей системы данных об объекте.

2.2. Природа плохо формализуемых факторов, их моделирование и анализ. Характер функционирования любых объектов зависит от множества факторов. При математическом моделировании факторы приобретают вид математических конструкций — таких, как параметры состояния объекта и среды, ограничения на область совместного и согласованного варьирования этих параметров (при удовлетворении этих ограничений объект функционирует нормально, не разрушая своей структуры), целевые функционалы, балансовые соотношения, другие соотношения между параметрами, отражающие различные виды взаимодействия (физико-химические, биологические, социальные) и т. д.

Однако в случае достаточно сложных объектов многие факторы оказываются плохо формализуемыми, т. е. их отражение математическими конструкциями оказывается либо невозможным, либо существенно неадекватным. Человеческий фактор — один из главных источников плохой формализуемости некоторых этапов процесса принятия решений. Другие источники: недостаточные научно-техническое и информационное обеспечение.

Приведем некоторые примеры. Плохо формализуемые факторы (ограничения, связи, цели) могут встретиться: при моделировании отображений (рис. 1), когда отсутствуют сведения о внутренней структуре преобразователя Π (в частности, в задаче оценки параметров смесей веществ);

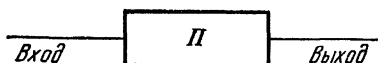


Рис. 1.

при управлении вычислительными процедурами: выбор *субъективных* констант алгоритмов, агрегирование исходных данных.

С точки зрения возможности применения формальных средств анализа, в первую очередь математического программирования, задачи планирования и управления можно разбить на следующие группы (в порядке возрастания сложности):

элементарная переработка информации по стандартным формулам;

полностью формализованные задачи оптимизации;

плохо формализованные задачи;

неформализованные задачи;

неформализуемые задачи.

Неформализованные задачи требуют исследования новых свойств управляемых систем. Пока же они решаются на основе опыта, интуиции, ассоциативного мышления. В плохо формализованных задачах существенно использование экспертиз и экспериментов в сочетании с формальными методами оптимизации.

Рассмотрим природу плохой формализуемости. Формальный подход в основу исследования объектов и явлений кладет их форму. Он предполагает подчеркивание

понятийно-рационального аспекта в противоположность содержательному. На самом деле формализация является лишь промежуточным этапом, допускающим автоматизацию мыслительных операций. На этом этапе содержание объектов исследуется с помощью выделенных элементов их формы.

Этап формализации предполагает построение некоторой структуры, согласно которой распределяют и связывают друг с другом высказывания (причем это делается в соответствии с чисто формальными признаками). Целью построения такой структуры является логическое понимание. Таким образом, формализация есть придание порядка, структуризация. Поэтому плохо формализуемыми являются аморфные, слабо структуризованные явления.

При формальном рассмотрении принимают во внимание не содержание высказываний, а их синтаксис и структуру; при этом выражения преобразуются согласно строгим правилам. Если рассматриваются явления, где возможна неоднозначность отдельных операций, где присутствуют неопределенные отношения, то также возникают плохо формализуемые факторы.

Вообще говоря, все реальные объекты и отношения являются неформализуемыми, но мы употребляем термин «плохая формализуемость» в том смысле, насколько возможно получение реальных результатов при отражении реальных отношений математическими моделями.

Конкретными причинами плохой формализуемости могут быть следующие:

многокомпонентность систем, разнообразная природа выделяемых признаков, их неоднородность;

недостаточное эмпирическое знание о моделируемом объекте;

невозможность проникновения «внутри» объекта, в его строение (при попытке такого проникновения объект, например, разрушается как таковой);

наличие человеческого фактора (люди часто не склонны раскрывать причины их деятельности, реальные, а не формальные связи);

сложные взаимозависимости, для анализа которых требуются слишком большие затраты ресурсов (в случае, например, диффузных систем);

наличие интуитивно оцениваемых (экспертами) показателей и интуитивных оценок допустимости вариантов по ряду критериев.

2.3. Некоторые примеры. Приведем примеры плохо формализуемых ограничений:

зависимости результатов процессов от входных переменных (качество смесей материалов, выход химических процессов, качество продукции);

требования, внутренние для системы (балансовые соотношения, пределы вариации параметров, ограничения на область действия или границы процессов, на химический состав материалов, условия обеспечения биологических процессов, физические и технологические условия и др.);

внешние ограничения (на выбор места строительства предприятия, на качество территории и окружающей среды, условия сохранения биогеоценозов, требования к составу выпускаемой продукции);

управленческие ограничения;

ограничения на допустимость планов с точки зрения динамики, отдаленных последствий принимаемых решений (например, возможность восстановления природных ресурсов);

ограничения на интенсивности использования технологических способов, при которых реакции способов укладываются в заданную аналитическую схему (например, соответствуют аксиомам аддитивности и однородности);

ограничения на допустимость технологических способов;

допущения, обеспечивающие численное решение;

соотношения, связывающие переменные моделей, написанных на разных языках;

ограничения, диктуемые интересами надсистемы (при согласовании моделей, образующих иерархическую систему);

условия оптимальности;

условия, отвечающие противоречивым наборам критериев;

условия, грубо отсекающие заведомо плохие планы (требования безопасности, ограничения уровня производительности, социальные, психо-физиологические и др.);

логические ограничения.

Приведем более конкретные примеры задач с плохо формализуемыми ограничениями:

прогнозирование качественных показателей или дискретных коэффициентов;

прогноз нормативов на новые процессы;

структуризация множества возможных вариантов неопределенной информации, выделение эталонных вариантов информации;

этап технико-экономического анализа решения задачи оптимального планирования;

задачи размещения, районирования, оценки ресурсов, типологии и классификации, управления производством, проектирования, прогнозирования, социологии, нахождения производственных функций, унификации и стандартизации, агрегирования и др.

Плохо формализуемые ограничения могут также возникать при моделировании процессов с неполной информацией.

Возможны различные содержательные интерпретации процессов при неполной системе исходных данных. В частности, речь может идти относительно ситуаций, в которых:

поведение моделируемой системы детерминировано, но информация о ней известна неполностью;

нестационарность определяющей информации в точности соответствует нестационарности самой моделируемой системы;

нестационарность информации является искусственным приемом в вычислительном алгоритме, когда информационная составляющая выступает в качестве управления алгоритмом;

нестационарность информации отвечает плохой формализуемости ограничений.

2.4. Традиционный и новый подходы к моделированию плохо формализуемых факторов. Обычно вид плохо формализуемых зависимостей указывается «заказчиком», который зачастую вынужден представлять эти зависимости неадекватно. Необходимо этот этап сделать более объективным, включив в него формальные средства, экспертизы и обработку их результатов, в частности с помощью методов распознавания образов.

Первоначальный этап построения модели часто состоит в «инвентаризации» всех имеющихся представлений и све-

дений (количественных и качественных) о моделируемой системе или процессе. Определенным шагом к формализации этих представлений можно считать составление номенклатуры переменных и связывающих их ограничений. Выбор вида модели, а также идентификация элементов определяющей информации, могут быть осуществлены на основе распознавания образов.

В соответствии со сказанным могут быть развиты методы математического программирования, позволяющие учитывать плохо формализуемые ограничения и цели [37]. Смысл их в том, что алгоритмы (итерационные) включают процедуру обучения диагностике планов по результатам экспертиз и экспериментов. Эти методы предусматривают обращение к *эксперту* (либо эксперименту) для оценки по плохо формализуемым критериям исследуемого множества планов. Варианты алгоритмов отвечают различным формам диалога с экспертом и характеру его информации о планах. Выяснилось, что простейшие формы диалога, распространенные в практике решения экономико-математических задач, лишь при очень сильных предположениях дают сходимость последовательности возникающих векторов к оптимальному плану, а в обычных условиях они только позволяют моделировать плохо формализуемые ограничения.

§ 3. Нестационарные процессы математического программирования

Нестационарной моделью математического программирования будем называть модель класса математических программ, реализующуюся в форме дискретной (актуальной или потенциальной) последовательности ее состояний \mathfrak{M}_t ($\forall t \in \mathbb{N}$). Если φ_t — итерационное отображение (вообще говоря, точечно-множественное), соотнесенное модели \mathfrak{M}_t , то процесс $\{x_{t+1} \in \varphi_t(x_t)\}_{\forall t}$ назовем *нестационарным*. Понятие нестационарной модели (равным образом и нестационарного процесса — как итерационного процесса для нестационарной модели) мы производим от понятия нестационарной экономики, т. е. экономики с эволюционирующими количеством и качеством технологий. В этом смысле, например, задача линейного программирования (1.1.5) (в своей естественной экономико-технологической

интерпретации) является стационарной. Если ее вектор исходных данных $s = [A, b, c]$ эволюционирует (т. е. нестационарен), то, формализуя этот факт последовательностью $\{s_t\}$ реализаций этого векторного параметра, мы приходим к понятию нестационарной модели линейного программирования. Однако в понятие нестационарной модели и нестационарного процесса мы вкладываем более общий смысл, который будет обсужден ниже.

В предыдущем параграфе было указано на те трудности, которые сопровождают применение математического программирования к моделированию и оптимизации систем в условиях большой размерности, недостаточности априорной информации, нестационарности, наличия плохо формализуемых факторов, «расплывчатости» критериев оценки принимаемых решений и т. д.

На наш взгляд, эти трудности могут преодолеваться (во всяком случае, для многих конкретных систем, в частности технико-экономических) путем обобщения и развития итерационных методов математического программирования.

3.1. Характеристика итерационных методов. Пусть имеется некоторая задача конструктивного выделения элемента неявно задаваемого множества $M \subset \mathbb{R}^n$ (им может быть, например, допустимое или оптимальное множество в задаче математического программирования). Под *итерационным отображением* (вообще говоря, точечно-множественным) для этой задачи будем понимать отображение

$$\varphi \in \{\mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}\}$$

такое, что при произвольном начальном приближении x_0 любая последовательность, порождаемая рекуррентно включением

$$x_{t+1} \in \varphi(x_t), \quad (1.3.1)$$

ограничена и все ее предельные точки содержатся в M .

Мы будем исходить из такого понимания итерационных процессов, т. е. процессов, порождаемых итерационным отображением φ , когда в качестве их характеристических свойств выступают следующие:

1) бесконечность процесса (это свойство было уже отмечено);

- 2) логическая и арифметическая простота итерации (что ведет к простоте программных реализаций);
- 3) помехоустойчивость;
- 4) инвариантность логической структуры итерационного отображения относительно изменения исходных данных о задаче;
- 5) медленная сходимость.

Интерес к итерационным методам, понимаемым в указанном смысле, определяется тем, что сфера их приложений существенно шире по сравнению с конечными методами. Такие их свойства, как 2) и 3), позволяют, например, преодолевать трудности, связанные с решением задач оптимизации большой размерности. Что касается свойства медленной сходимости, то оно перестает быть серьезной помехой по мере развития вычислительной техники. К тому же конструирование итерационных методов, существенно использующих специфику задачи (в частности, специфику определяющих ее функций), частично снимает остроту этих вопросов уже в условиях средней производительности вычислительных машин. Подчеркнем, что здесь свойство 5) выступает как плата за другие преимущества итерационных методов. Свойство 4) дает возможность развить итерационные процессы для нестационарных задач планирования и управления.

Применительно к процессам, изучаемым в настоящей книге, помехоустойчивость можно понимать в следующем смысле:

- 1) конечное число сбоев в вычислении элементов итерационной последовательности не влияет на конечный результат (это эквивалентно возможности произвольного выбора начального приближения);

- 2) если процесс (1.3.1) заменен на процесс с помехами δ_t :

$$\{x_{t+1} \in \Phi(x_t) + \delta_t\}_{\forall t}, \quad (1.3.2)$$

то при $\{\delta_t\} \rightarrow 0$ конечный результат будет тем же, т. е. $\{x_t\}' \subset M$;

- 3) если в (1.3.2) $|\delta_t| \leq \varepsilon, \forall t$, то при некотором $K > 0$, не зависящем от ε , справедливо неравенство

$$\sup \{ \|x' - M\| : x' \in \{x_t\}' \} \leq K\varepsilon.$$

3.2. Итерационный метод оптимизации с одновременной идентификацией модели по системе исходных данных. Из свойств итерационных методов видно, что организация самого процесса не обязательно должна базироваться на точном знании всей системы числовых данных о задаче. Вычислительный процесс можно вести с системой все более уточняющейся информации. Пусть φ — итерационный оператор, например для задачи (1.1.5) $\mathbf{s} = [A, \mathbf{b}, c]$ — вектор ее исходных данных. Вектор \mathbf{s} можно выделить в φ в качестве параметра: $\varphi(\cdot) = \varphi[\mathbf{s}](\cdot)$. Тогда вместо процесса, задаваемого в силу (1.3.1), можно рассматривать процесс

$$\{\mathbf{x}_{t+1} \in \varphi[\mathbf{s}_t](\mathbf{x}_t)\}_{vt}. \quad (1.3.3)$$

Вектор \mathbf{s}_t на каждом шаге может выбираться (или поступать на вход независимо от вычислителя) из некоторого *информационного множества* S_t . Сужение последнего от шага к шагу будет означать сужение неопределенности, связанной с выбором (поступлением) вектора \mathbf{s} . Если $\bigcap_{(t)} S_t = \{\bar{\mathbf{s}}\}$, $\bar{\mathbf{s}} = [\bar{A}, \bar{\mathbf{b}}, \bar{c}]$, то асимптотически идентифицированной моделью по системе исходных данных будет

$$\max \{(\bar{c}, \mathbf{x}) : \bar{A}\mathbf{x} \leq \bar{\mathbf{b}}, \mathbf{x} \geq 0\}.$$

При определенных условиях для φ мы вправе ожидать, что для последовательности, порожденной процессом (1.3.3), будет выполняться свойство

$$\{\mathbf{x}_t\}' \subset \text{Arg} (1.3.4).$$

Точные результаты о процессе (1.3.3) получены в гл. III.

3.3. Синтез нестационарных процессов. Хорошо известно, например, из практики математического программирования, что организация итерационных процедур сопряжена с необходимостью регулирования той или иной системой констант, характеризующих либо текущее состояние модели, либо состояние самого вычислительного процесса, в частности констант метода, промежуточных констант точности и т. д. Процесс может характеризоваться также наличием экзогенных неконтролируемых параметров, отражающих состояние среды, в которой он реализуется, а также наличием подвижной системы свя-

зей, т. е. соотношений между элементами модели. Последние, как правило, имеют форму предикатов, в числе которых могут быть \leq , $<$, $=$,

$$p_c(x, y) = \begin{cases} 1, & x + y = c, \\ 0, & x + y \neq c, \end{cases}$$

$$p'_c(x, y) = \begin{cases} 1, & x - y = c, \\ 0, & x - y \neq c \end{cases}$$

и т. д.

Сказанное позволяет сделать дальнейшее формальное обобщение процесса (1.3.3), заменив его на

$$\{x_{t+1} \in \Phi[s_t, \omega_t, u_t, \sigma_t](x_t)\}_{\forall t}, \quad (1.3.4)$$

здесь s_t — информационная составляющая модели на момент t , ω_t — вектор экзогенных параметров, u_t — вектор управления (система регулируемых эндогенных параметров), σ_t — система связей. Вектор x удобно интерпретировать как вектор состояния объекта (элемент фазового пространства, пространства признаков, план).

Реализация каждой из последовательностей $\{s_t\}$, $\{\omega_t\}$, $\{u_t\}$ и $\{\sigma_t\}$ может носить как детерминированный, так и стохастический характер и осуществляться в силу как формальных, так и неформальных средств. Добавим, что генерирование последовательности $\{u_t\}$ может отвечать как программному, так и позиционному управлению. Что касается правила преобразования u_t в u_{t+1} , то оно может либо иметь вид точного предписания (формулы, алгоритма), либо быть эвристическим. В частности, последнее реализуется при интерактивном (диалоговом) режиме ведения вычислительного процесса. Наконец, отметим, что формирование контролируемых параметров может осуществляться на основе учета обратных связей:

$$\{[s_t, \omega_t, u_t, \sigma_t, x_t]\}_{t=0}^T \rightarrow [s_{T+1}, \omega_{T+1}, u_{T+1}, \sigma_{T+1}].$$

Подчеркнем, что на итерационные процессы типа (1.3.3) в ряде случаев (разностные методы математической физики, процессы экономической динамики, динамические процессы межпопуляционного баланса и др.) можно смотреть как на содержательно интерпретируемый переход моделируемой системы из состояния в момент t

в последующее состояние на основе тех или иных модельных представлений.

В целом можно говорить о процессе

$$\{x_{t+1} \in I_t(x_t)\}_{vt}, \quad (1.3.5)$$

где I_t — оператор имитации поведения нестационарного объекта, интегрирующий всю систему формальных и неформальных средств (доступных и используемых в момент t) для описания перехода

$$x_t \xrightarrow{I_t} x_{t+1}.$$

Формирование моделей математического программирования и построение для них итерационных операторов существенно затрудняется в том случае, когда имеются плохо формализуемые ограничения (связи) или плохо формализуемые цели. Формальным средством внесения определенности в названные элементы, как уже отмечалось ранее, является математический аппарат распознавания образов. С целью оттенить это обстоятельство процесс (1.3.5) можно было бы переписать в виде

$$\{x_{t+1} \in I_t[R](x_t)\}_{vt}, \quad (1.3.6)$$

где под R понимается оператор, вносящий (средствами распознавания образов) определенность в описание плохо формализуемых элементов модели.

Если Φ — итерационный оператор для задачи математического программирования, то, соединяя (1.3.4) и (1.3.6), получим процесс

$$\{x_{t+1} \in \Phi[s_t, \omega_t, u_t, \sigma_t; R](x_t)\}_{vt}, \quad (1.3.7)$$

называемый *нестационарным процессом математического программирования*. Свойства процесса (1.3.7) показывают, что он потенциально является средством моделирования и оптимизации сложных эволюционирующих систем. Важно, по-видимому, также отметить, что процесс (1.3.7) реализует синтез идентификации и оптимизации модели слабо структуризованных систем в условиях неполноты информации, ее неточности и неопределенности.

3.4. Интерпретации нестационарных процессов математического программирования. Содержательные интерпретации процесса (1.3.7) мы будем относить к задаче

математического программирования, зависящей от векторного параметра y и заданной последовательностью реализаций

$$\max \{f_0(x) : f_j(x, y_t) \leq 0 \ (j = 1, \dots, m), \ x \in M\}_{\forall t} \quad (1.3.8)$$

по параметру y . Не теряя в общности, можно считать, что целевая функция не зависит от параметра y . Действительно, если бы она имела вид $f_0(x, y)$, то можно было бы эквивалентным образом перейти к задаче максимизации x_{n+1} при прежних ограничениях с добавлением нового неравенства $x_{n+1} \leq f_0(x, y)$. Под φ будем понимать итерационное отображение, соответствующее этой задаче.

1. Пусть φ имеет вид $\varphi[s](x)$, т. е. $y = s$ — вектор исходных данных. Тогда процессу (1.3.7), принимающему вид (1.3.3), можно придать интерпретацию, содержащуюся в п. 3.2 настоящего параграфа.

В этом случае процесс можно рассматривать следующим образом: модель (1.3.8) описывает некоторую динамическую ситуацию, а процесс (1.3.3) отслеживает оптимальное множество \bar{M}_t модели, обеспечивая в регулярном (в том или ином смысле) случае малость расстояния от x_t до \bar{M}_t при достаточно больших t . Такой процесс можно назвать *следящим*. Отметим, что модели следящего программирования удобны для машинного экспериментирования, связанного с прогнозированием поведения нестационарных объектов, и изучения влияния на систему вариаций исходных данных.

2. Если $\varphi = \varphi[\omega](x)$ и ω — случайный вектор, то

$$\{x_{t+1} \in \varphi[\omega_t](x_t)\}_{\forall t} \quad (1.3.9)$$

— стохастический процесс (типа стохастической аппроксимации).

3. Пусть $y = u$ — вектор параметров метода: штрафные константы — в методах штрафных функций, шаговый коэффициент — в методах субградиентного спуска, шаговый коэффициент и коэффициент релаксации — в методах фейеровских приближений (см. гл. II), коэффициенты инерции — в методах «тяжелого шарика», промежуточные константы точности и т. д. Тогда

$$\{x_{t+1} \in \varphi[u_t](x_t)\}_{\forall t} \quad (1.3.10)$$

— управляемый по u процесс, направленный на повышение эффективности метода. Этот вопрос будет подробно рассмотрен в § 5 гл. V.

4. В задаче (1.3.8) функции f_j , их число, а также вид соответствующих ограничений могут меняться, т. е. f_j могут принимать вид f_j^t , $j \in N_m(t)$, а отображения τ_j^t :

$$f_j^t \xrightarrow{\tau_j^t} \{\geq, \leq, =\},$$

определяют тип ограничений. Тогда

$$\sigma_t = \{f_j^t, \tau_j^t (\forall j \in N_m(t))\}$$

— эволюционирующая система связей модели (1.3.8), а процесс

$$\{x_{t+1} \in \Phi[\sigma_t](x_t)\}_{\forall t} \quad (1.3.11)$$

отражает динамику системы связей.

5. В (1.3.7) с помощью неформального оператора R могут устраняться недостаточность информации s , неопределенность системы связей σ и т. д. Оператор R реализует средства «дообучения», работы с экспертизами, экспериментами и др.

В связи с этим можно отметить, что при традиционном подходе к моделированию математик требует от «заказчика» указывать вид критериев, целей, зависимостей, ограничений, обычно не интересуясь тем, каким способом «заказчик» формирует эту информацию. Пытаясь учесть «внемодельные» ограничения, эксперту предъявляют не только рассчитанный оптимальный план задачи, не включающий эти ограничения, но и некоторое множество близких к оптимальному (*субоптимальных*) планов, из которых эксперт выбирает с его точки зрения наиболее удовлетворительный. Явное включение подобных средств (в форме оператора R) в итерационное отображение Φ позволяет в более систематизированном виде изучать прикладные возможности диалоговых процедур принятия решений.

6. Частные интерпретации нестационарных процессов (1.3.9)—(1.3.11) очевидным образом порождают варианты интерпретаций комбинированного типа.

7. Если для обычной задачи математического программирования (1.1.9) имеется конечная совокупность итера-

Тогда балансовые уравнения запишутся следующим образом:

$$\mathbf{x}_{t+1} = \Phi[\mathbf{y}_t, \mathbf{u}_t](\mathbf{x}_t) = \mathbf{x}_t - A[\mathbf{y}_t](\mathbf{x}_t) + \Delta[\mathbf{y}_t](\mathbf{x}_t) - \\ - \delta[\mathbf{y}_t](\mathbf{x}_t) + m[\mathbf{y}_t](\mathbf{x}_t) + \mathbf{u}_t,$$

где Φ — итерационный оператор.

Заметим, что на приведенную модель следует смотреть лишь как на общую схему, применение которой возможно только при наличии, во-первых, адекватного описания механизмов (пассивных и активных) регуляции численности популяций, вбираемых составляющими $\Delta[\mathbf{y}_t]$ и $\delta[\mathbf{y}_t]$, и, во-вторых, хорошо разработанной проблемы мониторинга. Управление \mathbf{u}_t может интерпретироваться широко и преследовать разные цели, например цель выхода на заданную структуру численности популяций:

$$\{\mathbf{x}_t\} \rightarrow k \bar{\mathbf{x}} (t \rightarrow +\infty), \quad k > 0.$$

Машинное экспериментирование по данной модели может носить весьма содержательный характер.

9. В связи с процессами типа $\{\mathbf{x}_{t+1} \in \Phi[\mathbf{s}](\mathbf{x}_t)\}_{\forall t}$ могут успешно решаться вопросы *информационной восстановимости*. Под этим мы понимаем следующее. Предположим, что модель процесса, выраженная написанным соотношением, верна, при этом значения \mathbf{x}_t наблюдаемы (т. е. координаты этого вектора могут быть измерены), а информационная составляющая неизвестна, полностью или частично. Возникает задача отыскания информационной составляющей модели \mathbf{s} . Здесь могут быть предложены различные процедуры, смысл которых так или иначе связан с решением задачи (или последовательности задач) относительно \mathbf{s} вида

$$\mathbf{s}_t = \arg \min \left\{ \sum_{i=1}^t |\mathbf{x}_{i+1} - \Phi[\mathbf{s}](\mathbf{x}_i)|^2 : \mathbf{s} \in S \right\}. \quad (1.3.13)$$

Эти процедуры поддаются модификации и на случай нестационарных ситуаций.

Дадим несколько иную интерпретацию задаче типа (1.3.13), связанную с идентификацией оператора Φ , преобразующего сигнал $\mathbf{x} \in \mathbb{R}^n$ в сигнал $\mathbf{y} \in \mathbb{R}^m$ (см. рис. 1). Предположим, что имеются последовательность сигналов, наблюдаемых на входе $\{\mathbf{x}_1, \dots, \mathbf{x}_t\}$, и последователь-

ность соответствующих им сигналов, наблюдаемых на выходе $\{y_1, \dots, y_t\}$, т. е. $\varphi(x_i) = y_i$ ($i = 1, \dots, t$). Пытаясь аппроксимировать оператором φ_t из класса Φ оператор Φ , задачу отыскания φ_t можно поставить в форме

$$\varphi_t = \arg \min \left\{ \sum_{i=1}^t |y_i - \varphi(x_i)|^2 : \varphi \in \Phi \right\}.$$

Если Φ — некоторый класс линейных операторов, то предыдущую задачу можно переписать в виде

$$A_t = \arg \min \left\{ \sum_{i=1}^t |y_i - Ax_i| : A \in \mathfrak{G} \right\}, \quad (1.3.14)$$

где A — переменная матрица размера $m \times n$, \mathfrak{G} — множество матриц, отвечающих линейным преобразованиям из Φ .

Число учитываемых сигналов может расти, и тогда вид отыскиваемой в (1.3.14) матрицы A_t будет уточняться. Если \mathfrak{G} — выпуклое замкнутое множество в $\mathbf{R}^{m \times n}$, то решение задачи (1.3.14) можно осуществить методами фейеровского типа (см. § 2 гл. II).

При гипотезе линейности оператора φ возможно его точное отыскание при достаточном числе сигналов и их достаточном разнообразии.

10. Возможны и другие интерпретации (1.3.7), в частности:

$s_t = \delta_t$ — вектор помех (см. (1.3.2));

нестационарность исходной информации о задаче является искусственным приемом в вычислительных алгоритмах, когда информационная составляющая модели выступает в качестве управления алгоритмами (методы отсечений Кэли [10]);

нестационарность проявляется в последовательности аппроксимаций составляющих задачу функций, либо

в последовательности аппроксимаций бесконечной системы ограничений ее конечными подсистемами в задаче оптимизации.

§ 4. Процесс обмена информацией в некоторых моделях иерархии управления

Рассмотрим подробнее пример нестационарного процесса, реализующего процедуру обмена информацией в некоторой иерархической системе управления.

В последнее время проявляется большой интерес к изучению проблем синтеза и анализа иерархических структур управления. Такие структуры могут быть достаточно сложными, при их изучении ощущается недостаток разработанных методов анализа. Естественным и полезным может быть подход, ориентирующий как на возможность содержательных интерпретаций выбранных моделей иерархии, так и на возможность их анализа средствами разработанного аппарата, например средствами выпуклого программирования и его обобщениями на нестационарный случай. При этом в числе основных условий, налагаемых на математическую модель, выступают условия выпуклости.

4.1. Основные понятия и обозначения. Рассмотрим одну из наиболее простых и естественных иерархических систем управления \mathfrak{M} , задаваемую структурой «правильно» ориентированного дерева (рис. 2).

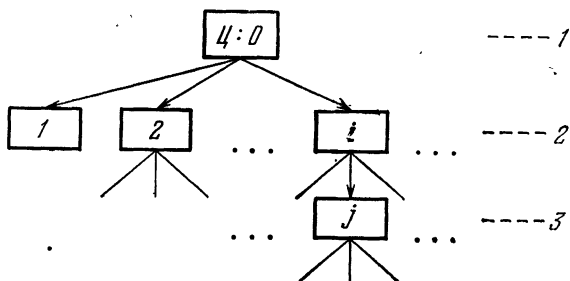


Рис. 2.

При описании системы употребляются следующие термины: *узел управления* u_i (на схеме ему соответствует номер i); *концевой узел* (на схеме таким является, например, узел u_i при $i=1$); *уровень управления*; *звено*; *цепь управления*.

Каждому узлу u_i соответствует множество π_i номеров узлов u_j , инцидентных узлу u_i , т. е. $\pi_i = \{j: i \rightarrow j\}$. Если $i \neq k$, то $\pi_i \cap \pi_k = \emptyset$. Концевому узлу u_i соответствует $\pi_i = \emptyset$. Пара (i, j) составляет звено управления, если $j \in \pi_i$; в этом случае будем писать $i = \pi_j^{-1}$. Последовательность звеньев $(i_1, i_2), (i_2, i_3), \dots, (i_k, i_{k+1})$ назы-

вается цепью управления длины k . Если максимальная из длин цепей равна $n-1$, то n называется высотой иерархии системы \mathfrak{M} (а сама система — n -уровневой). Заметим, что уровню $k \in \mathbb{N}_n$ принадлежат те и только те узлы управления, которые соединяются с узлом $u_0 = \Pi$ n -звенной цепью.

4.2. Модель узла управления и цель управления системой. В качестве отдельных элементов системы \mathfrak{M} выступают узлы управления u_i с собственными целями. С каждым узлом u_i свяжем математическую программу, моделирующую задачу оптимизации вектора x_i эндогенных переменных при фиксированном векторе y_i экзогенных параметров, задаваемых узлом управления u_i при $l = \pi_i^{-1}$.

Рассмотрим случай конечного узла, которому поставим в соответствие программу

$$\max \{f_i(x_i): g_i(x_i, y_i^t) \leq 0\} = \tilde{f}_i(y_i^t), \quad (1.4.1)$$

где $f_i(x_i)$ — скалярная функция векторного аргумента x_i ; $g_i(x_i, y_i)$ — векторная функция векторных аргументов x_i и y_i ; \max в (1.4.1) берется по x_i ; y_i^t — фиксированное состояние вектора y_i . Пусть u_i — неконцевой узел. Ему поставим в соответствие программу

$$\max \{f_i(x_i, \{\tilde{f}_j(y_j), \forall j \in \pi_i\}): g_i(x_i, \{y_j, \forall j \in \pi_i\}, y_i^t) \leq 0\} = \tilde{f}_i(y_i^t). \quad (1.4.2)$$

Здесь \max берется по допустимым x_i и $\{y_j, \forall j \in \pi_i\}$, а векторный параметр y_i^t является фиксированным и задается узлом $l = \pi_i^{-1}$. При $i=0$ вектор y_0^t либо постоянен, либо диктуется внешними для \mathfrak{M} условиями (средой). Заметим, что функционал f_i в (1.4.2) от $\{y_j, \forall j \in \pi_i\}$ зависит неявным образом.

Цель управления системой \mathfrak{M} моделируется задачей (1.4.2) при $i=0$ (*задача Центра*). По смыслу самой постановки ее решение должно носить децентрализованный характер и основываться на организации некоторой итерационной процедуры обмена информацией, каждый шаг которой сопряжен с пересмотром параметров y_i^t и передачей некоторой информации как из u_i в u_j , так и из u_j в u_i , $j \in \pi_i$.

4.3. Двухуровневая модель. В двухуровневой модели задачей первого уровня (задачей Центра) выступает

$$\max \{f_0(\mathbf{x}_0, \{\tilde{f}_i(\mathbf{y}_i), \forall i \in N_m\}) : g_0(\mathbf{x}_0, \{\mathbf{y}_i, \forall i \in N_m\}) \leq 0\} = \tilde{f}, \quad (1.4.3)$$

а второй уровень моделируется задачами

$$\max \{f_i(\mathbf{x}_i) : g_i(\mathbf{x}_i, \mathbf{y}_i) \leq 0\} = \tilde{f}_i(\mathbf{y}_i) \quad (\forall i \in N_m). \quad (1.4.4)$$

В задаче (1.4.3) \max ищется по всем допустимым \mathbf{x}_0 и $\mathbf{y} = [\mathbf{y}_1, \dots, \mathbf{y}_m]$; $[\bar{\mathbf{x}}_0, \bar{\mathbf{y}}]$ допустим, если существуют $\{\bar{\mathbf{x}}_i\}$ такие, что вектор $[\bar{\mathbf{x}}_0, \bar{\mathbf{y}}]$ удовлетворяет ограничениям задачи (1.4.3), а $[\bar{\mathbf{x}}_i, \bar{\mathbf{y}}_i]$ — ограничениям задачи (1.4.4).

Рассмотрим пример и экономическую интерпретацию. Примем интерпретацию задачи ЛП, в соответствии с которой столбцы матрицы коэффициентов системы ограничений моделируют технологические способы, вектор свободных членов — наличие исходной системы ингредиентов (ресурсов в широком смысле), коэффициент в линейной целевой функции — оценку продукции, получающейся при применении соответствующего технологического способа с единичной интенсивностью. В аспекте такой интерпретации выпишем частную реализацию модели (1.4.3), (1.4.4):

$$\max \{f_0(\mathbf{x}_0, \{\tilde{f}_i(\mathbf{u}_i, \mathbf{v}_i), \forall i \in N_m\}) : D_0 \mathbf{x}_0 \leq \mathbf{d}_0, \mathbf{x}_0 \geq 0; B_i \mathbf{u}_i \leq \mathbf{v}_i, \sum_{i=1}^m \mathbf{v}_i \leq \bar{\mathbf{v}}, \mathbf{u}_i \geq 0, \mathbf{v}_i \geq 0, \forall i\}; \quad (1.4.5)$$

$$\max_{(\mathbf{x}_i)} \{f_i(\mathbf{x}_i) : A_i \mathbf{x}_i + B_i \mathbf{u}_i \leq \mathbf{v}_i, \mathbf{x}_i \geq 0, \forall i\} = \tilde{f}_i(\mathbf{u}_i, \mathbf{v}_i) \quad (i = 1, \dots, m). \quad (1.4.6)$$

Смысл задачи (1.4.5), (1.4.6) следующий: продукция i -го предприятия распадается на две группы, при этом план \mathbf{x}_i по первой группе определяется самим предприятием, а план \mathbf{u}_i по второй — диктуется Центром, который диктует и ресурсы \mathbf{v}_i . В модели (1.4.5) Центра участвуют соотношения, балансирующие ресурсы, общий объем которых определяется вектором $\bar{\mathbf{v}}$; кроме того, в модели фигурирует вектор \mathbf{x}_0 и ограничения на него, выражающие те или иные внутренние задачи (интересы) Центра.

Эта модель может быть полностью линейной, если, например, положить

$$f_i(\mathbf{x}_i) = (c_i, \mathbf{x}_i),$$

$$f_0(\mathbf{x}_0, \{\beta_j\}) = (c_0, \mathbf{x}_0) + \sum_{i=1}^m \alpha_i \beta_j, \quad \alpha_i > 0 \quad (\forall i \in N_m).$$

4.4. Процедура обмена информацией в двухуровневой модели. Сделаем следующие предположения относительно модели (1.4.3), (1.4.4): $\{g_i(\mathbf{x}_i, \mathbf{y}_i)\}$ выпуклы и дифференцируемы по $[\mathbf{x}_i, \mathbf{y}_i]$; $g_0(\mathbf{x}_0, \mathbf{y})$ выпукла и дифференцируема по $[\mathbf{x}_0, \mathbf{y}]$; $\{f_i(\mathbf{x}_i)\}$ вогнуты и дифференцируемы; $f_0(\mathbf{x}_0, \beta_1, \dots, \beta_m)$ вогнута и дифференцируема по $[\mathbf{x}_0, \beta_1, \dots, \beta_m]$. При таких условиях легко проверяется

Утверждение 1.4.1. Функции $\{\tilde{f}_i(\mathbf{y}_i)\}$ вогнуты; функция $f_0(\mathbf{x}_0, \{\tilde{f}_i(\mathbf{y}_i)\})$ вогнута по $[\mathbf{x}_0, \mathbf{y}] = [\mathbf{x}_0, \mathbf{y}_1, \dots, \mathbf{y}_m]$ в предположении *изотонности* $f_0(\mathbf{x}_0, \beta_1, \dots, \beta_m)$ по $\mathbf{z} = [\beta_1, \dots, \beta_m]$, т. е. из $\mathbf{z}_1 \geq \mathbf{z}_2$ следует $f_0(\mathbf{x}_0, \mathbf{z}_1) \geq f_0(\mathbf{x}_0, \mathbf{z}_2)$.

Это утверждение позволяет средствами выпуклого программирования строить итерационные процедуры для задачи (1.4.3), согласующиеся с той схемой обмена информацией в модели (1.4.3), (1.4.4), которая диктуется структурой рассматриваемой иерархии управления. Такие процедуры предполагают умение вычислять субградиенты (опорные функционалы) функции

$$\bar{f}_0(\mathbf{x}_0, \mathbf{y}) = f_0(\mathbf{x}_0, \{\tilde{f}_i(\mathbf{y}_i)\})$$

в произвольной точке $[\mathbf{x}_0^t, \mathbf{y}^t]$. Множество субградиентов произвольной вогнутой функции $g(\mathbf{x})$ (субдифференциал) в точке $\bar{\mathbf{x}} \in \mathbb{R}^n$ обозначается через $\partial g(\bar{\mathbf{x}})$, т. е.

$$\partial g(\bar{\mathbf{x}}) = \{\mathbf{h} \in \mathbb{R}^n: (\mathbf{h}, \mathbf{x} - \bar{\mathbf{x}}) \geq g(\mathbf{x}) - g(\bar{\mathbf{x}}), \forall \mathbf{x} \in \mathbb{R}^n\}.$$

Для выпуклой функции субдифференциал определяется аналогично, нужно только в определении заменить \geq на \leq .

В наших предположениях если $\mathbf{h}_i^t \in \partial \tilde{f}_i(\mathbf{y}_i^t)$, то

$$\mathbf{h}^t = \left[\frac{\partial f_0}{\partial \mathbf{x}_0}, \frac{\partial f_0}{\partial \beta_1} \mathbf{h}_1^t, \dots, \frac{\partial f_0}{\partial \beta_m} \mathbf{h}_m^t \right] \in \partial \bar{f}(\mathbf{x}_0^t, \mathbf{y}_1^t, \dots, \mathbf{y}_m^t)$$

при $[\mathbf{x}_0, \beta_1, \dots, \beta_m] = [\mathbf{x}_0^t, \tilde{f}_1(\mathbf{y}_1^t), \dots, \tilde{f}_m(\mathbf{y}_m^t)]$.

Итерационный оператор для задачи (1.4.3), перерабатывающий вектор $[\mathbf{x}_0^t, \mathbf{y}^t]$ в вектор $[\mathbf{x}_0^{t+1}, \mathbf{y}^{t+1}]$, обозначим

$$\Phi(\mathbf{x}_0^t, \mathbf{y}^t) = [\mathbf{x}_0^{t+1}, \mathbf{y}^{t+1}].$$

Если учесть, что h_i^t вырабатывается в результате решения i -й задачи (1.4.4), то схему движения информации в итерационной процедуре для модели (1.4.3), (1.4.4) можно изобразить с помощью рис. 3.

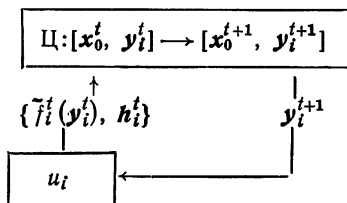


Рис. 3.

Выше оператор Φ был введен формально. Теперь мы дадим (на основе применения метода штрафных функций) один из способов конструктивного построения этого оператора. Для каждой из задач (1.4.4) введем множества

$$Q_i = \{\mathbf{y}_i: \exists \mathbf{x}_i \Rightarrow g_i(\mathbf{x}_i, \mathbf{y}_i) \leq 0\}$$

и

$$Q = Q_1 \times \dots \times Q_m.$$

Будем говорить, что задача (1.4.3) *равномерно регулярна* на Q , если:

- для произвольного $\bar{\mathbf{y}} = [\bar{\mathbf{y}}_1, \dots, \bar{\mathbf{y}}_m] \in Q$ задача (1.4.3) разрешима;
- функция Лагранжа

$$F(\mathbf{x}_0, \mathbf{u}_0) = \bar{f}(\mathbf{x}_0, \bar{\mathbf{y}}) - (\mathbf{u}_0, \mathbf{g}_0(\mathbf{x}_0, \bar{\mathbf{y}})),$$

отвечающая задаче (1.4.3), имеет седловую точку $[\bar{\mathbf{x}}_0, \bar{\mathbf{u}}_0(\bar{\mathbf{y}})]$;

$$\text{в) } \sup \{|\bar{\mathbf{u}}_0(\bar{\mathbf{y}})| : \bar{\mathbf{y}} \in Q\} = \delta < +\infty.$$

Если для задачи (1.4.3) выполнено условие равномерной регулярности, то при $\mathbf{r}_0 > [\delta, \dots, \delta]$ (вектор \mathbf{r}_0 имеет размерность вектора \mathbf{g}_0) и любом фиксированном $\mathbf{y} \in Q$ она эквивалентна задаче

$$\max \{\bar{F}(\mathbf{x}_0, \mathbf{y}) = f_0(\mathbf{x}_0, \{\bar{f}_i(\mathbf{y}_i)\}) - (\mathbf{r}_0, \mathbf{g}_0^+(\mathbf{x}_0, \mathbf{y}))\}$$

в смысле совпадения их оптимальных значений и оптимальных множеств [24]. Функция $\bar{F}(\mathbf{x}_0, \mathbf{y})$ является вогнутой по $[\mathbf{x}_0, \mathbf{y}]$. Если множество оптимальных векторов $[\tilde{\mathbf{x}}_0, \tilde{\mathbf{y}}]$ для (1.4.3) ограничено, то процесс

$$\{[\mathbf{x}_0^{t+1}, \mathbf{y}^{t+1}] = [\mathbf{x}_0^t, \mathbf{y}^t] + \alpha_t \bar{\mathbf{l}}_t\}_{\forall t},$$

где

$$\{\alpha_t > 0\} \rightarrow 0, \quad \sum_{t=1}^{+\infty} \alpha_t = +\infty,$$

$$\mathbf{l}_t \in \partial \bar{F}(\mathbf{x}_0^t, \mathbf{y}^t), \quad \bar{\mathbf{l}}_t = \mathbf{l}_t / |\mathbf{l}_t|,$$

сходится по функционалу \bar{F} к оптимуму, т.е. $\{\bar{F}(\mathbf{x}_0^t, \mathbf{y}^t)\} \rightarrow \bar{f}$ [42]. Вектор \mathbf{l}_t может быть представлен разностью $\mathbf{h}_t \in \partial \bar{f}_0(\mathbf{x}_0^t, \mathbf{y}^t)$ и $\mathbf{s}_t \in \partial (r_0, g_0^+(\mathbf{x}_0^t, \mathbf{y}^t))$. Вопрос о вычислении \mathbf{h}_t был уже решен, субградиент \mathbf{s}_t ищется по простым правилам субдифференцирования функций вида

$p(\mathbf{x}) = \sum_{i=1}^k \beta_i p_i^+(\mathbf{x})$, где $\beta_i \geq 0$ ($i = 1, \dots, k$), и $\{p_i(\mathbf{x})\}$ — выпуклые функции, а именно:

$$\partial p(\mathbf{x}) = \sum_{i=1}^k \beta_i \partial p_i^+(\mathbf{x}),$$

при этом

$$\partial p_i^+(\mathbf{x}) = \begin{cases} 0, & p_i(\mathbf{x}) < 0, \\ \partial p_i(\mathbf{x}), & p_i(\mathbf{x}) > 0, \\ \{\alpha \partial p_i(\mathbf{x}) : 0 \leq \alpha \leq 1\}, & p_i(\mathbf{x}) = 0. \end{cases}$$

Таким образом, нами построена для модели (1.4.3), (1.4.4) итерационная процедура, отвечающая определенной схеме обмена информацией, которая приводит к оптимальному (*равновесному*) состоянию этой двухуровневой модели управления.

Сделаем некоторые уточнения на случай модели (1.4.5), (1.4.6). Нами пока оставлен без внимания вопрос о субдифференцируемости функций оптимума $\tilde{f}_i(\mathbf{y}_i)$ в задачах второго уровня, т.е. вопрос о конструктивном нахождении элементов $\mathbf{h}_i \in \partial \tilde{f}_i(\mathbf{y}_i)$. В ряде случаев один из субградиентов \mathbf{h}_i может быть найден на основе знания решения двойственной к (1.4.4) задачи. Пусть, например,

$$g_i(\mathbf{x}_i, \mathbf{y}_i) = g_i(\mathbf{x}_i) - \mathbf{b}_i(\mathbf{y}_i)$$

и выполнены предположения п. 4.3. Решая задачу

$$\max_{(\mathbf{x}_i)} \{f_i(\mathbf{x}_i): g_i(\mathbf{x}_i) \leq \mathbf{b}_i(\mathbf{y}_i)\},$$

получаем вектор $\tilde{\mathbf{u}}_i$ двойственных оценок, характеризующих меру изменения $\tilde{f}_i(\mathbf{y}_i)$ в зависимости от вариаций $\mathbf{b}_i = \mathbf{b}_i(\mathbf{y}_i)$. Если $\tilde{f}_i(\mathbf{y}_i)$ рассматривать как функцию \mathbf{b}_i , т. е. $\tilde{f}_i(\mathbf{y}_i) = \tilde{f}_i(\mathbf{b}_i)$, то $\tilde{\mathbf{u}}_i \in \partial \tilde{f}_i(\mathbf{b}_i)$. Пусть $\tilde{\mathbf{u}}_i = \{\tilde{u}_i^k\}$, $\mathbf{y}_i = \{\mathbf{y}_i^s\}$, $\mathbf{b}_i(\mathbf{y}_i) = \{b_i^k(\mathbf{y}_i)\}$. Тогда, положив $\mathbf{h}_i^s = \sum_{(k)} \tilde{u}_i^k \frac{\partial b_i^k(\mathbf{y}_i)}{\partial y_i^s}$, будем иметь

$$\mathbf{h}_i = [h_i^1, \dots] \in \partial \tilde{f}_i(\mathbf{y}_i).$$

В частности, описанная схема получения \mathbf{h}_i пригодна для задач (1.4.6), в которых достаточно положить $\mathbf{b}_i(\mathbf{y}_i) = -B_i \mathbf{u}_i + \mathbf{v}_i$. Необходимые двойственные оценки формируются автоматически при решении задачи (1.4.6) при фиксированном $\mathbf{y}_i = [\mathbf{u}_i, \mathbf{v}_i]$, например модифицированным симплекс-методом (в предположении линейности $f_i(\mathbf{x}_i)$).

4.5. Пример применения в блочном программировании. Пусть дана задача линейного программирования блочной структуры

$$\max \left\{ \sum_{i=1}^m (c_i, \mathbf{x}_i): B_i \mathbf{x}_i = \mathbf{b}_i \quad (\forall i \in N_m), \right. \\ \left. \sum_{i=1}^m A_i \mathbf{x}_i = \mathbf{b}, \mathbf{x}_i \geq 0 \quad (\forall i \in N_m) \right\}. \quad (1.4.7)$$

Она допускает эквивалентное сведение к двухуровневой задаче оптимизации следующего вида. Выпишем программы 2-го и 1-го уровней соответственно:

$$\max \{ (c_i, \mathbf{x}_i): B_i \mathbf{x}_i = \mathbf{b}_i, A_i \mathbf{x}_i = \mathbf{y}_i, \mathbf{x}_i \geq 0 \} = \tilde{f}_i(\mathbf{y}_i) \\ (\forall i \in N_m), \\ \max \left\{ \sum_{i=1}^m \tilde{f}_i(\mathbf{y}_i): \sum_{i=1}^m \mathbf{y}_i = \mathbf{b} \right\}.$$

Такая редукция может быть положена в основу создания итерационного процесса для решения задачи (1.4.7) в соответствии со схемой, изложенной в п. 4.3.

4.6. Процедура обмена информацией в многоуровневой модели. Здесь мы рассмотрим лишь принципиальную схему обмена информацией в многоуровневой задаче оптимизации, отнеся ее описание к случаю трех уровней. Будем исходить из того, что с использованием приема штрафных функций программы всех узлов, за исключением конечных, записаны в форме математических программ без ограничений. Все предположения выпуклости, отнесенные к рассматриваемой задаче, будем считать выполненными.

Рассмотрим обмен информацией в трехуровневой модели. Выпишем задачи нулевого, 1-го и 2-го уровней соответственно:

$$\max f_0(\mathbf{x}_0, \{\tilde{f}_i(\mathbf{y}_i), i \in \pi_0\}, \{\mathbf{y}_i, i \in \pi_0\}) = \tilde{f}_0; \quad (1.4.8)$$

$$\max f_i(\mathbf{x}_i, \{\tilde{f}_j(\mathbf{y}_j), j \in \pi_i\}, \{\mathbf{y}_j, j \in \pi_i\}, \mathbf{y}_i) = \tilde{f}_i(\mathbf{y}_i); \quad (1.4.9)$$

$$\max f_j(\mathbf{x}_j, \mathbf{y}_j) = \tilde{f}_j(\mathbf{y}_j) \quad (j \in \pi_i, i \in \pi_0). \quad (1.4.10)$$

Интересующий нас итерационный процесс должен по схеме, близкой к описанной выше, перерабатывать систему значений

$$\{\mathbf{x}_k^i, \mathbf{y}_k^i, k \in \pi_0 \cup \pi_1 \cup \dots \cup \{0\}\}$$

из состояния t в состояние $t+1$. Введем краткое обозначение f_i^t для функций из задач (1.4.8) и (1.4.9) при $\mathbf{x}_i = \mathbf{x}_i^t$ ($i \in \{0\} \cup \pi_0$), $\mathbf{y}_j = \mathbf{y}_j^t$ ($j \in \pi_i$), $\mathbf{y}_i = \mathbf{y}_i^t$ ($i \in \pi_0$). Обозначим через ∂f_i^t множество субградиентов функции f_i^t по той или иной системе ее аргументов. Договоримся, что в явном виде будут перечисляться лишь те аргументы, по которым берется соответствующее множество. Так, например, $\partial f_i^t(\mathbf{y}_i^t)$ означает множество субградиентов функции f_i^t по \mathbf{y}_i в точке \mathbf{y}_i^t . Введем обозначения:

$$f_j^t = \tilde{f}_j(\mathbf{y}_j^t), \quad \mathbf{h}_j^t \in \partial \tilde{f}_j(\mathbf{y}_j^t) \quad (j \in \pi_i), \\ \mathbf{h}_i^t \in \partial f_i^t(\mathbf{y}_i^t) \quad (i \in \pi_0).$$

На рис. 4 изображена схема движения информации при реализации итерационного процесса.

К схеме сделаем следующие замечания. Поиск $\{h_j^i\}$ обсужден в конце п. 4.3. Итерационный шаг в узле u_i основывается на знании элементов $\{h_j^i\}$, $s_i^i \in \partial f_i^i(x_i)$ и

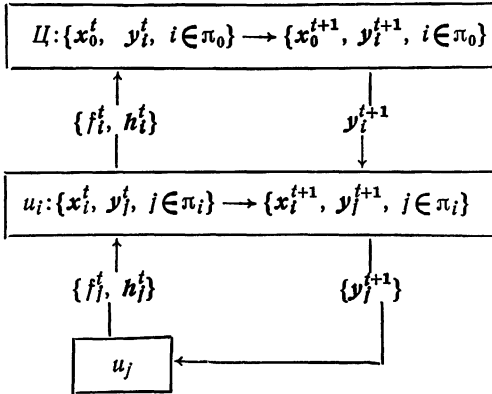


Рис. 4.

реализуется в соответствии с общей схемой субградиентного спуска (см. п. 4.3). Шаг, осуществляемый в u_0 (Центре), использует такую информацию:

$$s_0^i \in \partial f_0^i(x_0^i), \quad \bar{h}_i^i \in \partial f_0^i(y_i^i).$$

В предположениях выпуклости и равномерной регулярности можно доказать сходимость (по функционалу f_0) описанной итерационной процедуры к оптимуму задачи (1.4.8). В общем случае схема итерационной процедуры копирует схему трехуровневой модели, хотя при желании описать ее подробно нам понадобилось бы прибегнуть к громоздким обозначениям, выражениям и т. д., а также более тщательно дать некоторые определения (например, определение равномерной регулярности), выполнимость которых гарантирует сходимость процедуры. Изложение этих вопросов мы опускаем.

Отмеченными в § 3 гл. I свойствами итерационных процессов, которые мы положили в основу понимания итерационных методов вообще, характеризуются методы фейеровских приближений. Построение последних связано с использованием фейеровских отображений, выступающих в роли итерационных для задачи нахождения элемента из некоторого неявно заданного множества (например, заданного системой неравенств). Остановимся на этом вопросе подробнее.

Пусть имеется задача, состоящая в построении конструктивного способа выделения элемента из некоторого множества M (назовем его *допустимым*) или некоторого его подмножества \tilde{M} (назовем его *оптимальным* или *множеством эффективности*). Простой пример: пусть

$$\max \{(c, x) : Ax \leq b\}$$

— задача линейного программирования. Тогда в роли допустимого и эффективного множеств могут выступать

$$M = \{x : Ax \leq b\} \text{ и } \tilde{M} = \text{Arg max} \{(c, x) : x \in M\}.$$

Если же речь идет просто о решении системы линейных неравенств $Ax \leq b$, то можно положить $M = \tilde{M}$; однако в последнем случае можно выделить особое z -нормальное решение $\tilde{z} = \text{arg min} \{\|x - z\| : x \in M\}$ (z — некоторый фиксированный элемент из \mathbb{R}^n) и положить $\tilde{M} = \{\tilde{z}\}$. Выше элемент \tilde{z} , являющийся проекцией z на M , был определен однозначно.

Напомним ранее введенное понятие итерационного оператора (или отображения). Под *итерационным оператором* для задачи отыскания элемента $\bar{x} \in M$ ($\bar{x} \in \tilde{M}$) будем понимать такое отображение $\varphi \in \{\mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}\}$ (в общем случае точечно-множественное), которое при тех или иных ограничениях обеспечивает для последо-

вательности $\{x_t\}$, порожденной рекуррентно включением $x_{t+1} \in \varphi(x_t)$ при заданном начальном $x_0 \in \mathbb{R}^n$, свойство ограниченности последовательности $\{x_t\}$ и включение $\{x_t\}' \subset M$ ($\{x_t\}' \subset \bar{M}$); здесь $\{x_t\}'$ — множество пределов сходящихся подпоследовательностей последовательности $\{x_t\}$.

В дальнейшем понятие итерационного оператора будет соотнесено к случаю, когда роль M и \bar{M} будут играть множества соответственно допустимое и оптимальное для задачи выпуклого программирования

$$\max \{f_0(x) : f_j(x) \leq 0, \forall j \in N_m, x \in C\},$$

следовательно, здесь

$$M = \{x \in C : f_j(x) \leq 0; \forall j \in N_m\},$$

$$\bar{M} = \text{Arg max } \{f_0(x) : x \in M\}.$$

Требование выпуклости выписанной программы диктует выполнимость следующих свойств: функции $\{-f_0(x), f_1(x), \dots, f_m(x)\}$ выпуклы, множество C выпуклое, и замкнутое.

Если φ — итерационный оператор для задачи нахождения $\bar{x} \in M$, а ψ — для задачи $\max_{(x)} f_0(x)$, то конструирование итерационного оператора для

$$\max \{f_0(x_0) : x \in M\}$$

можно осуществлять путем композиции операторов φ и ψ в одной из форм $F_1(x) = \varphi(\psi(x))$ или $F_2(x) = \psi(\varphi(x))$; здесь под $\varphi(\psi(x))$ понимается отображение

$$x \rightarrow \bigcup_{z \in \psi(x)} \varphi(z).$$

В соответствии с этой схемой нами и будут развиваться итерационные методы решения задач выпуклого программирования в предположении, что отображение φ является фейеровским (методы фейеровских приближений).

§ 1. Фейеровские отображения и порождаемые ими последовательности

1.1. Определения и простейшие свойства. Пусть $\varphi \in \{\mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}\}$. Пусть $M(\varphi) = \{x : x = \varphi(x)\}$ — это множество неподвижных точек оператора φ .

Определение 2.1.1. Отображение φ называется *M-фейеровским*, если

$$|z - y| < |x - y|, \quad \forall z \in \varphi(x), \quad \forall y \in M(\varphi), \quad \forall x \notin M(\varphi).$$

Чтобы подчеркнуть связь фейеровского отображения с определяемым им *множеством неподвижности*, будем это множество называть также *M-фейеровским*.

В связи с определением 2.1.1 обратим внимание на следующий простой факт: если $\varphi(x)$ — однозначный расширяющий оператор из $\{\mathbb{R}^n \rightarrow \mathbb{R}^n\}$ с множеством неподвижности $M \neq \emptyset$, то оператор $\varphi_\alpha(x) = (1 - \alpha)\varphi(x) + \alpha x$ является *M-фейеровским* при любом $\alpha \in (0; 1)$.

Класс *M-фейеровских* отображений обозначим через F_M . Отметим легко проверяемое свойство: из $\varphi \in F_M$ вытекает выпуклость и замкнутость множества M .

Определение 2.1.2. Последовательность $\{x_t\}$ называется *фейеровской* относительно множества $Q \neq \emptyset$ (*Q-фейеровской*), если

$$|x_{t+1} - y| < |x_t - y| \quad (\forall y \in Q, \forall t \in \mathbb{N}).$$

Если Q — выпуклое, замкнутое множество из \mathbb{R}^n , то простейшим *Q-фейеровским* отображением $\varphi(x)$ является операция метрического проектирования на Q , т. е.

$$\varphi(x) = \arg \min \{ |x - z| : z \in Q \}.$$

Укажем на следующие простые свойства *Q-фейеровской* последовательности $\{x_t\}$.

Свойство 2.1.1. Если Q — выпуклое телесное множество (т. е. существует шар $S = \{x : |x - \bar{x}| \leq \varepsilon\}$ с центром в точке \bar{x} и радиусом $\varepsilon > 0$ такой, что $S \subset Q$), то $\{x_t\} \rightarrow x' \in \mathbb{R}^n$.

Свойство 2.1.2. Если $\{x_t\}$ — *Q-фейеровская*, то она будет и *со Q-фейеровской*; здесь $\text{со } Q$ — замыкание выпуклой оболочки $\text{со } Q$ множества Q .

Свойство 2.1.3. Если $\{x_t\}' \cap \overline{\text{со } Q} \neq \emptyset$, то $\{x_t\}$ сходится к точке из выпянутого пересечения.

Отметим, что если $\varphi \in F_M$, а $\{x_t\}$ — последовательность, порожденная отображением φ рекуррентно с помощью включения $x_{t+1} \in \varphi(x_t)$ при произвольном начальном $x_0 \in \mathbb{R}^n$, то из $\{x_t\} \cap M = \emptyset$ вытекает, что $\{x_t\}$ — *M-фейеровская*.

Определение 2.1.3. Отображение $\varphi \in \{\mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}\}$ называется *замкнутым*, если из того, что $\{x_t\} \rightarrow x'$, $\{y_t\} \rightarrow y'$, $y_t \in \varphi(x_t)$, $\forall t$, следует, что $y' \in \varphi(x')$.

Важным для дальнейшего примером замкнутого отображения является

$$x \rightarrow \partial f(x),$$

где $f(x)$ — определенная на \mathbb{R}^n выпуклая функция, $\partial f(x)$ — множество опорных к $f(x)$ в точке x линейных функционалов (см. [93]), т. е.

$$\partial f(x) = \{h \in \mathbb{R}^n: (h, y - x) \leq f(y) - f(x), \forall y \in \mathbb{R}^n\}.$$

Множество $\partial f(x)$ называют *субдифференциалом*, конкретный элемент из $\partial f(x)$ — *субградиентом*. Множество $\partial f(x)$ является непустым, выпуклым, замкнутым и ограниченным [91].

Теорема 2.1.1. Если отображение $\varphi \in F_M$ является замкнутым, то последовательность $\{x_t\}$, порожденная рекуррентно включением $x_{t+1} \in \varphi(x_t)$ при произвольном $x_0 \in \mathbb{R}^n$, сходится к $x' \in M$ (в более короткой записи: $\{x_{t+1} \in \varphi(x_t)\}_{\forall t} \rightarrow x' \in M$).

Доказательство. Следует убедиться в том, что в предположении $\{x_t\} \cap M = \emptyset$ последовательность $\{x_t\}$, являющаяся M -фейеровской, сходится к элементу $x' \in M$. Имея в виду ограниченность последовательности $\{x_t\}$, выделим подпоследовательность $\{x_{j_t}\} \rightarrow x'$ так, чтобы $\{x_{j_t+1}\} \rightarrow x''$. Если $x' \in M$, то по свойству 2.1.3 $\{x_t\} \rightarrow x'$, что и требовалось. Предположим $x' \notin M$, тогда в силу замкнутости φ будем иметь $x'' \in \varphi(x')$, а потому, взяв произвольный вектор $y \in M$, получим $|x'' - y| < |x' - y|$. Но так как, очевидно, $\{x_{j_t}\}' \subset \{x: |x - y| = \inf_{(t)} |x_{j_t} - y|\}$,

то таким образом получено противоречие. Теорема доказана.

Заметим, что может возникнуть как задача описания множества $M(\varphi)$ по заданному отображению $\varphi \in F_{M(\varphi)}$, так и задача построения отображения $\varphi \in F_M$ по заданному (например, системой неравенств) множеству $M \subset \mathbb{R}^n$. Вторая из них в применении к проблематике математического программирования является более актуальной. Именно ею мы и будем заниматься.

З а м е ч а н и е (к последующим формулировкам теорем о сходимости). Всюду, где не оговаривается выбор начального элемента x_0 итерационного процесса, предполагается, что он может быть произвольным.

1.2. Базовые конструкции фейеровских отображений. Пусть

$$P = \{x: l(x) = (h, x) - b \leq 0\}, \quad |h| \neq 0, \\ L_\lambda(x) = x - \lambda \frac{l^+(x)}{|h|^2} h, \quad 0 < \lambda < 2. \quad (2.1.1)$$

Параметр λ носит название *коэффициента релаксации*. Всюду в данной главе символами λ и λ_j будут обозначаться коэффициенты релаксации, поэтому мы не будем всякий раз оговаривать, что они принимают значения из интервала $\{\lambda: 0 < \lambda < 2\}$.

Лемма 2.1.1. *Отображение (2.1.1) является P-фейеровским.*

Это следует из того, что если $x \notin P$ и $y \in P$, то

$$|L_\lambda(x) - y|^2 = |x - y|^2 + 2\lambda \frac{l(x)l(y)}{|h|^2} - \lambda(2-\lambda) \frac{l^2(x)}{|h|^2}.$$

Заменив в (2.1.1) $L_\lambda(x)$ соотношением

$$L_\Delta(x) = \{L_\lambda(x): 0 < \lambda < 2\},$$

получим точно-множественное P-фейеровское отображение.

Обратим внимание на то обстоятельство, что верхнюю грань для значений λ в (2.1.1) увеличить нельзя, т. е. если положить $\lambda = 2 + \varepsilon$, $\varepsilon > 0$, то $L_\lambda(x)$ перестает быть P-фейеровским.

Пусть

$$\mu_f(x, h) = \begin{cases} \frac{f^+(x)}{|h|^2} h, & h \neq 0; \\ 0, & h = 0. \end{cases}$$

Лемма 2.1.2. *Если $f(x)$ — выпуклая функция и $M_f = \{x: f(x) \leq 0\} \neq \emptyset$, то отображение*

$$\varphi_f(x) = \{x - \lambda \mu_f(x, h): h \in \partial f(x)\} \quad (2.1.2)$$

является M_f -фейеровским ($0 < \lambda < 2$).

Это утверждение легко вытекает из предыдущей леммы с учетом того, что для $y \notin M_f$ справедливо очевидное

включение

$$\{x: (h, x-y) + f(y) \leq 0\} \supset M_f,$$

где $h \in \partial f(y)$.

Лемма 2.1.3. Если $\varphi_j \in F_{M_j}$ ($\forall j \in N_m$), то:

1) $\sum_{j=1}^m \alpha_j(x) \varphi_j(x) \in F_M$, где $M = \bigcap_{(j)} M_j \neq \emptyset$, $\{\alpha_j(x)\}$ — система неотрицательных функций, обладающих свойствами

$$\alpha_j(x) > 0 \quad (\forall x \notin M_j); \quad \sum_{j=1}^m \alpha_j(x) \leq 1 \quad (\forall x).$$

Кроме того, всюду ниже функции $\alpha_j(x)$ будут предполагаться непрерывными на R^n .

2) $\varphi_1(\dots\varphi_m(x)\dots) \in F_M$.

Проверяется непосредственно.

Следствие. Если степень F_M^2 определить как $\{\varphi(\psi(x)): \varphi \in F_M, \psi \in F_M\}$, то справедливо включение

$$\sum_{i=1}^m \alpha_i F_M^{n_i} \subset F_M,$$

каковы бы ни были натуральные n_i , m ($i \in N_m$); α_i такие, что $\sum_{i=1}^m \alpha_i = 1$, $\alpha_i \geq 0$ ($\forall i \in N_m$).

Пусть M — выпуклое замкнутое множество из R^n , $M \neq \emptyset$. Рассмотрим одну из важнейших конструкций M -фейеровского отображения безотносительно к способу конкретного задания множества M .

Определение 2.1.4. Пусть $d(x)$ — скалярная непрерывная функция, $h(x) \in \{R^n \rightarrow 2^{R^n}\}$; при этом

$$\{x: d(x) \leq 0\} = M, \quad h \neq 0 \quad (\forall h \in h(x), \forall x \notin M).$$

Пара $\{d(x), h(x)\}$ называется M -разделяющей, если для произвольных $p \notin M$ и $h \in h(p)$ полупространство, отвечающее неравенству $(h, x-p) + d(p) \leq 0$, содержит множество M .

Лемма 2.1.4. Если $\{d(x), h(x)\}$ — M -разделяющая пара, то отображение

$$\varphi(x) = \{x - \lambda \mu_d(x, h): h \in h(x)\}$$

является M -фейеровским, т. е. $\varphi \in F_M$, где

$$\mu_d(x, h) = \begin{cases} \frac{d^+(x)}{|h|^2} h, & h \neq 0, \\ 0, & h = 0. \end{cases}$$

Справедливость этого утверждения вытекает из леммы 2.1.1.

1.3. Применение базовых конструкций фейеровских отображений к случаю задания множества неподвижности системой выпуклых неравенств. Пусть

$$f_j(x) \leq 0 \quad (\forall j \in N_m) \quad (2.1.3)$$

— система выпуклых неравенств. Положим $M_j = \{x: f_j(x) \leq 0\}$, $M = \bigcap M_j$. Будем предполагать, что $M \neq \emptyset$.

Если $\varphi_j(x)$ — отображение (2.1.2), в котором $f = f_j$, $\lambda = \lambda_j$, $\mu_j(x, h) = \mu_{f_j}(x, h)$ ($= \mu_j(x, h)$), то каждое из отображений, сконструированных в лемме 2.1.3 из заданных $\{\varphi_j\}$, является M -фейеровским, где M — множество решений системы неравенств (2.1.3). Выпишем эти отображения:

$$\varphi^{(1)}(x) = \sum_{j=1}^m \alpha_j(x) \varphi_j(x), \quad (2.1.4)$$

$$\varphi^{(2)}(x) = \varphi_1(\dots \varphi_m(x)\dots), \quad (2.1.5)$$

где

$$\alpha_j(x) > 0 \quad (\forall j \notin M_j), \quad \sum_{j=1}^m \alpha_j(x) \leq 1 \quad (\forall x),$$

$$\varphi_j(x) = \{x - \lambda_j \mu_j(x, h): h \in \partial f_j(x)\}, \quad (2.1.6)$$

$$\mu_j(x, h) = \begin{cases} \frac{f_j^+(x)}{|h|^2} h, & h \neq 0; \\ 0, & h = 0. \end{cases} \quad (2.1.7)$$

Рассмотрим еще одну конструкцию M -фейеровского отображения, опирающуюся на лемму (2.1.4).

Пусть $g(y_1, \dots, y_m)$ — выпуклая на \mathbb{R}^m функция, обладающая следующими свойствами:

- 1) $g(0) = 0$;
- 2) $g(y) > 0 \quad (\forall y \in \mathbb{R}_+^m = \{y \geq 0\}), \quad y \neq 0$.

Примеры таких функций:

$$g_1(y) = \sum_{j=1}^m r_j y_j, \quad g_2(y) = \sum_{j=1}^m r_j y_j^2, \quad g_3(y) = \max_{(i)} r_j y_j$$

и т. д., где $r_j > 0$ ($\forall j \in N_m$).

Образую функцию $d(x) = g(f_1^+(x), \dots, f_m^+(x))$. Легко проверить, что она является выпуклой на R^n .

Выпишем отображение

$$\varphi^{(3)}(x) = \{x - \lambda \mu_d(x, h) : h \in \partial d(x)\}. \quad (2.1.8)$$

Пара $\{d(x), \partial d(x)\}$ удовлетворяет определению 2.1.4, а потому, согласно лемме 2.1.4, отображение (2.1.8) является M -фейеровским.

В связи с (2.1.8) целесообразно привести формулу

$$\partial d(x) = \bigcup_{y \in \partial g(y_x)} \sum_{i=1}^m y_i \partial f_i^+(x),$$

где

$$y = [y_1, \dots, y_m], \quad y_x = [f_1^+(x), \dots, f_m^+(x)],$$

$$\partial f_i^+(x) = \begin{cases} \partial f_i(x), & f_i(x) > 0, \\ \{0\}, & f_i(x) < 0, \\ \{\alpha \partial f_i(x) : 0 \leq \alpha \leq 1\}, & f_i(x) = 0. \end{cases}$$

Для случая, когда система (2.1.3) является линейной, т. е.

$$l_j(x) = (h_j, x) - b_j \leq 0 \quad (\forall j \in N_m), \quad (2.1.9)$$

в качестве примеров конкретных реализаций рассмотренных конструкций можно выписать следующие M -фейеровские отображения (M -множество решений системы 2.1.9):

$$L^{(1)}(x) = x - \frac{\lambda}{\delta} \sum_{j=1}^m l_j^+(x) h_j, \quad (2.1.10)$$

где $\delta = \sum_{j=1}^m |h_j|^2$;

$$L^{(2)}(x) = x - \lambda \frac{\max_{(l)} l_j^+(x)}{|h_{lx}|^2} h_{lx}, \quad (2.1.11)$$

где

$$j_x = \min \{j \in J(x)\}, \quad J(x) = \{i: \max_{(j)} l_j^+(x) = l_i^+(x)\};$$

$$L^{(3)}(x) = x - \lambda \frac{\sum_{j=1}^m [l_j^+(x)]^2}{\left| \sum_{j=1}^m l_j^+(x) h_j \right|^2} \sum_{j=1}^m l_j^+(x) h_j, \quad (2.1.12)$$

где по определению $L^{(3)}(x) = x$, если $s(x) = \{j: l_j(x) > 0\} = \emptyset$;

$$L^{(4)}(x) = \varphi_1(\dots \varphi_m(x)\dots), \quad (2.1.13)$$

где

$$\varphi_j(x) = x - \lambda_j \frac{l_j^+(x)}{|h_j|^2} h_j \quad (\forall j \in N_m);$$

$$L^{(5)}(x) = x - \lambda \frac{\sum_{j=1}^m l_j^+(x)}{\left| \sum_{j \in s(x)} h_j \right|^2} \sum_{j \in s(x)} h_j,$$

$$L^{(5)}(x) = x, \text{ если } s(x) = \emptyset. \quad (2.1.14)$$

Заметим, что отображения $L^{(1)}(x)$, $L^{(3)}(x)$ и $L^{(4)}(x)$ являются непрерывными. Поэтому в силу теоремы 1.1.1 процесс, порождаемый любым из них, сходится к некоторому решению системы (2.1.9).

З а м е ч а н и е. Если систему (2.1.3) заменить на

$$f_j(x) \leq 0 \quad (\forall j \in N_m), \quad x \in C, \quad (2.1.15)$$

где C — выпуклое замкнутое множество, то учет дополнительного ограничения $x \in C$ при конструировании фейеровского отображения относительно множества всех решений (2.1.15) может быть осуществлен весьма просто. Если φ — фейеровское отображение относительно $M = \{x: f_j(x) \leq 0, \forall j \in N_m\}$, полученное любым из базовых способов, то при произвольном $\psi \in F_C$ отображение $\psi(\varphi(x))$ будет $(M \cap C)$ -фейеровским; при этом, если φ и ψ замкнуты, то и их композиция замкнута.

Обычно при рассмотрении задачи математического программирования при ограничениях (2.1.15) в качестве ψ берут оператор $\pi_C(\cdot)$ метрического проектирования на C ,

а затем его композицию с итерационным оператором, соответствующим задаче оптимизации при ограничениях (2.1.15), однако без требования $\mathbf{x} \in C$. При этом исходят из предположения, что множество C достаточно просто устроено, например в смысле простоты выполнения проектирования на него произвольного элемента $\mathbf{x} = [x_1, \dots, x_n]$. В ряде случаев так и бывает. Если, скажем, $C = \mathbf{R}_+^n$, то проектирование на C действительно выполняется просто, а именно:

$$\pi_C(\mathbf{x}) = [x_1^+, \dots, x_n^+].$$

Но такая операция может оказаться и сложной, тогда есть смысл заменить ее на $\psi \in F_C$ с простой процедурой вычисления $\psi(\mathbf{x})$. Если C , например, задано некоторой системой выпуклых неравенств, то при построении ψ можно воспользоваться одной из простых конструкций C -фейеровского отображения.

Смысл разбиения системы ограничений на две подсистемы, вторая из которых задает множество C , может состоять в том, что первая подсистема допускает простую конструкцию итерационного оператора оптимизации $\tilde{\varphi}$. Тогда и для задачи в целом можно записать простой итерационный оператор $\tilde{\varphi}(\psi(\mathbf{x}))$, где $\psi \in F_C$.

§ 2. Применение к решению систем выпуклых неравенств

2.1. Замкнутость отображений $\varphi^{(i)}(\mathbf{x})$ ($i = 1, 2, 3$).

Лемма 2.2.1. Если $\{d(\mathbf{x}), h(\mathbf{x})\}$ — M -разделяющая пара, $h(\mathbf{x})$ замкнуто и ограничено (т. е. $\sup \{\|\mathbf{y}\| : \mathbf{y} \in h(\mathbf{x}), \mathbf{x} \in S\} < +\infty$ при любом ограниченном множестве $S \subset \mathbf{R}^n$), то отображение $\varphi(\mathbf{x})$ из леммы 2.1.4 замкнуто.

Доказательство. Для замкнутости отображения $\varphi^{(3)}(\mathbf{x})$, заданного согласно (2.1.8), достаточно доказать замкнутость отображения

$$H(\mathbf{x}) = \{\mu_d(\mathbf{x}, \mathbf{h}) : \mathbf{h} \in h(\mathbf{x})\},$$

где $\mu_d(\mathbf{x})$ задано соотношением вида (2.1.7), если вместо $f_j^+(\mathbf{x})$ подставить $d(\mathbf{x})$.

Заметим, что для величины $|\mu_d(\mathbf{x}, \mathbf{h})| = \frac{d(\mathbf{x})}{|\mathbf{h}|}$ ($\mathbf{x} \notin M$), являющейся расстоянием от элемента \mathbf{x} до гиперплоско-

сти с уравнением $(h, z-x) + d(x) = 0$ (z — свободная переменная), выполняется неравенство

$$|\mu_d(x, h)| \leq |x - y| \quad (\forall y \in M). \quad (2.2.1)$$

Итак, пусть $\{x_t\} \rightarrow x'$, $\{y_t\} \rightarrow y'$, при этом $y_t \in H(x_t)$ ($\forall t$). Нужно доказать, что $y' \in H(x')$.

Рассмотрим вначале случай $d(x') \leq 0$, т. е. $x' \in M$. Тогда, очевидно, $0 \in H(x')$. Если для некоторой подпоследовательности $\{x_{i_t}\}$ выполняется соотношение $d^+(x_{i_t}) = 0$ ($\forall t$), то $y' = 0 \in H(x')$. Если же почти для всех t (т. е. для всех t , начиная с некоторого t_0), будет $d(x_t) > 0$, то $|h_t| \neq 0$ ($\forall t \geq t_0$); здесь h_t определено условием $y_t = \frac{d(x_t)}{|h_t|^2} h_t$. Так как в силу (2.1.15)

$$|y_t| = |\mu_d(x_t, h_t)| \leq |x_t - x'| \rightarrow 0 \quad (t \rightarrow +\infty),$$

то $y' = 0 \in H(x')$. Таким образом, случай $d(x') \leq 0$ рассмотрен.

Пусть теперь $d(x') > 0$. Тогда, как легко видеть, $0 \notin h(x')$, что в силу ограниченности и замкнутости $h(x)$ ведет к $0 \notin \bigcup_{t \geq \bar{t}} h(x_t)$ при некотором \bar{t} . Это обстоятельство

позволяет считать, что $\{h_t\} \rightarrow h' \neq 0$, причем $h' \in h(x')$.

Но тогда $\{y_t\} \rightarrow y' = \frac{d(x')}{|h'|^2} h' \in H(x')$.

Лемма 2.2.2. Пусть $R^n \supset C$ ограничено, $f(x)$ выпукла на R^n ; тогда

$$\sup \{ |h| : h \in \partial f(x), x \in C \} < +\infty.$$

Доказательство. Пусть $\{x_t\} \subset C$, $h_t \in \partial f(x_t)$ и $\{|h_t|\} \rightarrow +\infty$. По свойству опорных функционалов имеем

$$(h_t, x - x_t) \leq f(x) - f(x_t) \quad (\forall x \in R^n).$$

Положив $x = x_t + \frac{h_t}{|h_t|}$, получим

$$|h_t| \leq f\left(x_t + \frac{h_t}{|h_t|}\right) - f(x_t) \leq \gamma < +\infty$$

(здесь использован факт непрерывности выпуклой функции $f(x)$ и, следовательно, ее ограниченности на любом ограниченном множестве).

Теорема 2.2.1. *Отображения $\varphi^{(i)}(\mathbf{x})$ ($i=1, 2, 3$), заданные согласно (2.1.4), (2.1.5) и (2.1.8), являются замкнутыми.*

Действительно, замкнутость $\varphi_j(\mathbf{x})$, заданного согласно (2.1.6), вытекает из леммы 2.2.1, если учесть, что пара $\{f_j(\mathbf{x}), \partial f_j(\mathbf{x})\}$ является M_j -разделяющей, а отображение $x \rightarrow \partial f_j(\mathbf{x})$ — замкнутым и ограниченным (см. лемму 2.2.1). Так как, кроме того, $\{\varphi_j(\mathbf{x})\}$ ограничены, а $\{\alpha_j(\mathbf{x})\}$ — непрерывны, то замкнутость $\varphi^{(1)}(\mathbf{x})$ отсюда вытекает очевидным образом. Замкнутость $\varphi^{(2)}(\mathbf{x})$ вытекает из замкнутости и ограниченности отображений $\{\varphi_j(\mathbf{x})\}$. Отображение $\varphi^{(3)}(\mathbf{x})$ является частным случаем отображения $\varphi(\mathbf{x})$ из леммы 2.1.5, поэтому оно также замкнуто.

2.2. Основная теорема (о сходимости). Из теорем 2.1.1 и 2.2.1 вытекает

Теорема 2.2.2. *Процесс*

$$\{\mathbf{x}_{t+1} \in \varphi^{(i)}(\mathbf{x}_t)\}_{\forall t},$$

порожденный любым из отображений $\varphi^{(i)}(\mathbf{x})$ ($i=1, 2, 3$) при любом начальном \mathbf{x}_0 , сходится к некоторому решению системы выпуклых неравенств (2.1.3) (кратко:

$$\{\mathbf{x}_{t+1} \in \varphi^{(i)}(\mathbf{x}_t)\}_{\forall t} \rightarrow \mathbf{x}' \in M).$$

Следствие. *Любой из процессов*

$$\{\mathbf{x}_{t+1} \in L^{(i)}(\mathbf{x}_t)\}_{\forall t} \quad (i=1, \dots, 5)$$

сходится к некоторому решению системы линейных неравенств (2.1.9); здесь $L^{(i)}(\mathbf{x})$ заданы согласно (2.1.10) — (2.1.14).

2.3. Фейеровские процессы для несовместных систем линейных неравенств. Некоторые варианты фейеровских процессов, отнесенных к случаю системы выпуклых неравенств, являются сходящимися и тогда, когда соответствующая система неравенств несовместна.

Рассмотрим некоторое обобщение отображения (2.1.10) на случай конечной системы линейных ограничений

$$l_j(\mathbf{x}) = 0 \quad (\forall j \in J_1), \quad l_j(\mathbf{x}) \leq 0 \quad (\forall j \in J_2), \quad (2.2.2)$$

а именно

$$L(\mathbf{x}) = \mathbf{x} - \frac{\lambda}{\delta} \left[\sum_{j \in J_1} r_j l_j(\mathbf{x}) \mathbf{h}_j + \sum_{j \in J_2} r_j l_j^+(\mathbf{x}) \mathbf{h}_j \right], \quad (2.2.3)$$

где

$$\delta = \sum_{j \in J_1 \cup J_2} r_j |\mathbf{h}_j|^2, \quad r_j > 0,$$

$$l_j(\mathbf{x}) = (\mathbf{h}_j, \mathbf{x}) - b_j \quad (\forall j \in J_1 \cup J_2).$$

Отображение $L(\mathbf{x})$ получается из

$$L_j(\mathbf{x}) = \mathbf{x} - \lambda \frac{l_j(\mathbf{x})}{|\mathbf{h}_j|^2} \mathbf{h}_j \quad (\forall j \in J_1),$$

$$L_j(\mathbf{x}) = \mathbf{x} - \lambda \frac{l_j^+(\mathbf{x})}{|\mathbf{h}_j|^2} \mathbf{h}_j \quad (\forall j \in J_2)$$

путем взятия их выпуклой комбинации с $\alpha_j = \frac{r_j |\mathbf{h}_j|^2}{\delta}$ ($\forall j \in N_m$). Для каждого из $L_j(\mathbf{x})$ выполняется неравенство

$$|L_j(\mathbf{x}) - L_j(\mathbf{y})| \leq |\mathbf{x} - \mathbf{y}|. \quad (2.2.4)$$

Действительно, пусть, например, $j \in J_2$ и $P_j = \{\mathbf{x} : l_j(\mathbf{x}) \leq 0\}$. Если один из элементов \mathbf{x}, \mathbf{y} лежит в P_j , то (2.2.4) справедливо в силу P_j -фейеровости $L_j(\mathbf{x})$ (см. лемму 2.1.1). Если же $\mathbf{x}, \mathbf{y} \notin P_j$, то (2.2.4) вытекает из непосредственно проверяемого тождества

$$|L_j(\mathbf{x}) - L_j(\mathbf{y})|^2 = |\mathbf{x} - \mathbf{y}|^2 - \frac{\lambda(2-\lambda)}{|\mathbf{h}_j|^2} [l_j(\mathbf{x}) - l_j(\mathbf{y})]^2.$$

Случай $j \in J_1$ аналогичен.

Так как $L(\mathbf{x}) = \sum_{j \in J_1 \cup J_2} \alpha_j L_j(\mathbf{x})$, то

$$|L(\mathbf{x}) - L(\mathbf{y})| \leq |\mathbf{x} - \mathbf{y}| \quad (\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n),$$

т. е. $L(\mathbf{x})$ — оператор нерастяжения.

Положим

$$f(\mathbf{x}) = \sum_{j \in J_1} r_j l_j^2(\mathbf{x}) + \sum_{j \in J_2} r_j [l_j^+(\mathbf{x})]^2.$$

Тогда, как легко видеть,

$$L(\mathbf{x}) = \mathbf{x} - \frac{\lambda}{2\delta} \nabla f(\mathbf{x}), \quad (2.2.5)$$

а потому

$$\{\mathbf{x} : L(\mathbf{x}) = \mathbf{x}\} = \{\mathbf{x} : \nabla f(\mathbf{x}) = \mathbf{0}\} = \underset{(x)}{\text{Arg min}} f(\mathbf{x}) \quad (= \tilde{M} \neq \emptyset).$$

Взяв $0 < \lambda_1 < 2$ и $0 < \alpha < 1$ такими, чтобы $\lambda_1 \alpha = \lambda$, будем иметь

$$L(\mathbf{x}) = (1-\alpha) \mathbf{x} + \alpha \left(\mathbf{x} - \frac{\lambda_1}{2\delta} \nabla f(\mathbf{x}) \right).$$

Отсюда, в силу замечания к определению фейеровского отображения (см. § 1 гл. II), получим

$$L(\cdot) \in F_{\bar{M}};$$

следовательно (см. теорему 2.1.1),

$$\{\mathbf{x}_{t+1} = L(\mathbf{x}_t)\}_{\forall t} \rightarrow \mathbf{x}' \in \bar{M}. \quad (2.2.6)$$

Естественно, если система (2.2.2) совместна и M —ее множество решений, то $M = \bar{M}$. Таким образом, процесс $\{\mathbf{x}_{t+1} \in L(\mathbf{x}_t)\}_{\forall t}$ выступает в качестве метода минимизации выпуклой функции вида

$$f(\mathbf{x}) = \sum_{j \in J_1} r_j l_j^2(\mathbf{x}) + \sum_{j \in J_2} r_j [l_j^+(\mathbf{x})]^2.$$

Этот метод мы обобщим на случай задачи

$$\min \{f(\mathbf{x}) : \mathbf{x} \in C\}, \quad (2.2.7)$$

где C —выпуклое, замкнутое множество из \mathbf{R}^n ($C \neq \emptyset$). Будем предполагать, что

$$\bar{M} = \text{Arg}(2.2.7) \neq \emptyset.$$

Пусть $\pi_C(\cdot)$ —оператор метрического проектирования на C (как известно, он является оператором нерастяжения). Образует оператор

$$\varphi(\mathbf{x}) = \pi_C(L(\mathbf{x})),$$

где $L(\mathbf{x})$ задано согласно (2.2.3).

Лемма 2.2.3. Справедливо равенство

$$\{\mathbf{x} : \varphi(\mathbf{x}) = \mathbf{x}\} = \text{Arg}(2.2.7).$$

Доказательство. Во-первых, заметим, что

$$\tilde{\mathbf{x}} \in \bar{M} \Leftrightarrow (\nabla f(\tilde{\mathbf{x}}), \mathbf{p} - \tilde{\mathbf{x}}) \geq 0 \quad (\forall \mathbf{p} \in C).$$

Далее, отбросим, как тривиальный, случай $\nabla f(\tilde{\mathbf{x}}) = \mathbf{0}$, $\tilde{\mathbf{x}} \in \bar{M}$. Пусть теперь $\tilde{\mathbf{x}} \in \bar{M}$, тогда гиперплоскость, отвечающая уравнению $(\nabla f(\tilde{\mathbf{x}}), \mathbf{x} - \tilde{\mathbf{x}}) = 0$, является опорной к C в точке $\tilde{\mathbf{x}}$, а потому проекция вектора $L(\tilde{\mathbf{x}}) = \tilde{\mathbf{x}} - \frac{\lambda}{2\delta} \nabla f(\tilde{\mathbf{x}}) \notin C$ (см. формулу (2.2.5)) на C совпадает

с \tilde{x} , но это и означает $\varphi(\tilde{x}) = \tilde{x}$. С другой стороны, пусть $\varphi(\tilde{x}) = \tilde{x}$, т. е. проекция вектора $L(\tilde{x}) = \tilde{x} - \frac{\lambda}{2\delta} \nabla f(\tilde{x})$ на C дает \tilde{x} . Но этот элемент является и проекцией $L(\tilde{x})$ на гиперплоскость с уравнением $(\nabla f(\tilde{x}), x - \tilde{x}) = 0$, что влечет $(\nabla f(\tilde{x}), p - \tilde{x}) \geq 0$ ($\forall p \in C$), т. е. $\tilde{x} \in \tilde{M}$. Лемма доказана.

Так как $\varphi(x)$ — оператор нерастяжения, то справедливо

С л е д с т в и е.

$$\varphi_\alpha(x) = (1 - \alpha)\varphi(x) + \alpha x \in F_{\tilde{M}}, \quad 0 < \alpha < 1,$$

поэтому

$$\{x_{t+1} = \varphi_\alpha(x_t)\}_{\forall t} \rightarrow x' \in \tilde{M}.$$

2.4. Распространение результатов на один нелинейный случай. Рассмотрим систему выпуклых неравенств (не обязательно совместную)

$$f_j(x) = |x - M_j|^2 \leq 0 \quad (\forall j \in N_m),$$

где $\{M_j\}$ — система выпуклых замкнутых множеств. Обозначим через $\pi_j(\cdot)$ оператор проектирования на M_j . Отметим, что $\nabla f_j(x) = 2(x - \pi_j(x))$. Отображение

$$\varphi_j(x) = x - \lambda \frac{f_j(x)}{|\nabla f_j(x)|^2} \nabla f_j(x)$$

является M_j -фейеровским. Взяв выпуклую комбинацию отображений $\{\varphi_j(x)\}$ при $\alpha_j(x) = |\nabla f_j(x)|^2 / \delta$, где

$\delta = \sum_{j=1}^m |\nabla f_j(x)|^2$, получим (после тривиальных преобразований)

$$\varphi(x) = x - \frac{\lambda}{2\delta} \nabla f(x),$$

где $f(x) = \sum_{j=1}^m f_j(x)$.

Отображение $\varphi(x)$ является оператором нерастяжения. Для доказательства достаточно проверить, что

$$|\varphi_j(x) - \varphi_j(y)| \leq |x - y| \quad \forall x, y \in \mathbb{R}^n, \quad (\forall j \in N_m).$$

Так как, очевидно,

$$\varphi_j(x) = x - \frac{\lambda}{2} (x - \pi_j(x)),$$

то

$$|\varphi_j(\mathbf{x}) - \varphi_j(\mathbf{y})| \leq (1 - \lambda/2) |\mathbf{x} - \mathbf{y}| + \frac{\lambda}{2} |\pi_j(\mathbf{x}) - \pi_j(\mathbf{y})| \leq |\mathbf{x} - \mathbf{y}|.$$

Точно так же, как и в предыдущем пункте, получаем

$$\varphi \in F_{\tilde{M}},$$

где $\tilde{M} = \underset{(x)}{\text{Arg min}} f(\mathbf{x})$. Поэтому

$$\{\mathbf{x}_{t+1} = \varphi(\mathbf{x}_t)\}_{\forall t} \rightarrow \mathbf{x}' \in \tilde{M}$$

(при $\tilde{M} \neq \emptyset$).

2.5. Вспомогательные результаты из теории штрафных функций. Рассмотрим разрешимую задачу выпуклого программирования

$$\min \{f_0(\mathbf{x}) : f_j(\mathbf{x}) \leq 0 \ (\forall j \in N_m), \mathbf{x} \in C\} (= \bar{f}); \quad (2.2.8)$$

в ней $\{f_i(\mathbf{x}) : i = 0, 1, \dots, m\}$ — выпуклые функции, C — выпуклое замкнутое множество. Введем обозначение $\tilde{M} = \text{Arg} (2.2.8) (\neq \emptyset)$. Пусть задача (2.2.8) удовлетворяет любой форме условий регулярности, гарантирующих для нее справедливость теоремы Куна — Таккера и, следовательно, обеспечивающих для $\tilde{\mathbf{x}} \in \tilde{M}$ существование $\mathbf{h}_i \in \partial f_i(\tilde{\mathbf{x}}) (\forall i \in N_m)$, $u_j \leq 0 (\forall j \in N_m)$ и \mathbf{h} таких, что

$$\mathbf{h}_0 = \sum_{j=1}^m u_j \mathbf{h}_j + \mathbf{h}, \quad u_j f_j(\tilde{\mathbf{x}}) = 0 \ (\forall j \in N_m), \quad (2.2.9)$$

при этом $(\mathbf{h}, \mathbf{x} - \tilde{\mathbf{x}}) \geq 0 (\forall \mathbf{x} \in C)$ [36].

Задаче (2.2.8) поставим в соответствие задачу

$$\inf \left\{ F(\mathbf{x}, r) = f_0(\mathbf{x}) + \sum_{j=1}^m r_j [f_j^+(\mathbf{x})]^2 : \mathbf{x} \in C \right\}, \quad (2.2.10)$$

$r_j > 0 (\forall j \in N_m)$. Будем предполагать, что $\tilde{M}(r) = \text{Arg} (2.2.10) \neq \emptyset$. Это выполняется, если, например, оптимальное множество \tilde{M} задачи (2.2.8) ограничено. Пусть $\tilde{\mathbf{x}} \in \tilde{M}(r)$, т. е. $\tilde{\mathbf{x}} \in C$ и $F(\tilde{\mathbf{x}}, r) = \min_{\mathbf{x} \in C} F(\mathbf{x}, r)$. Обозначим

$$|\mathbf{u}|_0 = \max_{(j)} |u_j|, \quad |\mathbf{u}|_1 = \sum_{(j)} |u_j|.$$

В сделанных предположениях справедлива Теорема 2.2.3. *Имеют место оценки*

$$1) \quad \bar{f} \geq F(\bar{\mathbf{x}}, r) \geq \bar{f} - \sum_{j=1}^m \frac{u_j^2}{4r_j}, \quad (2.2.11)$$

$$2) \quad d(\bar{\mathbf{x}}) = \max_{(i)} f_j^+(\bar{\mathbf{x}}) \leq \frac{1}{2r} (|\mathbf{u}| + |\mathbf{u}|_0), \quad (2.2.12)$$

$$3) \quad |f_0(\bar{\mathbf{x}}) - \bar{f}| \leq \frac{|\mathbf{u}_1|}{2r} (|\mathbf{u}| + |\mathbf{u}|_0), \quad (2.2.13)$$

где $r = \min_{(j)} r_j$.

Доказательство. 1. Так как $F(\bar{\mathbf{x}}, r) = \bar{f}$, то левое неравенство в (2.2.11) справедливо. Пусть $\mathbf{x} \in C$. Используя (2.2.9) и соотношение $(\mathbf{h}_0, \mathbf{x} - \bar{\mathbf{x}}) \leq f_0(\mathbf{x}) - f_0(\bar{\mathbf{x}})$, т. е. $f_0(\mathbf{x}) \geq (\mathbf{h}_0, \mathbf{x} - \bar{\mathbf{x}}) + \bar{f}$, будем иметь

$$\begin{aligned} F(\mathbf{x}, r) &\geq \bar{f} + \sum_{j=1}^m u_j (\mathbf{h}_j, \mathbf{x} - \bar{\mathbf{x}}) + (\mathbf{h}, \mathbf{x}) + \sum_{j=1}^m r_j [f_j^+(\mathbf{x})]^2 \geq \\ &\geq \bar{f} + \sum_{j=1}^m u_j [f_j(\mathbf{x}) - f_j(\bar{\mathbf{x}})] + \sum_{j=1}^m r_j [f_j^+(\mathbf{x})]^2 \geq \\ &\geq \bar{f} + \sum_{j=1}^m u_j f_j^+(\mathbf{x}) + \sum_{j=1}^m r_j [f_j^+(\mathbf{x})]^2 = \\ &= \bar{f} + \sum_{j=1}^m \left[r_j^{1/2} f_j^+(\mathbf{x}) + \frac{u_j}{2r_j} \right]^2 - \sum_{j=1}^m \frac{u_j^2}{4r_j} \geq \bar{f} - \sum_{j=1}^m \frac{u_j^2}{4r_j}, \end{aligned}$$

что и ведет к правому неравенству в (2.2.11).

2. Так как $\bar{f} \geq F(\bar{\mathbf{x}}, r)$, то с использованием соотношения $(\mathbf{h}_0, \bar{\mathbf{x}} - \bar{\mathbf{x}}) \geq \bar{f} - f_0(\bar{\mathbf{x}})$, получаем

$$\begin{aligned} \sum_{j=1}^m [f_j^+(\bar{\mathbf{x}})]^2 &\leq \frac{F(\bar{\mathbf{x}}, r) - f_0(\bar{\mathbf{x}})}{r} \leq \frac{\bar{f} - f_0(\bar{\mathbf{x}})}{r} \leq \\ &\leq \frac{1}{r} \left[\sum_{j=1}^m u_j (\mathbf{h}_j, \bar{\mathbf{x}} - \bar{\mathbf{x}}) + (\mathbf{h}, \bar{\mathbf{x}} - \bar{\mathbf{x}}) \right] \leq \\ &\leq \frac{1}{r} \sum_{j=1}^m |u_j| (\mathbf{h}_j, \bar{\mathbf{x}} - \bar{\mathbf{x}}) \leq \frac{1}{r} \sum_{j=1}^m |u_j| [f_j(\bar{\mathbf{x}}) - f_j(\bar{\mathbf{x}})] \leq \\ &\leq \frac{1}{r} \sum_{j=1}^m |u_j| f_j^+(\bar{\mathbf{x}}). \end{aligned}$$

Отсюда

$$\sum_{j=1}^m \left[f_j^+(\bar{x}) - \frac{|u_j|}{2r} \right]^2 \leq \sum_{j=1}^m \frac{u_j^2}{4r^2} = \frac{1}{4r^2} \|\mathbf{u}\|^2,$$

поэтому

$$\max_{(j)} \left| f_j^+(\bar{x}) - \frac{|u_j|}{2r} \right| \leq \frac{\|\mathbf{u}\|}{2r}.$$

Следовательно,

$$\max_{(j)} f_j^+(\bar{x}) \leq \frac{\|\mathbf{u}\|}{2r} + \frac{1}{2r} \max_{(j)} |u_j| = \frac{1}{2r} (\|\mathbf{u}\| + \|\mathbf{u}\|_0),$$

что и требовалось.

3. В выкладках п. 2 содержится неравенство

$$0 \leq \frac{\tilde{f} - f_0(\bar{x})}{r} \leq \frac{1}{r} \sum_{j=1}^m |u_j| f_j^+(\bar{x}),$$

которое с применением (2.2.12) и дает требуемое соотношение (2.2.13).

2.6. Применение процесса (2.2.6) для приближенного решения задачи линейного программирования. Рассмотрим задачу ЛП

$$\min \{(\mathbf{h}_0, \mathbf{x}) : (2.2.2), \mathbf{x} \geq 0\} (= \tilde{f}) \quad (2.2.14)$$

с допустимым множеством $M \neq \emptyset$.

Если взять число γ таким, что $\gamma + \tilde{f} > 0$, то задача (2.2.14) эквивалентна, очевидно, задаче

$$\min \{[(\mathbf{h}_0, \mathbf{x}) + \gamma]^2 : \mathbf{x} \in M\} (= \tilde{\gamma}), \quad (2.2.15)$$

где $\tilde{\gamma} = (\tilde{f} + \gamma)^2$. Последняя же асимптотически эквивалентна

$$\min_{\mathbf{x} \geq 0} \left\{ l_0^2(\mathbf{x}) + \sum_{j \in J_1} r_j l_j^2(\mathbf{x}) + \sum_{j \in J_2} r_j [l_j^+(\mathbf{x})]^2 \right\}, \quad (2.2.16)$$

где $l_0(\mathbf{x}) = (\mathbf{h}_0, \mathbf{x}) + \gamma$. Более точно это можно выразить так: если $\mathbf{x}(r) \in \text{Arg}(2.2.16)$, то

$$\begin{aligned} d(\mathbf{x}(r)) &= \max \{ |l_j(\mathbf{x}(r))|, j \in J_1; l_j^+(\mathbf{x}(r)), j \in J_2 \} \leq \\ &\leq \frac{1}{2r} (\|\mathbf{u}\| + \|\mathbf{u}\|_0), \end{aligned} \quad (2.2.17)$$

$$|l_0(\mathbf{x}(r)) - \tilde{\gamma}| \leq \frac{\|\mathbf{u}\|_1}{2r} (\|\mathbf{u}\| + \|\mathbf{u}\|_0), \quad (2.2.18)$$

где $r = \min r_j$, $\mathbf{u} = [u_1, \dots, u_m]$ — вектор оптимальных множителей ⁽ⁱ⁾ Лагранжа, отвечающих ограничениям $l_j(\mathbf{x}) = 0$ ($\forall j \in J_1$), $l_j(\mathbf{x}) \leq 0$ ($\forall j \in J_2$) в задаче (2.2.15). Выписанные неравенства (2.2.17) и (2.2.18) являются реализациями неравенств (2.2.12) и (2.2.13) на случай задачи (2.2.15).

Что же касается задачи (2.2.16), то для нее можно построить метод, аналогичный (2.2.6), а именно: пусть $f_0(\mathbf{x}, r)$ — оптимизируемая функция в задаче (2.2.16) и

$$\varphi_0(\mathbf{x}, r) = \left[\mathbf{x} - \frac{\lambda}{2\delta} \nabla_{\mathbf{x}} f_0(\mathbf{x}, r) \right]^+,$$

где $\delta = |\mathbf{h}_0|^2 + \sum_{j \in J_1 \cup J_2} r_j |\mathbf{h}_j|^2$, $+$ над вектором означает замену его отрицательных координат нулями (эта операция реализует метрическое проектирование вектора на положительный ортант $\mathbf{R}_+^n = \{\mathbf{x} \geq 0\}$ пространства \mathbf{R}^n). В силу основного результата п. 2.3

$$\{\mathbf{x}_{t+1} = \varphi(\mathbf{x}_t, r)\}_{\forall t} \rightarrow \mathbf{x}(r) \in \text{Arg}(2.2.16), \quad (2.2.19)$$

при этом, согласно (2.2.18),

$$[(\mathbf{h}_0, \mathbf{x}(r)) + \gamma]^2 \rightarrow [\tilde{f} + \gamma]^2 \quad (r \rightarrow +\infty),$$

что влечет $(\mathbf{h}_0, \mathbf{x}(r)) \rightarrow \tilde{f}$ ($r \rightarrow +\infty$). Из (2.2.17) вытекает, что $d(\mathbf{x}(r)) \rightarrow 0$ ($r \rightarrow +\infty$). Отсюда в силу неравенства (см. § 4 гл. II)

$$|\mathbf{x} - \tilde{M}| \leq K \max \{d(\mathbf{x}), |(\mathbf{h}_0, \mathbf{x}) - \tilde{f}|\} \quad (\forall \mathbf{x} \geq 0),$$

где K — некоторое положительное число, получаем

$$|\mathbf{x}(r) - \tilde{M}| \rightarrow 0 \quad (r \rightarrow +\infty),$$

где $\tilde{M} = \text{Arg}(2.2.14)$ — множество, задаваемое системой

$$l_j(\mathbf{x}) = 0 \quad (\forall j \in J_1), \quad l_j(\mathbf{x}) \leq 0 \quad (\forall j \in J_2), \quad (\mathbf{h}_0, \mathbf{x}) - \tilde{f} = 0.$$

Этим показано, что процесс (2.2.19) решает приближенно задачу (2.2.14), притом тем точнее, чем больше $r = \min r_j$. (i)

§ 3. Декомпозиционный метод построения фейеровских отображений

3.1. Учет диагональной структуры ограничений. Построение M -фейеровских отображений, соответствующих системе выпуклых неравенств

$$f_j(\mathbf{x}) \leq 0 \quad (\forall j \in N_m) \quad (2.3.1)$$

с множеством решений $M \neq \emptyset$, может осуществляться путем экономного учета особенностей структуры ее ограничений (выражающихся, например, в особенностях структуры матрицы коэффициентов в случае линейной системы). Пусть $N_m = \bigcup_{i=1}^s J_i$ — произвольное разбиение множества индексов системы (2.3.1); при этом пересечения $J_i \cap J_j$ не обязательно пусты при $i \neq j$. Если каждая из подсистем $f_j(x) \leq 0$ ($\forall j \in J_i$) с множеством решений M_i допускает простое построение (с точки зрения логики и арифметики реализации) M_i -фейеровского отображения φ_i (например, проектирование на M_i), то M -фейеровское отображение $(M = \bigcap_{i=1}^s M_i)$, соответствующее всей системе, может быть получено в силу любой из базовых конструкций, рассмотренных в § 1. Применим сказанное к системе вида

$$\begin{aligned} f_j(x^i) \leq 0 \quad (\forall j \in J_i) \quad (2.3.2)_i \quad (i = 1, \dots, s), \\ f_j(x^1, \dots, x^s) \leq 0 \quad (\forall j \in J_0) \quad (2.3.2)_0; \end{aligned} \quad (2.3.2)$$

здесь $x^i \in \mathbb{R}^{n_i}$, $\sum_{i=1}^s n_i = n$. Пусть M_i — множество решений системы (2.3.2)_i и $\varphi_i \in F_{M_i}$ ($i = 0, 1, \dots, s$). Тогда

$$\varphi(x) = \varphi_0(\varphi_1(x^1), \dots, \varphi_s(x^s)) \in F_M,$$

где M — множество решений системы (2.3.2). При этом, если $\{\varphi_i\}_{i=0}^s$ замкнуты, то φ также замкнуто; следовательно, любой порожденный им процесс сходится к решению системы (2.3.2).

3.2. Применение к одновременному решению прямой и двойственной задач линейного программирования. Выпишем для задачи ЛП

$$\max \{(h_0, x) : (h_j, x) - b_j \leq 0 \quad (\forall j \in N_m), x \geq 0\} \quad (2.3.3)$$

(форма записи которой для нас, естественно, безразлична) ей двойственную:

$$\min \left\{ (b, u) : \sum_{j=1}^m u_j h_j \geq h_0, u \geq 0 \right\}, \quad (2.3.3)^*$$

где $\mathbf{b} = [b_1, \dots, b_m]$, $\mathbf{u} = [u_1, \dots, u_m]$. Из теории ЛП известно, что вектор $[\tilde{\mathbf{x}}, \tilde{\mathbf{u}}] \in \mathbf{R}^{n+m}$ тогда и только тогда решает одновременно выписанную пару задач, когда он является решением системы линейных неравенств

$$(\mathbf{h}_j, \mathbf{x}) - b_j \leq 0 \quad (j = 1, \dots, m), \quad \mathbf{x} \geq 0; \quad (2.3.4)_1$$

$$\sum_{j=1}^m u_j \mathbf{h}_j \geq \mathbf{h}_0, \quad \mathbf{u} \geq 0, \quad (2.3.4)_2 \quad (2.3.4)$$

$$(\mathbf{h}_0, \mathbf{x}) - (\mathbf{b}, \mathbf{u}) = 0. \quad (2.3.4)_3$$

Поставим теперь системе $(2.3.4)_1$ в соответствие отображение $\varphi \in F_M$, а системе $(2.3.4)_2$ — $\varphi^* \in F_{M^*}$, где M — допустимое множество в задаче $(2.3.3)$, а M^* — в задаче $(2.3.3)^*$.

Запишем также отображение

$$\varphi_0(\mathbf{x}, \mathbf{u}) = [\mathbf{x}, \mathbf{u}] - \frac{(\mathbf{h}_0, \mathbf{x}) - (\mathbf{b}, \mathbf{u})}{|\mathbf{h}_0|^2 + |\mathbf{b}|^2} [\mathbf{h}_0, -\mathbf{b}], \quad (2.3.5)$$

реализующее операцию проектирования вектора $[\mathbf{x}, \mathbf{u}]$ на гиперплоскость в пространстве $\mathbf{R}^n \times \mathbf{R}^m$, заданную уравнением $(2.3.4)_3$. Композиция отображений φ , φ^* и φ_0 дает

$$F(\mathbf{x}, \mathbf{u}) = [\varphi(\mathbf{x}), \varphi^*(\mathbf{u})] + \alpha(\mathbf{x}, \mathbf{u}) [\mathbf{h}_0, -\mathbf{b}],$$

или

$$\psi(\mathbf{x}) = \varphi(\mathbf{x}) + \alpha(\mathbf{x}, \mathbf{u}) \mathbf{h}_0, \quad (2.3.6)$$

$$\psi^*(\mathbf{u}) = \varphi^*(\mathbf{u}) - \alpha(\mathbf{x}, \mathbf{u}) \mathbf{b}, \quad (2.3.7)$$

где $\alpha(\mathbf{x}, \mathbf{u}) = \frac{(\mathbf{b}, \mathbf{u}) - (\mathbf{h}_0, \mathbf{x})}{|\mathbf{h}_0|^2 + |\mathbf{b}|^2}$, $F = [\psi, \psi^*]$. Из сказанного в предыдущем пункте следует

$$F(\mathbf{x}, \mathbf{u}) \in F_{\tilde{M} \times \tilde{M}^*},$$

здесь $\tilde{M} = \text{Arg}(2.3.3)$, $\tilde{M}^* = \text{Arg}(2.3.3)^*$. Если $\varphi(\mathbf{x})$ и $\varphi^*(\mathbf{u})$ замкнуты, то отображение $F(\mathbf{x}, \mathbf{u})$ замкнуто; поэтому любой процесс, им порождаемый, сходится к $[\tilde{\mathbf{x}}, \tilde{\mathbf{u}}]$; при этом $\tilde{\mathbf{x}}$ будет оптимальным вектором задачи $(2.3.3)$, $\tilde{\mathbf{u}}$ — оптимальным вектором задачи $(2.3.3)^*$. Примерами непрерывных (а потому и замкнутых) отображений $\varphi(\mathbf{x})$

и $\varphi^*(u)$ могут служить (см. (2.1.10))

$$\varphi(x) = \left[x - \frac{\lambda_1}{\delta_1} \sum_{j=1}^m l_j^+(x) h_j \right]^+,$$

где $l_j(x) = (h_j, x) - b_j$, $\delta_1 = \sum_{j=1}^m |h_j|^2$, $0 < \lambda_1 < 2$;

$$\varphi^*(u) = \left[u - \frac{\lambda_2}{\delta_2} \sum_{i=1}^n a_i^+(u) a_i \right]^+,$$

где $a_i(u) = (a_i, u) + h_0^i$, $a_i = -[h_{i1}^i, \dots, h_{im}^i]$, $h_0 = [h_{01}^1, \dots, h_{0n}^n]$,

$h_j = [h_j^1, \dots, h_j^n]$ ($\forall i \in N_n, \forall j \in N_m$), $0 < \lambda_2 < 2$, $\delta_2 = \sum_{i=1}^n |a_i|^2$.

Замечание. Процесс $\{[x_{t+1}, u_{t+1}] = F(x_t, u_t)\} \forall t$ в силу (2.3.6) и (2.3.7) можно записать в виде

$$x_{t+1} = \varphi(x_t) + \alpha_t h_0, \quad (2.3.6)'$$

$$u_{t+1} = \varphi^*(u_t) - \alpha_t b, \quad (2.3.7)'$$

$\alpha_t = \alpha(x_t, u_t)$. Такого рода процесс называется сопряженным. Однако можно попытаться расщепить его на два самостоятельных, обеспечив каждый из них некоторым правилом генерирования шаговых коэффициентов α_t ($t=0, 1, \dots$). В § 6 будет показано, в частности, что если для $\{\alpha_t\}$ выполнены условия $\{\alpha_t \geq 0\} \rightarrow 0$ и $\sum_{(t)} \alpha_t = +\infty$, то процесс (2.3.6)' ограничен и

$$\{x_t\}' \subset \tilde{M}$$

(соответственно процесс (2.3.7)' ограничен и $\{u_t\}' \subset \tilde{M}^*$).

3.3. Применение к одновременному решению прямой и двойственной задач квадратичного программирования. Рассмотрим аналогичную конструкцию для выпуклой квадратичной программы

$$\max \{f(x) = -(x, Bx) + (c, x): Ax \leq b\}, \quad (2.3.8)$$

где B — неотрицательно определенная матрица, $A = \{a_{ji}\}_{m,n}$; $c, x \in R^n$, $b \in R^m$. Двойственной к ней будет [33]:

$$\min \{F(x, u) = f(x) - (u, Ax - b): \\ \nabla_x F(x, u) = 0, u \geq 0\}. \quad (2.3.8)'$$

Заметим, что $\nabla_x F(\mathbf{x}, \mathbf{u}) = -2B\mathbf{x} + \mathbf{c} - A^T\mathbf{u}$. Допустимые множества этих задач обозначим через $M_{\underline{u}}$ и M^* , оптимальные — через \tilde{M} и \tilde{M}^* . Из теории двойственности для задач выпуклого программирования известно [33], что:

1) если $\tilde{\mathbf{x}} \in \tilde{M}$, то при некотором $\tilde{\mathbf{u}} \geq 0$: $[\tilde{\mathbf{x}}, \tilde{\mathbf{u}}] \in \tilde{M}^*$, при этом $[\tilde{\mathbf{x}}, \tilde{\mathbf{u}}]$ удовлетворяет системе

$$A\mathbf{x} \leq \mathbf{b}, \nabla_x F(\mathbf{x}, \mathbf{u}) = 0, \mathbf{u} \geq 0; f(\mathbf{x}) \geq F(\mathbf{x}, \mathbf{u}); \quad (2.3.9)$$

2) если $[\tilde{\mathbf{x}}, \tilde{\mathbf{u}}]$ — произвольное решение системы (2.3.9), то $\tilde{\mathbf{x}} \in \tilde{M}$ и $[\tilde{\mathbf{x}}, \tilde{\mathbf{u}}] \in \tilde{M}^*$.

В развернутом виде система (2.3.9) записывается следующим образом:

$$\begin{aligned} A\mathbf{x} \leq \mathbf{b}, \quad -2B\mathbf{x} + \mathbf{c} - A^T\mathbf{u} = 0, \\ (\mathbf{u}, A\mathbf{x} - \mathbf{b}) \geq 0, \quad \mathbf{u} \geq 0. \end{aligned} \quad (2.3.10)$$

Если из первого соотношения найти $A^T\mathbf{u}$ и подставить во второе (переписав его в форме $(A^T\mathbf{u}, \mathbf{x}) - (\mathbf{b}, \mathbf{u}) \geq 0$), то (2.3.10) примет вид

$$\begin{aligned} A\mathbf{x} \leq \mathbf{b}, \quad -2B\mathbf{x} + \mathbf{c} - A^T\mathbf{u} = 0, \quad \mathbf{u} \geq 0; \\ 2(\mathbf{x}, B\mathbf{x}) - (\mathbf{c}, \mathbf{x}) + (\mathbf{b}, \mathbf{u}) \leq 0. \end{aligned}$$

Полученная система содержит линейные ограничения и одно выпуклое неравенство, следовательно, к ней можно применить уже рассмотренные способы построения \tilde{M}^* -фейеровских отображений, порождающих сходящиеся к $[\tilde{\mathbf{x}}, \tilde{\mathbf{u}}] \in \tilde{M}^*$ процессы.

Конкретные реализации таких построений многочисленны и останавливаться на них нет необходимости.

§ 4. Анализ скорости сходимости

Итерационные методы, понимаемые в смысле перечисленных ранее их характеристических черт (§ 3 гл. I), являются медленно сходящимися. Это связано с широтой их применимости, общностью и слабостью предположений, лежащих в основе теорем о сходимости соответствующих процессов.

Фейеровские процессы для негладких систем неравенств дают в лучшем случае сходимость со скоростью геометрической прогрессии. Последнее понимается в

следующем смысле: если $\varphi \in F_M$, то

$$|x_{t+1} - M| \leq \theta |x_t - M|, \quad \forall t \in \mathbb{N},$$

где $x_{t+1} \in \varphi(x_t)$, $0 < \theta < 1$.

Однако при практической реализации итерационных методов не обязательно ориентироваться на теоретическую скорость сходимости. Напротив, применение тех или иных мер управления процессом (в том числе эвристических, за счет управления константами метода и др.) не только может, но и должно вести к скорости сходимости, более высокой по сравнению с теоретической. Этот вопрос будет развит в § 3 гл. V.

4.1. Вспомогательные леммы. Пусть

$$f_j(x) \leq 0 \quad (\forall j \in N_m) \quad (2.4.1)$$

— система выпуклых неравенств с множеством решений $M \neq \emptyset$, S_0 — такое множество из \mathbb{R}^n , что $S_0 \cap M = \emptyset$ и

$$\{\pi_M(x) : x \in S_0\} \text{ ограничено,} \quad (2.4.2)$$

где $\pi_M(\cdot)$ — оператор метрического проектирования на M .

Если в (2.4.1) $f_j(x) = (h_j, x) - b_j$, т. е. неравенство является линейным, то будем его считать собственным, т. е. $h_j \neq 0$.

Для (2.4.1) будем предполагать выполненным условие регулярности в форме Слейтера

$$\exists \rho: f_j(\rho) < 0 \quad (\forall j \in N_m). \quad (R_0)$$

Обозначим $J(x) = \{j: f_j(x) = 0\}$; для краткости положим $\bar{x} = \pi_M(x)$.

Лемма 2.4.1. Если выполнено условие (2.4.2), то

$$\max_{(j)} \sup \{ \|h\| : h \in \partial f_j(\bar{x}), x \in S_0 \} = K < +\infty;$$

если, кроме того, выполнено условие (R_0) , то

$$\inf \{ \|h\| : h \in \partial f_j(\bar{x}), j \in J(\bar{x}), x \in S_0 \} = \delta > 0. \quad (2.4.3)$$

Доказательство. Первое из выписанных соотношений вытекает из леммы (2.2.2). Докажем второе. Предположив противное, можно выделить индекс $j' \in N_m$, последовательности $\{x_t\} \subset S_0$ и $\{h_t, h_t \in \partial f_{j'}(x_t)$ такие, что $\{x_t\} \rightarrow x'$ и $\{h_t\} \rightarrow h' = 0$. Отсюда $f_{j'}(x') = 0$ в силу непрерывности $f_{j'}(x)$, и $h' = 0 \in \partial f_{j'}(x')$ в силу замкнутости

отображения $x \rightarrow \partial f_{j'}(x)$. Но тогда $0 \leq f_{j'}(x') = \min_{(x)} f_{j'}(x)$, что противоречит условию (R_0) . Лемма доказана.

Лемма 2.4.2. Пусть выполнены условия (2.4.2) и (R_0) , тогда существует $\mu \in (0; 1)$ такое, что

$$\frac{\left(h_j, \sum_{J(\tilde{x}) \setminus j} u_k h_k \right)}{|h_j| \cdot \left| \sum_{J(\tilde{x}) \setminus j} u_k h_k \right|} \geq -\mu$$

при любых $h_j \in \partial f_j(\tilde{x})$, $j \in J(\tilde{x})$, $h_k \in \partial f_k(\tilde{x})$, $x \in S_0$ и $\{u_k \geq 0\}$ таких, что $\left| \sum_{J(\tilde{x}) \setminus j} u_k h_k \right| \neq 0$.

Доказательство. Заметим, что при условии $\left(\frac{a}{|a|}, \frac{b}{|b|} \right) = -1 + \varepsilon$ имеем $\left| \frac{a}{|a|} + \frac{b}{|b|} \right| = \sqrt{2\varepsilon}$. Предположим теперь, что лемма неверна. Тогда для любого $\varepsilon > 0$ найдутся $x \in S_0$ и $\{u_k\}$ такие, что для некоторого $j \in J(\tilde{x})$

$$\frac{(h_j, a)}{|h_j| \cdot |a|} = -1 + \varepsilon',$$

где $h_j \in \partial f_j(\tilde{x})$, $a = \sum_{J(\tilde{x}) \setminus j} u_k h_k$, $h_k \in \partial f_k(\tilde{x})$, $\varepsilon' \leq \varepsilon$. В силу сделанного замечания имеем

$$\frac{a}{|a|} = -\frac{h_j}{|h_j|} + a_{\varepsilon'}, \quad |a_{\varepsilon'}| = \sqrt{2\varepsilon'}. \quad (2.4.4)$$

Пусть p — вектор из условия (R_0) , K — число из леммы

2.4.1, тогда, положив $\delta_0 = -\frac{\max_{(i)} f_j(p)}{K}$, будем иметь

$$\begin{aligned} \left(\frac{a}{|a|}, p - \tilde{x} \right) &\leq \\ &\leq \frac{\sum_{J(\tilde{x}) \setminus j} u_k f_k(p)}{\left| \sum_{J(\tilde{x}) \setminus j} u_k h_k \right|} \leq \frac{\max_{(k)} f_k(p)}{K} \frac{\sum_{J(\tilde{x}) \setminus j} u_k}{\sum_{J(\tilde{x}) \setminus j} u_k} = -\delta_0 < 0. \end{aligned}$$

С другой стороны, согласно (2.4.4)

$$\begin{aligned} \left(\frac{a}{|a|}, p - \tilde{x} \right) &= -\left(\frac{h_j}{|h_j|} + a_{\varepsilon'}, p - \tilde{x} \right) \geq \\ &\geq -\frac{f_j(p)}{|h_j|} + (a_{\varepsilon'}, p - \tilde{x}) \geq (a_{\varepsilon'}, p - \tilde{x}), \end{aligned}$$

следовательно, $\delta_0 \leq (\alpha_{\varepsilon'}, \tilde{x} - p) \leq \sqrt{2\varepsilon'} K'$, где $K' = \sup \{|\tilde{x} - p| : x \in S_0\} < +\infty$, $\varepsilon' \leq \varepsilon$. Полученное неравенство $0 < \delta_0 \leq \sqrt{2\varepsilon'}$ порождает противоречие, так как ε' может быть как угодно малым, а δ_0 от ε' не зависит.

Лемма 2.4.3. Пусть $e = \sum_{i=1}^n u_i e_i$ и система векторов $\{e_i : \forall i \in N_n\}$ такова, что

$$\frac{(e_i, \sum_{k \neq i} u_k e_k)}{|e_i| \cdot \left| \sum_{k \neq i} u_k e_k \right|} \geq -\mu, \quad \forall i \in N_n,$$

где $u = [u_1, \dots, u_n] \geq 0$, $\left| \sum_{k \neq i} u_k e_k \right| \neq 0$, $\mu \in (0; 1)$. Если $|e_i| \geq \delta > 0$ ($\forall i \in N_n$), то

$$\max_{(i)} u_i \leq \frac{|e|}{\delta} \sqrt{\frac{1}{1-\mu}}.$$

Действительно, из

$$\begin{aligned} |e|^2 &= u_i^2 |e_i|^2 + 2u_i (e_i, \sum_{k \neq i} u_k e_k) + \left| \sum_{k \neq i} u_k e_k \right|^2 \geq \\ &\geq u_i^2 |e_i|^2 - 2u_i \mu |e_i| \left| \sum_{k \neq i} u_k e_k \right| + \left| \sum_{k \neq i} u_k e_k \right|^2 = \\ &= \left(\left| \sum_{k \neq i} u_k e_k \right| - u_i \mu |e_i| \right)^2 + u_i |e_i|^2 (1 - \mu^2) \end{aligned}$$

следует

$$u_i \leq \frac{|e|}{|e_i|} \sqrt{\frac{1}{1-\mu^2}}.$$

Поэтому лемма верна.

Пусть $d(x)$ — выпуклая функция, $M = \{x : d(x) \leq 0\} \neq \emptyset$ и $h(x) = \partial d(x)$, так что $\{d(x), h(x)\}$ — M -разделяющая пара (см. определение 2.1.4). Отображение

$$\varphi(x) = \{x - \lambda \mu(x, h) : h \in \partial d(x)\}, \quad (2.4.5)$$

где $0 < \lambda < 2$ и

$$\mu(x, h) = \begin{cases} \frac{d^+(x)}{|h|^2} h, & h \neq 0, \\ 0, & h = 0, \end{cases}$$

реализует одну из основных базовых конструкций замкнутого M -фейеровского отображения.

Лемма 2.4.4. Пусть $\mathbb{R}^n \supset S_0$ ограничено, $S_0 \cap M = \emptyset$ и существует $C > 0$ такое, что

$$|x - M| \leq C d(x), \quad \forall x \in S_0. \quad (2.4.6)$$

Тогда при некотором $\theta \in (0; 1)$

$$|z - M| \leq \theta |x - M|,$$

$\forall x \in S_0, \forall z \in \varphi(x)$, где $\varphi(x)$ задано согласно (2.4.5). Роль θ может играть число

$$\left(1 - \frac{2\lambda - \lambda^2}{C^2 N^2}\right)^{1/2},$$

где $N = \sup \{ |h| : h \in \partial d(x), x \in S_0 \}$.

Доказательство. Пусть $x \in S_0$, тогда $d(x) > 0$, но так как $\{d(x), h(x)\}$ — M -разделяющая пара, то $(h, y - x) + d(x) \leq 0, \forall y \in M$, поэтому

$$(h, x - \tilde{x}) \geq d(x) > 0,$$

$\forall h \in h(x) = \partial d(x), \tilde{x}$ — проекция x на M . Используя полученное неравенство, будем иметь для $z \in \varphi(x)$

$$\begin{aligned} |z - M|^2 &\leq |z - \tilde{x}|^2 = |x - \tilde{x}|^2 - \\ &\quad - 2\lambda (h, x - \tilde{x}) \frac{d(x)}{|h|^2} + \lambda^2 \frac{d^2(x)}{|h|^2} \leq \\ &\leq |x - \tilde{x}|^2 - (2\lambda - \lambda^2) \frac{d^2(x)}{|h|^2}. \end{aligned}$$

Отсюда

$$\frac{|z - M|^2}{|x - M|^2} \leq 1 - (2\lambda - \lambda^2) \frac{d^2(x)}{|h|^2} \frac{1}{|x - M|^2}. \quad (2.4.7)$$

Заметим, что так как $d(x)/|h|$ есть расстояние точки x до полупространства, отвечающего неравенству $(h, y - x) + d(x) \leq 0$ с переменным вектором y , а M в нем содержится, то

$$\frac{d(x)}{|h|} \leq |x - M|;$$

следовательно,

$$1 \geq \frac{d^2(x)}{|h|^2 |x - M|^2} \geq \frac{1}{|h|^2 C^2} \geq \frac{1}{N^2 C^2} > 0.$$

Отсюда и из (2.4.7) вытекает, что

$$\frac{|z-M|}{|x-M|} \leq \theta^2 = 1 - \frac{2\lambda - \lambda^2}{N^2 C^2};$$

при этом $\theta \in (0; 1)$, что и требовалось. |

З а м е ч а н и е. Если $d(x)$ — выпуклая кусочно-линейная функция, заданная, например, в форме $d(x) = \max_{(j)} l_j^+(x)$, $l_j(x) = (h_j, x) - b_j$ ($\forall j \in N_m$), то (2.4.6) имеет место при некотором $C > 0$ и $\forall x \in R^n$ [25] без каких-либо дополнительных предположений.

Пусть $|\cdot|_B$ — произвольная банахова норма в конечномерном пространстве B_m . Примерами таких норм являются:

$$|y|_0 = \max_{(j)} |y_j|, \quad |y|_1 = \sum_{j=1}^m |y_j|.$$

Исходя из произвольной нормы $|\cdot|_B$, построим функцию

$$d(x) = |[f_1^+(x), \dots, f_m^+(x)]|_B. \quad (2.4.8)$$

Так как любые две нормы $|\cdot|_{(1)}$ и $|\cdot|_{(2)}$ такого рода связаны соотношением $|\cdot|_{(1)} \leq C |\cdot|_{(2)}$ при некотором $C > 0$, то

$$\sum_{j=1}^m f_j^+(x) \leq C_1 d(x), \quad (2.4.9)$$

где $C_1 > 0$, а функция $d(x)$ задана согласно (2.4.8).

Лемма 2.4.5. Пусть для системы (2.4.1) выполнены условия леммы 2.4.2 и $\{\mu, \delta\}$ — величины, определенные в соответствии с (2.4.3) и леммой 2.4.2. Если $d(x)$ — функция (2.4.8), M — непустое множество решений системы (2.4.1), то

$$|x-M| \leq C d(x), \quad \forall x \in S_0, \quad (2.4.10)$$

где роль C может играть, например, число

$$\frac{C_1}{\delta \sqrt{\frac{1}{1-\mu^2}}},$$

C_1 — из (2.4.9).

Доказательство. Пусть $p \in S_0$, \tilde{p} — проекция p на M , т. е.

$$\tilde{p} = \arg \min \{ |x-p|^2 : x \in M \}.$$

Тогда в силу условия (R_0) и (2.2.9) будем иметь

$$-\nabla |x-p|^2 \Big|_{x=\tilde{p}} = -2(\tilde{p}-p) = \sum_{j=1}^m u_j^{(p)} h_j, \quad (2.4.11)$$

$$u_j^{(p)} f_j(\tilde{p}) = 0,$$

где $h_j \in \partial f_j(\tilde{p})$, $u_j^{(p)} \geq 0$ ($\forall j \in N_m$); при этом можно считать, что $\{h_j; u_j^{(p)} > 0\}$ линейно независимы. Согласно леммам 2.4.1—2.4.3 получаем

$$\max_{(j)} u_j^{(p)} \leq \frac{2|p-\tilde{p}|}{\delta} \sqrt{\frac{1}{1-\mu^2}}.$$

Теперь, используя (2.4.11), (2.4.9) и предыдущее соотношение, оценим $|p-M|^2$:

$$|p-M|^2 = |p-\tilde{p}|^2 = \frac{1}{2} \left(\sum_{j=1}^m u_j^{(p)} h_j, p-\tilde{p} \right) \leq$$

$$\leq \frac{|p-\tilde{p}|}{\delta} \sqrt{\frac{1}{1-\mu^2}} \sum_{j=1}^m f_j^*(p) \leq |p-\tilde{p}| \frac{C_1}{\delta} \sqrt{\frac{1}{1-\mu^2}} d(p).$$

Сокращая полученное неравенство на $|p-\tilde{p}| \neq 0$, получаем (2.4.10). Лемма доказана.

4.2. Оценка скорости сходимости фейеровского процесса, сконструированного с помощью M -разделяющей пары. Пусть M — множество решений совместной системы выпуклых неравенств (2.4.1), $d(x)$ — функция (2.4.8), $\varphi(x)$ — отображение (2.4.5), в котором

$$\mu(x, h) = \begin{cases} \frac{d(x)}{|h|^2} h, & h \neq 0, \\ 0, & h = 0. \end{cases}$$

Рассмотрим процесс

$$\{x_{t+1} \in \varphi(x_t)\} \forall t. \quad (2.4.12)$$

Теорема 2.4.1. Пусть система (2.4.1) удовлетворяет условию Слейтера. Тогда процесс (2.4.12) сходится к M по геометрической прогрессии, т. е. при некотором $\theta \in (0; 1)$:

$$|x_{t+1} - M| \leq \theta |x_t - M|, \quad \forall t. \quad (2.4.13)$$

Доказательство. В силу теоремы 2.2.2 процесс (2.4.12) сходится к элементу из M . Положим $S_0 = \{x_t\}$. Если $S_0 \cap M \neq \emptyset$, т. е. при некотором $t_0: x_{t_0} \in M$, то

$x_t = x_{t_0}$, $\forall t > t_0$, поэтому (2.4.13) справедливо при любом $\theta \in (0; 1)$. Если $S_0 \cap M = \emptyset$, то доказываемое соотношение (2.4.13) является итогом лемм 2.4.1—2.4.5, точнее, следует из лемм 2.4.5 и 2.4.4, при доказательстве которых в свою очередь были использованы предыдущие леммы 2.4.1—2.4.3. Теорема доказана.

Замечание. Если система (2.4.1) является линейной, то, как уже говорилось в замечании к лемме 2.4.4, существование константы C в неравенстве (2.4.10) обеспечивается и без условия Слейтера. Поэтому в силу той же леммы 2.4.4 гарантируется существование $\theta \in (0; 1)$ такого, что справедливо (2.4.13) для процесса (2.4.12), в котором конструкция отображения $\varphi(x)$ соответствует системе линейных неравенств $f_j(x) = (h_j, x) - b_j \leq 0$ ($j = 1, \dots, m$) с множеством решений $M \neq \emptyset$. Отсюда, в частности, следует, что процессы, порождаемые любым из отображений (2.1.11), (2.1.12), (2.1.4) сходятся к M со скоростью геометрической прогрессии.

4.3. Скорость сходимости фейеровского процесса, порожденного выпуклой комбинацией заданных M_j -фейеровских отображений. Речь будет идти о скорости сходимости процесса (2.4.12), порожденного многозначным отображением

$$\varphi(x) = \sum_{j=1}^m \alpha_j \varphi_j(x), \quad (2.4.14)$$

в котором

$$\alpha_j > 0 \quad (\forall j \in N_m), \quad \sum_{j=1}^m \alpha_j = 1;$$

$$\varphi_j(x) = \left\{ x - \lambda_j \frac{f_j^+(x)}{|h_j|^2} h_j : h_j \in \partial f_j(x) \right\} \quad (\forall j \in N_m).$$

В последнем соотношении второе слагаемое полагаем равным нулю при $h_j = 0$. Отображение $\varphi_j(x)$, как уже ранее отмечалось, является M_j -фейеровским, $\varphi(x)$ — M -фейеровским, здесь $M_j = \{x : f_j(x) \leq 0\}$, $M = \bigcap_{(j)} M_j \neq \emptyset$.

Заметим, что если для системы (2.4.1) выполнено условие (R_0) , то

$$|x - M| \leq C_0 \max_{(j)} f_j^+(x), \quad \forall x \in S_0 \quad [(2.4.15)]$$

при некотором $C_0 > 0$, здесь S_0 — ограничено и $M \cap S_0 = \emptyset$. Это соотношение является частным случаем неравенства (2.4.10).

Теорема 2.4.2. Пусть система (2.4.1) удовлетворяет условию (R_0) , S_0 ограничено и $M \cap S_0 = \emptyset$. Если, кроме того, для (2.4.14) выполнено требование $\{\alpha_j \lambda_j < \frac{2}{m}$ ($\forall j \in N_m$), то

$$|z - M| \leq \theta |x - M|, \quad \forall x \in S_0, \quad \forall z \in \varphi(x), \quad (2.4.16)$$

при некотором $\theta \in (0; 1)$. Отсюда, в частности, следует, что любой процесс, порожденный рассматриваемым отображением $\varphi(x)$, сходится к M со скоростью геометрической прогрессии.

При этом, если система (2.4.1) является линейной, то условие (R_0) является излишним, а в (2.4.16) можно считать $S_0 = \mathbb{R}^n$.

Доказательство. Итак, пусть $x \in S_0$, $z \in \varphi(x)$, \tilde{x} — проекция x на M . Для z имеем представление

$$z = x - \sum_{j=1}^m \alpha_j \lambda_j \frac{f_j^+(x)}{|h_j|^2} h_j$$

при некоторых $h_j \in \partial f_j(x)$ ($\forall j \in N_m$). Непосредственно проверяется, что

$$|z - M|^2 \leq |z - \tilde{x}|^2 = |x - M|^2 - 2 \sum_{j=1}^m \alpha_j \lambda_j \left[\frac{f_j^+(x)}{|h_j|} \right]^2 + \left| \sum_{j=1}^m \alpha_j \lambda_j \frac{f_j^+(x)}{|h_j|^2} h_j \right|^2. \quad (2.4.17)$$

Обозначив последнее слагаемое в (2.4.17) через γ , матрицу Грама системы векторов $\{h_j/|h_j|\}$ — через H (если $h_j = 0$, то полагаем $h_j/|h_j| = 0$), вектор

$$\left[\alpha_1 \lambda_1 \frac{f_1^+(x)}{|h_1|}, \dots, \alpha_m \lambda_m \frac{f_m^+(x)}{|h_m|} \right]$$

— через f , будем иметь

$$\gamma = f' H f \leq |f|^2 \|H\| \leq |f|^2 m = m \sum_{j=1}^m \alpha_j^2 \lambda_j^2 \left[\frac{f_j^+(x)}{|h_j|} \right]^2,$$

где $\|H\|$ — норма матрицы H , не превосходящая m . Отсюда в соответствии с (2.4.17) получаем

$$\begin{aligned} |z - M|^2 &\leq |x - M|^2 - \sum_{j=1}^m (2\alpha_j \lambda_j - m\alpha_j^2 \lambda_j^2) \left[\frac{f_j^+(x)}{|h_j|} \right]^2 \leq \\ &\leq |x - M|^2 - \theta_1 \sum_{j=1}^m \left[\frac{f_j^+(x)}{|h_j|} \right]^2, \end{aligned} \quad (2.4.18)$$

где

$$\theta_1 = \min_{(j)} (2\alpha_j \lambda_j - m\alpha_j^2 \lambda_j^2) = \min_{(j)} \left[\frac{1}{m} - m(\alpha_j \lambda_j - 1)^2 \right] < 1$$

(последнее соотношение допускает равенство $\theta_1 = 1$ лишь при $m = 1$ и $\lambda_j = 1$, но в этом случае доказываемое неравенство (2.4.16) является частным случаем теоремы 2.4.1). Из (2.4.18) следует

$$\begin{aligned} |z - M|^2 &\leq |x - M|^2 \left(1 - \frac{\theta_1}{|x - M|^2} \max_{(j)} \left[\frac{f_j^+(x)}{|h_j|} \right]^2 \right) \leq \\ &\leq |x - M|^2 (1 - \theta_1 \theta_2), \end{aligned}$$

где

$$\theta_2 = \inf_{x \in S_0} \max_{(j)} \left[\frac{f_j^+(x)}{|h_j| \cdot |x - M|} \right]^2;$$

при этом:

1) $\theta_2 \leq 1$, так как, очевидно,

$$|x - M| \geq \frac{f_j^+(x)}{|h_j|}, \quad \forall x \in S_0, \quad \forall j \in N_m;$$

2) $\theta_2 > 0$ в силу (2.4.15) и того, что

$$\sup \{ |h_j| : h_j \in \partial f_j(x), \quad j \in N_m, \quad x \in S_0 \} < +\infty$$

(см. лемму 2.2.2). Следовательно, в доказываемом соотношении (2.4.16) можно положить $\theta = (1 - \theta_1 \theta_2)^{1/2}$, при этом $\theta \in (0; 1)$.

Наконец, так как в проведенном доказательстве условие (R_0) было использовано лишь для того, чтобы обеспечить неравенство (2.4.15), а последнее для случая линейных $f_j(x)$ ($\forall j \in N_m$) справедливо без каких-либо предположений (см. замечание к теореме 2.4.1), то поэтому заключительная часть доказываемой теоремы также верна.

Следствие. Для отображения (2.1.10), а именно

$$L^1(x) = x - \frac{\lambda}{\delta} \sum_{j=1}^m l_j^1(x) h_j,$$

где $\lambda \in (0; 2)$, $\delta = \sum_{j=1}^m |h_j|^2$, справедливо при некотором $\theta \in (0; 1)$ неравенство

$$|L^{(1)}(x) - M| \leq \theta |x - M|, \quad \forall x \in \mathbb{R}^n, \quad (2.4.19)$$

здесь M — множество решений системы линейных неравенств (2.1.9), т. е. системы $(h_j, x) - b_j \leq 0$ ($\forall j \in N_m$).

Действительно, так как в этом случае $\lambda_j = \lambda$, $\alpha_j = |h_j|^2/\delta$ ($\forall j \in N_m$), то требование $0 < \alpha_j \lambda_j < 2/m$ переписывается в виде $0 < \lambda < 2\delta/m |h_j|^2$ ($\forall j \in N_m$), поэтому должно быть

$$\lambda < \min_{(j)} \frac{2\delta}{m |h_j|^2}.$$

Но так как

$$\min_{(j)} \frac{2\delta}{m |h_j|^2} = \frac{2\delta}{m (\max_{(j)} |h_j|^2)} \leq 2,$$

то условие $\alpha_j \lambda_j < 2/m$ выполняется автоматически.

Следовательно, к ситуации следствия приложима теорема 2.4.2.

Соотношение (2.4.19) дает, в частности, сходимость процесса $\{x_{t+1} = L^{(1)}(x_t)\}_{\forall t}$ к множеству M решений системы (2.1.9) со скоростью геометрической прогрессии.

4.4. Скорость сходимости фейеровского процесса, порожденного композицией M_j -фейеровских отображений. Пусть $\varphi_j(x)$ ($\forall j \in N_m$) — система M_j -фейеровских отображений, $M = \bigcap_{(j)} M_j \neq \emptyset$. Тогда, как уже отмечалось,

$$\varphi(x) = \varphi_m(\dots \varphi_1(x) \dots) \in F_M,$$

т. е. является M -фейеровским отображением. Пусть $S_0 \subset \mathbb{R}^n$, и для каждого $\varphi_j(x)$ выполнено соотношение

$$|z - M_j| \leq \theta_j |x - M_j|, \quad \forall x \in S_0, \quad \forall z \in \varphi_j(x), \quad (2.4.20)$$

где $\theta_j \in (0; 1)$ ($\forall j \in N_m$). Будем предполагать систему множеств $\{M_j\}$ такой, что обеспечивается существование

константы $C(J) > 0$, при которой

$$|x - \bigcap_{j \in J} M_j| \leq C(J) \max_{j \in J} |x - M_j|, \quad \forall x \in S_0, \quad (2.4.21)$$

для любого $J \subset N_m$. Если S_0 ограничено и $M = \bigcap_{(j)} M_j$ телесно, то (2.4.21) выполняется.

Образуем для $\alpha \in (0; 1)$

$$\varphi_j^\alpha(x) = (1 - \alpha) \varphi_j(x) + \alpha x \quad (\forall j \in N_m).$$

Очевидно, $\varphi_j^\alpha \in F_{M_j}$ ($\forall j \in N_m$).

Ниже в доказательстве леммы нам понадобится следующий простой факт.

Если $R^n \supset Q$ — выпуклое замкнутое множество и $\psi \in \{R^n \rightarrow 2^{R^n}\}$, то $\psi \in F_Q$ в том и только том случае, когда:

$$1) \psi(y) = y, \quad \forall y \in Q;$$

$$2) (z - x, y - x) > \frac{1}{2} |z - x|^2, \quad \forall y \in Q, \quad \forall x \notin Q,$$

$\forall z \in \psi(x)$.

Геометрически более очевидным является эквивалентное неравенство

$$\left(z - x, y - \frac{x+z}{2}\right) > 0.$$

Отсюда следует, что

$$(z - x, y - x) > \frac{1}{2(1-\alpha)} |z - x|^2, \quad (2.4.22)$$

$$\forall y \in M_j, \quad \forall x \notin M_j, \quad \forall z \in \varphi_j^\alpha(x).$$

Лемма 2.4.6. Из (2.4.20) при любых $\alpha \in (0; 1)$ и $z \in \varphi_j^\alpha(x)$ следует

$$|z - M_j| \leq \theta_j^\alpha |x - M_j|, \quad \forall x \in S_0, \quad \forall z \in \varphi_j^\alpha(x),$$

где

$$\theta_j^\alpha = [1 - \alpha(1 - \alpha)(1 - \theta_j)^2]^{1/2} \in (0; 1).$$

Доказательство. Непосредственно из (2.4.20) вытекает неравенство

$$|z - x| \geq (1 - \theta_j)(1 - \alpha) |x - M_j|. \quad (2.4.23)$$

Полагая $\tilde{x} = \pi_j(x)$ и учитывая (2.4.22), (2.4.23), будем иметь

$$\begin{aligned} |z - M_j|^2 &\leq |z - \tilde{x}|^2 = |x - M_j|^2 + |z - x|^2 - \\ &\quad - 2(z - x, \tilde{x} - x) \leq |x - M_j|^2 + |z - x|^2 - \\ &\quad - \frac{1}{1-\alpha} |z - x|^2 = |x - M_j|^2 - \frac{\alpha}{1-\alpha} |z - x|^2 \leq \\ &\leq |x - M_j|^2 - \frac{\alpha(1-\alpha)^2(1-\theta_j)^2}{(1-\alpha)} |x - M_j|^2 = (\theta_j^\alpha)^2 |x - M_j|^2, \end{aligned}$$

что и требовалось.

Положим $\varphi_\alpha(x) = \varphi_m^\alpha(\dots \varphi_1^\alpha(x)\dots)$.

Теорема 2.4.3. Пусть $\varphi_j \in F_{M_j}$ ($\forall j \in N_m$). Тогда из (2.4.20) и (2.4.21) следует, что при некотором $\theta \in (0; 1)$

$$|z - M| \leq \theta |x - M|, \quad \forall x \in S_0, \quad \forall z \in \varphi_\alpha(x).$$

Доказательство. Метод индукции позволяет здесь ограничиться доказательством * теоремы для случая $\varphi_\alpha(x) = \varphi_2^\alpha(\varphi_1^\alpha(x))$. Будем считать $S_0 \cap M = \emptyset$. Пусть $z \in \varphi_\alpha(x)$, тогда $z \in \varphi_2^\alpha(z_1)$, $z_1 \in \varphi_1^\alpha(x)$. Рассмотрим два случая.

1. Пусть $d(x) = \max_{i=1, 2} |x - M_i| = |x - M_1|$, тогда $x \notin M_1$. Поскольку

$$|x - M_1| \leq |z_1 - x| + |z_1 - M_1| \leq |z_1 - x| + \theta_1^\alpha |x - M_1|,$$

то

$$|z_1 - x| \geq (1 - \theta_1^\alpha) |x - M_1| \geq \frac{1 - \theta_1^\alpha}{C_2} |x - M|, \quad (2.4.24)$$

где C_2 — константа $C(J)$ из (2.4.21) при $J = \{1, 2\}$. Согласно (2.4.22)

$$(z_1 - x, y - x) > \frac{1}{2(1-\alpha)} |z_1 - x|^2, \quad \forall y \in M,$$

отсюда с применением (2.4.24) получаем

$$\begin{aligned} |z_1 - M|^2 &\leq |z_1 - \tilde{x}|^2 = |z_1 - x|^2 + |x - M|^2 - \\ &\quad - 2(z_1 - x, \tilde{x} - x) \leq |x - M|^2 - \frac{\alpha}{1-\alpha} |z_1 - x|^2 \leq \\ &\leq |x - M|^2 \left[1 - \frac{\alpha}{1-\alpha} \frac{(1 - \theta_1^\alpha)^2}{C_2^2} \right] = (\theta_1 |x - M|)^2, \end{aligned}$$

где θ_1 — квадратный корень из выражения, стоящего в квадратных скобках (очевидно, $\theta_1 \in (0; 1)$). Так как $|z - M| \leq |z_1 - M|$, то теорема для рассматриваемого случая верна.

2. Пусть $d(x) = |x - M_2|$. Рассмотрим два подслучая:

а) если $d(z_1) = |z_1 - M_2|$, то по аналогии с (2.4.24) будем иметь: так как

$$|z_1 - M_2| \leq |z - z_1| + |z - M_2| \leq |z - z_1| + \theta_2^\alpha |z_1 - M_2|,$$

то

$$|z - z_1| \geq (1 - \theta_2^\alpha) |z_1 - M_2| \geq \frac{1 - \theta_2^\alpha}{C_2} |z_1 - M|;$$

равным образом по аналогии с (2.4.22)

$$(z - z_1, y - z_1) > \frac{1}{2(1 - \alpha)} |z - z_1|^2, \quad \forall y \in M;$$

положив $\tilde{z} = \pi_M(z_1)$, отсюда получаем

$$\begin{aligned} |z - M|^2 &\leq |z - \tilde{z}|^2 = |z - z_1|^2 + |z_1 - M|^2 - \\ &- 2(z - z_1, \tilde{z} - z_1) \leq |z_1 - M|^2 - \frac{\alpha}{1 - \alpha} |z - z_1|^2 \leq \\ &\leq |z_1 - M|^2 - \frac{\alpha}{1 - \alpha} \cdot \frac{(1 - \theta_2^\alpha)^2}{C_2^2} |z_1 - M|^2 = \\ &= (\theta_2 |z_1 - M|)^2 \leq (\theta_2 |x - M|)^2, \end{aligned}$$

где $\theta_2 = \left[1 - \frac{\alpha}{1 - \alpha} \frac{(1 - \theta_2^\alpha)^2}{C_2^2} \right]^{1/2} \in (0; 1)$;

б) пусть теперь $d(z_1) = |z_1 - M_1|$, тогда

$$|z_1 - M_2| \leq |z_1 - M_1| \leq \theta_1^\alpha |x - M_1| \leq \theta_1^\alpha |x - M_2|;$$

отсюда

$$|x - M_2| \leq |z_1 - x| + |z_1 - M_2| \leq |z_1 - x| + \theta_1^\alpha |x - M_2|;$$

поэтому

$$|z_1 - x| \geq (1 - \theta_1^\alpha) |x - M_2| \geq \frac{1 - \theta_1^\alpha}{C_2} |x - M|. \quad (2.4.25)$$

По аналогии с п. 1, используя (2.4.25), получаем

$$|z_1 - M|^2 \leq \dots \leq |x - M|^2 \left[1 - \frac{\alpha}{1 - \alpha} \frac{(1 - \theta_1^\alpha)^2}{C_2^2} \right] = \\ = (\theta_1 |x - M|)^2,$$

где θ_1 имеет прежний смысл. Так как $|z - M| \leq |z_1 - M|$, то и в этом случае теорема верна.

Доказательство теоремы, таким образом, завершено.

З а м е ч а н и е 1. Пусть $\{M_j\}$ — собственные полупространства пространства R^n , т. е. $M_j = \{x: (h_j, x) - b_j \leq 0\}$, $h_j \neq 0$ ($\forall j \in N_m$). Тогда в силу замечания к лемме 2.4.4 неравенство (2.4.21) справедливо без условия телесности множества $M = \bigcap_{(j)} M_j$; при этом $C(j)$ не зависит от S_0 ;

более того, можно считать $S_0 = R^n$. Следовательно, для этого случая доказанная теорема принимает такую редакцию:

Если $\{M_j: \forall j \in N_m\}$ — собственные полупространства пространства R^n , $M = \bigcap_{(j)} M_j \neq \emptyset$ и выполнено (2.4.20), то при некотором $\theta \in (0; 1)$ справедливо неравенство

$$|z - M| \leq \theta |x - M|, \quad \forall x \in R^n, \quad \forall z \in \varphi_\alpha(x), \quad (2.4.26)$$

где $\varphi_\alpha(x) = \varphi_m^\alpha(\dots \varphi_1^\alpha(x) \dots)$.

Можно дать даже более сильную редакцию этого факта, а именно:

Если $\{M_j: \forall j \in N_m\}$ — выпуклые полиэдральные множества из R^n с непустым пересечением $M = \bigcap_{(j)} M_j$, для которых выполнено (2.4.20), то при некотором $\theta \in (0; 1)$ справедливо (2.4.26).

З а м е ч а н и е 2. Если $M_j = \{x: f_j(x) \leq 0\} \neq \emptyset$ и $f_j(x)$ — выпуклая функция ($\forall j \in N_m$), $\varphi_j[\lambda_j](x) = \varphi_j(x)$ из (2.4.14), то, представив число $\lambda_j \in (0; 2)$ в виде $\lambda_j = \alpha \lambda_j^\alpha$, где $\lambda_j^\alpha \in (0; 2)$, $\alpha \in (0; 1)$, будем иметь

$$\varphi_j[\lambda_j](x) = (1 - \alpha) \varphi_j[\lambda_j^\alpha](x) + \alpha x;$$

при этом $\varphi_j[\lambda_j^\alpha](\cdot) \in F_{M_j}$. Это показывает, что если для $z \in \varphi_j[\lambda_j^\alpha](x)$ выполнено соотношение (2.4.20), то оно в силу леммы 2.4.6 выполнено и для $z \in \varphi_j[\lambda_j](x)$ (вообще говоря, при другом значении θ_j). Но тогда в теореме 2.4.3 неравенство $|z - M| \leq \theta |x - M|$ справедливо и для

$$\forall z \in \varphi(x) = \varphi_m(\dots \varphi_1(x) \dots), \quad \forall x \in S_0,$$

т. е. для обеспечения написанного неравенства переход от $\{\varphi_j\}$ и φ к соответственно $\{\varphi_j^\alpha\}$ и φ_α не обязателен.

Сказанное позволяет сформулировать следующие утверждения.

Следствие 1. Пусть система выпуклых неравенств (2.4.1) удовлетворяет условию (R_0) , $\{\varphi_j(\mathbf{x}): \forall j \in N_m\}$ — отображения из (2.4.14). Тогда для любого ограниченного множества $S_0 \subset \mathbb{R}^n$ при некотором $\theta \in (0; 1)$ справедливо неравенство $|\mathbf{z} - M| \leq \theta |\mathbf{x} - M|$ для любых $\mathbf{x} \in S_0$ и $\mathbf{z} \in \Phi(\mathbf{x}) = \Phi_m(\dots \Phi_1(\mathbf{x})\dots)$. Следовательно, процесс

$$\{\mathbf{x}_{t+1} \in \Phi(\mathbf{x}_t)\} \forall t$$

сходится к M со скоростью геометрической прогрессии (при произвольном начальном $\mathbf{x}_0 \in \mathbb{R}^n$).

Следствие 2. Пусть система (2.4.1) — линейная, т. е. $f_j(\mathbf{x}) = (\mathbf{h}_j, \mathbf{x}) - b_j$, и $\mathbf{h}_j \neq \mathbf{0}$ ($\forall j \in N_m$). Если

$$L_j(\mathbf{x}) = \mathbf{x} - \lambda_j \frac{f_j^+(\mathbf{x})}{|\mathbf{h}_j|^2} \mathbf{h}_j, \quad \lambda_j \in (0; 2),$$

и

$$L(\mathbf{x}) = L_m(\dots L_1(\mathbf{x})\dots),$$

то при некотором $\theta \in (0; 1)$

$$|L(\mathbf{x}) - M| \leq \theta |\mathbf{x} - M|,$$

где M — непустое множество решений рассматриваемой системы.

Действительно, здесь следует только учесть, что, поскольку метрическая проекция $\tilde{\mathbf{x}}$ элемента \mathbf{x} на полупространство $M_j = \{\mathbf{x}: (\mathbf{h}_j, \mathbf{x}) - b_j \leq 0\}$ вычисляется по формуле

$$\tilde{\mathbf{x}} = \mathbf{x} - \frac{f_j^+(\mathbf{x})}{|\mathbf{h}_j|^2} \mathbf{h}_j,$$

причем

$$|\mathbf{x} - \tilde{\mathbf{x}}| = |\mathbf{x} - M_j| = f_j^+(\mathbf{x}) / |\mathbf{h}_j|,$$

то для

$$L_j[\lambda_j](\mathbf{x}) = \mathbf{x} - \lambda_j \frac{f_j^+(\mathbf{x})}{|\mathbf{h}_j|^2} \mathbf{h}_j$$

и любых $\lambda_j \in (0; 2)$ выполняется соотношение

$$\begin{aligned} |L_j[\lambda_j](\mathbf{x}) - M_j| &\leq |L_j[\lambda_j](\mathbf{x}) - \tilde{\mathbf{x}}| = \\ &= |\lambda_j - 1| \cdot |\mathbf{x} - M_j| = \theta_j |\mathbf{x} - M_j|, \end{aligned}$$

где $\theta_j = |\lambda_j - 1| \in (0; 1)$.

4.5. Скорость сходимости сопряженного фейеровского процесса для двойственной пары задач линейного программирования. Рассмотрим процесс (2.3.6)', (2.3.7)', порожденный отображением

$$F(\mathbf{x}, \mathbf{u}) = [\varphi(\mathbf{x}), \varphi^*(\mathbf{u})] + \alpha(\mathbf{x}, \mathbf{u}) [\mathbf{h}_0, -\mathbf{b}],$$

в котором

$$\begin{aligned} \varphi(\mathbf{x}) &= \left[\mathbf{x} - \frac{\lambda_1}{\delta_1} \sum_{j=1}^m l_j^+(\mathbf{x}) \mathbf{h}_j \right]^+ \in F_M, \\ \varphi^*(\mathbf{u}) &= \left[\mathbf{u} - \frac{\lambda_2}{\delta_2} \sum_{i=1}^n a_i^+(\mathbf{u}) \mathbf{a}_i \right]^+ \in F_{M^*}, \\ \alpha(\mathbf{x}, \mathbf{u}) &= \frac{(\mathbf{b}, \mathbf{u}) - (\mathbf{h}_0, \mathbf{x})}{|\mathbf{h}_0|^2 + |\mathbf{b}|^2} \end{aligned}$$

(см. п. 3.2 § 3 гл. II). Этот процесс сходится к $[\bar{\mathbf{x}}, \bar{\mathbf{u}}] \in \bar{M} \times \bar{M}^*$. Отображение $F(\mathbf{x}, \mathbf{u})$ строится из $\varphi(\mathbf{x})$, $\varphi^*(\mathbf{u})$ и

$$\varphi_0(\mathbf{x}, \mathbf{u}) = [\mathbf{x}, \mathbf{u}] - \alpha(\mathbf{x}, \mathbf{u}) [\mathbf{h}_0, -\mathbf{b}]$$

путем их композиции, при этом для каждого из них обеспечивается существование своего знаменателя прогрессии, характеризующего скорость сходимости порожденного им процесса. Достаточно разъяснить сказанное относительно отображения $\varphi(\mathbf{x})$. Оно представляет собой результат композиции отображений

$$\varphi'(\mathbf{x}) = \mathbf{x} - \frac{\lambda_1}{\delta_1} \sum_{j=1}^m l_j^+(\mathbf{x}) \mathbf{h}_j \quad \text{и} \quad \pi_i(\mathbf{x}) \quad (\forall i \in N_n),$$

где $\pi_i(\mathbf{x})$ — оператор проектирования на полупространство

$$\{\mathbf{x} = [x_1, \dots, x_n]: x_i \geq 0\},$$

так что $\varphi(\mathbf{x}) = \pi_m(\dots \pi_1(\varphi'(\mathbf{x}))\dots)$. В силу следствия из теоремы 2.4.2 и следствия 2 из теоремы 2.4.3 справедливо неравенство

$$|\varphi(\mathbf{x}) - M| \leq \theta_1 |\mathbf{x} - M|, \quad \forall \mathbf{x} \in \mathbb{R}^n, \quad \theta_1 \in (0; 1).$$

В силу тех же соображений

$$|\varphi^*(\mathbf{u}) - M^*| \leq \theta_2 |\mathbf{u} - M^*|, \quad \forall \mathbf{u} \in \mathbb{R}^m, \quad \theta_2 \in (0; 1).$$

Учитывая, наконец, замечание 2 к теореме 2.4.3, мы на основе самой теоремы 2.4.3 получаем

$$|F(x, u) - \bar{M} \times \bar{M}^*| \leq \theta |[x, u] - \bar{M} \times \bar{M}^*| \quad (2.4.27)$$

при некотором $\theta \in (0; 1)$, $\forall [x, u] \in \mathbb{R}^n \times \mathbb{R}^m$. Это показывает, что процесс

$$\{[x_{t+1}, u_{t+1}] = F(x_t, u_t)\}_{\forall t}$$

сходится к $\bar{M} \times \bar{M}^*$ по геометрической прогрессии со знаменателем θ из (2.4.27).

4.6. Вопросы помехоустойчивости фейеровских процессов. Пусть $\varphi \in \mathbf{F}_M$ и задан процесс (1.3.2), т. е.

$$\{x_{t+1} \in \varphi(x_t) + \delta_t\}_{\forall t}, \quad (2.4.28)$$

где $|\delta_t| \leq \varepsilon$, $\forall t$. Помехоустойчивость итерационного процесса была охарактеризована тремя свойствами (см. п. 3.1 § 2 гл. II), первое из которых для фейеровских процессов выполняется автоматически.

Прежде чем обсудить условия, при которых выполняются два других, коснемся вопроса об ограниченности процесса (2.4.28). В сущности (2.4.28)—это тип нестационарного процесса, условия ограниченности которого приводятся в § 2 гл. III. Одно из них мы здесь воспроизведем.

Условие регулярного сжатия на бесконечности: существуют $r > 0$ и $\delta > 0$ такие, что

$$|z - M| \leq |x - M| - \delta, \quad \forall x \notin M^r, \quad \forall z \in \varphi(x). \quad (2.4.29)$$

Лемма 2.4.7. Пусть $\varphi \in \mathbf{F}_M$, M ограничено, и выполнено условие (2.4.29). Тогда при $\varepsilon \leq \delta$ процесс (2.4.28) ограничен.

Доказательство. Если x_0 —начальный элемент процесса (2.4.28), то возьмем $R \geq \max\{r + \varepsilon, |x_0 - M|\}$. Докажем, что $\{x_t\} \subset M^R$. Очевидно, $x_0 \in M^R$. Пусть, уже доказано, что $x_t \in M^R$; докажем, что $x_{t+1} \in M^R$. Рассмотрим два случая:

1) пусть $x_t \in M^r$, тогда

$$|x_{t+1} - M| \leq |z_t - M| + \varepsilon \leq |x_t - M| + \varepsilon \leq r + \varepsilon \leq R,$$

т. е. $x_{t+1} \in M^R$; здесь z_t —некоторый элемент из $\varphi(x_t)$;

2) пусть $x_t \notin M^r$, тогда

$$|z_t - M| \leq |x_t - M| - \delta,$$

поэтому

$$|x_{t+1} - M| \leq |x_t - M| - \delta + r \leq R - \delta + r \leq R.$$

Лемма доказана.

Лемма 2.4.8. Если M -фейеровское отображение φ обладает свойством

$$|z - M| \leq \theta |x - M|, \quad \forall z \in \varphi(x), \quad \forall x, \quad (2.4.30)$$

где $\theta \in (0; 1)$, то при ограниченном M процесс (2.4.28) ограничен.

Действительно, взяв $R \geq \max \left\{ \frac{\varepsilon}{1-\theta}, |x_0 - M| \right\}$, будем иметь: если $x_t \in M^R$, то

$$|x_{t+1} - M| \leq \theta |x_t - M| + \varepsilon \leq \theta R + \varepsilon \leq R.$$

Следовательно, $\{x_t\} \subset M^R$, что и требовалось.

Теорема 2.4.4. Пусть $F_M \ni \varphi$ замкнуто, $\{\delta_t\} \rightarrow 0$ и процесс (2.4.28) ограничен. Тогда

$$\{x_t\}' \subset M. \quad (2.4.31)$$

Доказательство. Пусть x' — наиболее удаленная от M точка из $\{x_t\}'$. Выделим последовательность $\{x_{i_t}\} \rightarrow x'$ такую, что $x_{i_t+1} = z_t + \delta_{i_t}$ и $\{z_t\} \rightarrow z'$, $\{x_{i_t+1}\} \rightarrow x''$; здесь $z_t \in \varphi(x_{i_t})$. В силу замкнутости $\varphi(x)$ отсюда получаем, что $x'' = z' \in \varphi(x')$. Если $x' \in M$, то по правилу выбора точки x' получаем (2.4.31). Пусть $x' \notin M$. Тогда $|x'' - y| < |x' - y|$, $y \in M$, что противоречит принципу выбора x' . Теорема доказана.

Следствие 1. Пусть $F_M \ni \varphi$ замкнуто, M ограничено и выполнено условие (2.4.29). Если $\{\delta_t\} \rightarrow 0$, то при достаточно малом $\varepsilon_0 = \sup_{(t)} |\delta_t|$ процесс (2.4.8) ограничен и выполнено соотношение (2.4.31).

Следствие 2. Пусть $F_M \ni \varphi$ замкнуто, M ограничено и выполнено условие (2.4.30). Если $\{\delta_t\} \rightarrow 0$, то процесс (2.4.8) ограничен и выполнено соотношение (2.4.31).

Теорема 2.4.5. Пусть $F_M \ni \varphi$ замкнуто, M ограничено и выполнено условие (2.4.30). Если $|\delta_t| \leq \varepsilon$, $\forall t$, то

процесс (2.4.28) ограничен, при этом

$$\sup_{\bar{x} \in \{x_t\}'} |\bar{x} - M| \leq \frac{\varepsilon}{1-\theta}. \quad (2.4.32)$$

Доказательство. Ограниченность процесса имеет место в силу следствия 2 из теоремы 2.4.4. Применив схему доказательства последней и поменяв в ней роли элементов x' и x'' (т. е. x'' будет наиболее удаленной от M точкой из $\{x_t\}'$), получим $|x'' - y| \leq |z' - M| + \varepsilon$, какой бы ни была точка $y \in M$; здесь $z' \in \varphi(x')$. Отсюда $|x'' - M| \leq |z' - M| + \varepsilon \leq \theta|x' - M| + \varepsilon \leq \theta|x'' - M| + \varepsilon$, т. е. $|x'' - M| \leq \frac{\varepsilon}{1-\theta}$, что и ведет к (2.4.32).

Замечание 1. Можно рассматривать и такой вариант фейеровского процесса с помехами:

$$\{x_{t+1} \in \varphi(x_t + \delta_t)\}_{\forall t}. \quad (2.4.33)$$

Все утверждения этого пункта остаются верными и для процесса (2.4.33); при этом их обоснования требуют лишь тривиальных модификаций: в доказательстве леммы (2.4.8) надо взять $R \geq \frac{\theta\varepsilon}{1-\theta}$; при доказательстве теоремы (2.4.4) соотношение $x_{i+1} = z_i + \delta_{i_t}$ заменить на $z_i = x_{i_t} + \delta_{i_t}$, $\varphi(z_i) \ni x_{i+1}$.

Замечание 2. Если $\delta_t = \alpha_t c$, $c \neq 0$, $\alpha_t \geq 0$, $\forall t$, $\sum_{i=1}^{+\infty} \alpha_i = +\infty$, то при условиях теоремы (2.4.4) имеет место включение

$$\{x_t\}' \subset \text{Arg max } \{(c, x) : x \in M\};$$

при этом безразлично, идет ли речь о процессе (2.4.28) или (2.4.33).

Сформулированное утверждение является частным случаем результатов из § 3 гл. III.

§ 5. Рандомизация фейеровских процессов

Итерационный процесс решения той или иной задачи (например, задачи нахождения хотя бы одного решения системы неравенств) может генерироваться несколькими

итерационными операторами, применяемыми в том или ином порядке их чередования. Пусть таковыми будут

$$\{\varphi_i(\mathbf{x}): i = 1, \dots, m\} \subset \{\mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}\}.$$

Тогда можно организовать процесс

$$\{\mathbf{x}_{t+1} \in \Phi[\mathbf{p}_t](\mathbf{x}_t)\}_{\forall t} \quad (2.5.1)$$

со смешанной стратегией \mathbf{p}_t использования этих отображений на шаге t ; здесь

$$\mathbf{p}_t \in P = \left\{ \mathbf{p} = [p^1, \dots, p^m] \geq 0: \sum_{i=1}^m p^i = 1 \right\},$$

$$\Phi[\mathbf{p}](\mathbf{x}) = \begin{cases} \varphi_1(\mathbf{x}) & \text{с вероятностью } p^1, \\ \dots & \dots \\ \varphi_m(\mathbf{x}) & \text{с вероятностью } p^m. \end{cases} \quad (2.5.2)$$

Генерирование последовательности $\{\mathbf{p}_t\}$ в (2.5.1) может быть подчинено соображениям, носящим как точный характер, так и эвристический, и направленным на повышение эффективности вычислительного процесса.

5.1. Теоремы о сходимости. Положим

$$\Phi[\mathbf{p}](\mathbf{x}) = \sum_{i=1}^m p^i \varphi_i(\mathbf{x}); \quad (2.5.3)$$

здесь $\mathbf{p} \in P$.

Теорема 2.5.1. Если отображения $\{\varphi_i: \forall i \in \mathbb{N}_m\} \subset F_M$ замкнуты, то процесс

$$\{\mathbf{x}_{t+1} \in \Phi[\mathbf{p}_t](\mathbf{x}_t)\}_{\forall t} \quad (2.5.4)$$

сходится к элементу из M , какова бы ни была последовательность $\{\mathbf{p}_t\} \subset P$.

Доказательство. Заметим, во-первых, что $\Phi[\mathbf{p}](\cdot) \in F_M$, $\forall \mathbf{p} \in P$, и $\Phi[\mathbf{p}](\mathbf{x})$ замкнуто по $[\mathbf{x}, \mathbf{p}] \in \mathbb{R}^n \times \mathbb{R}^m$. Во-вторых, в силу свойства

$$|\mathbf{x}_{t+1} - \mathbf{y}| \leq |\mathbf{x}_t - \mathbf{y}|, \quad \forall t, \quad \forall \mathbf{y} \in M,$$

достаточно доказать, что $\{\mathbf{x}_t\}' \cap M \neq \emptyset$.

Пусть $\{\mathbf{x}_{i_t}\} \rightarrow \mathbf{x}'$, можно считать, что $\{\mathbf{x}_{i_t+1}\} \rightarrow \mathbf{x}''$ и $\{\mathbf{p}_{i_t}\} \rightarrow \mathbf{p}' \in P$. Так как $\Phi[\mathbf{p}](\mathbf{x})$ замкнуто по $[\mathbf{x}, \mathbf{p}]$, то $\mathbf{x}'' \in \Phi[\mathbf{p}'](\mathbf{x}')$. Если $\mathbf{x}' \in M$, то, как уже отмечалось, теорема верна. В случае же $\mathbf{x}' \notin M$ получаем для $\mathbf{y} \in M$

неравенство $|\mathbf{x}'' - \mathbf{y}| < |\mathbf{x}' - \mathbf{y}|$, которое противоречиво, так как все точки из $\{\mathbf{x}_t\}'$ лежат на одинаковом расстоянии от \mathbf{y} (непосредственно вытекает из определения М-фейеровской последовательности). Доказательство теоремы завершено.

Следствие 1. В предположениях теоремы 2.5.1 процесс (2.5.1) сходится к элементу из M .

Следствие 2. Пусть τ — произвольное отображение натурального ряда чисел N на N_m . В сделанных относительно отображений $\{\varphi_i\}$ предположениях процесс

$$\{\mathbf{x}_{t+1} \in \varphi_{\tau(t)}(\mathbf{x}_t)\} \forall t$$

сходится к элементу из M .

Рассмотрим теперь случай, когда $\varphi_i \in F_{M_i} (\forall i \in N_m)$ и $M = \bigcap M_i \neq \emptyset$. Положим

$$P_\varepsilon = \{p = [p^1, \dots, p^m] \in P: p^i \geq \varepsilon (\forall i \in N_m)\},$$

где $0 < \varepsilon < 1/m$. Образует из заданных $\{\varphi_i\}$ отображение $\Phi_\varepsilon[p](\mathbf{x})$ по формуле (2.5.3).

Теорема 2.5.2. Если отображения $\varphi_i \in F_{M_i} (\forall i \in N_m)$ замкнуты, то процесс

$$\{\mathbf{x}_{t+1} \in \Phi_\varepsilon[p_t](\mathbf{x}_t)\} \forall t$$

сходится к элементу из M , какова бы ни была последовательность $\{p_t\} \subset P_\varepsilon$.

Обоснование этой теоремы почти полностью повторяет доказательство теоремы 2.5.1, поэтому мы его опускаем.

Рассмотрим еще одно отображение, а именно $\varphi_\varepsilon[p](\mathbf{x})$, которое получается из $\varphi_i \in F_{M_i} (\forall i \in N_m)$ согласно правилу (2.5.2).

Теорема 2.5.3. Если $\varphi_i \in F_{M_i} (\forall i \in N_m)$ — замкнуты и $M = \bigcap M_i$ телесно, то процесс

$$\{\mathbf{x}_{t+1} \in \varphi_\varepsilon[p_t](\mathbf{x}_t)\} \forall t$$

сходится к элементу из M , какова бы ни была последовательность $\{p_t\} \subset P_\varepsilon$.

Доказательство. Если при некотором t' : $\mathbf{x}_{t'} \in M$, то $\mathbf{x}_t = \mathbf{x}_{t'}$, $\forall t > t'$, но тогда, очевидно, теорема верна. Поэтому предположим $\{\mathbf{x}_t\} \cap M = \emptyset$. Последовательность $\{\mathbf{x}_t\}$ в этом случае (если не считать возможных повторений ее членов) является М-фейеровской, а потому, со-

гласно свойству 2.1.1 (см. § 1 гл. II): $\{x_t\} \rightarrow x'$. Пусть $i \in N_m$. Так как $p_i^t \geq \varepsilon > 0$, то существует подпоследовательность $\{x_{it}\} \subset \{x_t\}$ такая, что $x_{it} \in \Phi_i(x_{it-1})$, $\forall i$. Из замкнутости $\Phi_i(x)$ вытекает $x' \in \Phi_i(x')$, поэтому $x' = \Phi_i(x')$, т. е. $x' \in M_i$. В силу произвольности $i \in N_m$ получаем $x \in M$, что и требовалось.

5.2. Управление фейеровским процессом. Рассмотрим процесс (2.5.1) с точки зрения управления параметром p_t с целью повышения эффективности вычислительного процесса. Заметим, во-первых, что если каждое из отображений Φ_i обладает свойством

$$|z - M| \leq \theta_i |x - M|, \quad \forall x \in S_0, \quad \forall z \in \Phi_i(x)$$

при любом ограниченном $S_0 \subset R^n$ (θ_i может, естественно, зависеть от S_0), то и $\Phi[p](x)$ будет обладать аналогичным свойством, а именно:

$$|z - M| \leq \theta |x - M|, \quad \forall x \in S_0, \quad \forall z \in \Phi[p](x);$$

при этом в роли θ может выступить число $\max \{\theta_i; \forall i \in N_m\}$; выше p — произвольный элемент из P .

Следовательно, в этих условиях для процесса (2.5.4) гарантируется сходимость со скоростью геометрической прогрессии. Но так как числа θ_i ($\forall i \in N_m$) — это внутренние характеристики отображений $\{\Phi_i\}$, как правило, неизвестные вычислителю, то упорядочение их по предпочтительности невозможно. Поэтому и возникает вопрос об управлении последовательностью $\{p_t\}$, обеспечивающем организацию процесса, выводящего на использование наиболее эффективного итерационного отображения из числа $\{\Phi_i\}$, или на использование некоторой эффективной смешанной стратегии их чередования.

Рассмотрим некоторые способы управления последовательностью $\{p_t\}$ в процессе (2.5.4). Пусть $d(x)$ — выпуклая функция дефекта для множества M : $d(y) = 0$, $\forall y \in M$; $d(x) > 0$, $\forall x \notin M$. Если M задано системой выпуклых неравенств (2.4.1), то в качестве $d(x)$ можно взять любую функцию вида (2.4.8), т. е., например, $d(x) = \sum_{(j)} f_j^+(x)$.

Если $x_t \in \Phi[p_{t-1}](x_{t-1})$, то

$$x_t = \sum_{i=1}^m p_{t-1}^i y_i^{t-1},$$

где $y_i^{t-1} \in \varphi_i(x_{t-1})$. Положим

$$\Delta_i^t = [d(y_i^{t-1}) - d(y_i^t)]^+.$$

Величина Δ_i^t может пониматься как «вклад» итерационного отображения φ_i на шаге t в задачу минимизации дефекта $d(x)$, реализовавшегося на $(t-1)$ -м шаге. В пропорции с этим «вкладом» может быть пересчитана стратегия p_{t-1} . В силу сказанного организация процесса (2.5.4) может выглядеть следующим образом.

Возьмем любое число $t_0 \in \mathbb{N}$, $t_0 > 1$. Предположим, что тем или иным образом реализован отрезок x_0, x_1, \dots, x_{t_0} процесса (2.5.4), например при постоянной стратегии p_0 (т. е. при $p_t = p_0, \forall t \in \mathbb{N}_{t_0}$). Далее индуктивно: если получен элемент $x_t, t > t_0$, то x_{t+1} берется из $\Phi[p_t](x_t)$, а координаты p_i^t вектора p_t вычисляются по формуле

$$p_i^t = \frac{\sum_{k=t-t_0}^{t-1} \Delta_i^k}{\sum_{i=1}^m \sum_{k=t-t_0}^{t-1} \Delta_i^k} \quad (\forall i \in \mathbb{N}_m). \quad (2.5.5)$$

При решении систем линейных неравенств такой метод дал хорошие результаты [69].

Если $\varphi_i \in F_{M_i}$ и $M = \bigcap_{(i)} M_i \neq \emptyset$, то управление вектором p в (2.5.4) можно осуществлять в соответствии с соотношением (2.5.5); при этом в нем следует положить

$$\Delta_i^k = [d_i(y_i^{k-1}) - d_i(y_i^k)]^+,$$

где $d_i(x)$ — функция дефекта для $M_i (\forall i \in \mathbb{N}_m)$. При $M_i = \{x: f_i(x) \leq 0\}$ в качестве $d_i(x)$ можно взять $f_i^+(x)$. Более простым способом пересчета вектора p является следующий:

$$p_i^t = \frac{d_i(x_t)}{\sum_{k=1}^m d_k(x_t)} \quad (\forall i \in \mathbb{N}_m). \quad (2.5.6)$$

Применительно к системе линейных неравенств $(h_i, x) - b_i \leq 0 (i = 1, \dots, m)$ метод (2.5.4), в котором генерирование последовательности $\{p_t\}$ осуществляется

согласно (2.5.6) при $d_i(\mathbf{x}) = [(\mathbf{h}_i, \mathbf{x}) - b_i]^+$, оказался эффективным [107].

Если в описываемой ситуации в качестве φ_i взять

$$\varphi_i(\mathbf{x}) = \mathbf{x} - \lambda \frac{d_i(\mathbf{x})}{|\mathbf{h}_i|^2} \mathbf{h}_i, \quad 0 < \lambda < 2,$$

то (2.5.4) принимает вид

$$\mathbf{x}_{t+1} = \mathbf{x}_t - \frac{\lambda}{\delta_t} \sum_{i=1}^m \frac{d_i^2(\mathbf{x})}{|\mathbf{h}_i|^2} \mathbf{h}_i,$$

где

$$\delta_t = \sum_{i=1}^m d_i(\mathbf{x}_t).$$

§ 6. Фейеровские процессы для задач выпуклого программирования с негладкими ограничениями

Как уже отмечалось во введении к настоящей главе, итерационный оператор для задачи

$$\max \{f_0(\mathbf{x}) : \mathbf{x} \in M\} \quad (2.6.1)$$

может строиться в виде композиции двух итерационных операторов φ и ψ , первый из которых является итерационным для задачи нахождения хотя бы одного элемента из M , а второй — для задачи $\max_{(x)} f_0(\mathbf{x})$. В методах фейе-

ровского типа (для задач выпуклого программирования) в качестве φ выступает M -фейеровское отображение, в качестве же ψ — отображение

$$\mathbf{x} \rightarrow \mathbf{x} + \alpha \partial f_0(\mathbf{x}),$$

где $\alpha > 0$, $\partial f_0(\mathbf{x})$ — субдифференциал вогнутой функции $f_0(\mathbf{x})$. Композиция φ и ψ дает отображение

$$\mathbf{x} \rightarrow \varphi(\mathbf{x} + \alpha \partial f_0(\mathbf{x})); \quad (2.6.2)$$

при этом под $\varphi(S)$ понимается $\bigcup_{y \in S} \varphi(y)$. Если $f_0(\mathbf{x}) = (c, \mathbf{x})$ (что, кстати, не делает постановку задачи (2.6.1) менее общей), то с точностью до замены переменных (2.6.2) можно переписать в виде

$$\mathbf{x} \rightarrow \Phi_0[\alpha](\mathbf{x}) = \varphi(\mathbf{x}) + \alpha c. \quad (2.6.3)$$

В целях определенных удобств будем оперировать ниже не отображением (2.6.2), а отображением (2.6.3), отнесенным к произвольной функции $f_0(x)$:

$$x \rightarrow \Phi[\alpha](x) = \varphi(x) + \alpha \partial f_0(x). \quad (2.6.4)$$

Отображению (2.6.4) соответствует процесс

$$\{x_{t+1} \in \Phi[\alpha_t](x_t)\}_{\forall t}, \quad (2.6.5)$$

решающий при определенных ограничениях (в частности, на последовательность $\{\alpha_t\}$) задачу (2.6.1).

Обратим внимание на то, что (2.6.5) является обобщением процесса (2.3.6), возникшего в связи с применением фейеровских методов к одновременному решению двойственной пары задач линейного программирования.

Множество M может быть задано, в частности, системой выпуклых неравенств

$$f_j(x) \leq 0 \quad (\forall j \in N_m). \quad (2.6.6)$$

На выпуклые функции $\{-f_0(x), f_1(x), \dots, f_m(x)\}$ никаких условий гладкости накладываться не будет, в этом смысле рассматриваемые ниже методы являются методами негладкой оптимизации.

В общей проблеме негладкой оптимизации мы выделим три ее составные части:

- 1) решение негладких систем неравенств;
- 2) максимизация (минимизация) негладких функций без ограничений;
- 3) максимизация (минимизация) негладких функций при негладких ограничениях.

С точки зрения применения к этим задачам фейеровских методов первая из них довольно полно решается в §§ 1—5 настоящей главы. Поскольку вторая задача является частным случаем третьей, то представляется возможным ограничиться рассмотрением только последней.

Заметим, что третья задача, наиболее общая, может быть сведена к задаче с линейной (а потому гладкой) оптимизационной функцией, однако далеко не всегда имеется практический смысл это делать. В пояснение к сказанному запишем очевидное соотношение:

$$\max f_0(x) : x \in M = \max \{x_{n+1} : x \in M, x_{n+1} \leq f_0(x)\}.$$

6.1. z -нормализованные фейеровские процессы. Пусть M — непустое, замкнутое и выпуклое множество из \mathbb{R}^n . Если $z \in \mathbb{R}^n$ и \tilde{z} — метрическая проекция на M , т. е.

$$\tilde{z} = \arg \min \{ \|x - z\| : x \in M \},$$

то элемент \tilde{z} будем называть z -нормальным. Элемент \tilde{z} определяется однозначно.

Для отыскания \tilde{z} может быть предложен процесс (называемый нами z -нормализованным):

$$\{x_{t+1} \in (1 - \alpha_t) \varphi(x_t) + \alpha_t z\}_{\forall t}, \quad (2.6.7)$$

в котором $\varphi \in F_M$,

$$\{\alpha_t \geq 0\} \rightarrow 0, \quad \sum_{(t)} \alpha_t = +\infty. \quad (2.6.8)$$

Теорема 2.6.1. Если отображение $\varphi \in F_M$ замкнуто и выполнено условие (2.6.8), то процесс (2.6.7) сходится к элементу \tilde{z} — z -нормальному элементу из M .

Это утверждение получается как частный случай теоремы 3.3.3, отнесенной к ситуации нестационарного процесса.

6.2. Теорема о сходимости базового процесса.

Теорема 2.6.2. Пусть $\varphi \in F_M$, φ замкнуто, M ограничено. Если выполнено условие (2.6.8), то при достаточно малом $\bar{\alpha} = \sup_{(t)} \alpha_t > 0$ процесс (2.6.5) ограничен и

$$\{x_t\}' \subset \bar{M},$$

где $\bar{M} = \text{Arg} \max \{f_0(x) : x \in M\}$.

Сформулированное утверждение является частным случаем доказываемой в гл. III теоремы 3.3.2.

Замечание. В формулировке теоремы 2.6.2 содержатся условия ограниченности M и достаточной малости числа $\bar{\alpha} = \sup_{(t)} \alpha_t > 0$. Такие ограничения носят неконструктивный характер, особенно второе ограничение. Чтобы избежать связанных с этим затруднений, при организации вычислительного процесса можно использовать известный прием «возврата». Он состоит в следующем. Тем или иным образом фиксируют шар S (или параллелепипед), содержащий начальную точку x_0 и имеющий с оптимальным множеством \bar{M} задачи (2.6.1) непустое

пересечение (в практических задачах это можно сделать, как правило, без труда). Модификация процесса теперь состоит в том, что если он дает точку, не лежащую в S , то последующим членом делают проекцию этой точки на S . Следовательно, организация процесса подчинена рекуррентному включению

$$x_{t+1} \in \pi_S(\Phi[\alpha_t](x_t)).$$

Для процесса, порожденного этим включением, утверждение теоремы (2.6.2) остается справедливым, однако условия ограниченности M и достаточной малости числа $\bar{\alpha}$ в этой ситуации излишни.

§ 7. Несобственные задачи линейного программирования

В связи с содержанием пп. 2.3 и 2.4 § 2 гл. II рассмотрим задачу линейного программирования

$$\max \{(c, x) : Ax \leq b, x \geq 0\}, \quad (2.7.1)$$

ограничения которой не обязательно совместны, т. е. возможно и $M = \{x \geq 0 : Ax \leq b\} = \emptyset$. Будем считать, что $x \in \mathbb{R}^n$, $b \in \mathbb{R}^m$. Двойственной для (2.7.1) является задача

$$\min \{(b, u) : A^T u \geq c, u \geq 0\}. \quad (2.7.2)$$

Обозначим $M^* = \{u \geq 0 : A^T u \geq c\}$, множество M^* также может быть пустым.

Определение 2.7.1. Назовем задачу (2.7.1) *несобственной*, если выполнено одно из следующих условий:

$$1) \quad M = \emptyset, \quad M^* \neq \emptyset; \quad (2.7.3)$$

$$2) \quad M \neq \emptyset, \quad M^* = \emptyset; \quad (2.7.4)$$

$$3) \quad M = \emptyset, \quad M^* = \emptyset. \quad (2.7.5)$$

В зависимости от выполнимости одного из этих условий будем говорить о *несобственной задаче* соответственно 1-го, 2-го или 3-го рода. Разрешимую задачу будем называть *собственной*.

Сделаем замечание к вопросу о важности рассмотрения таких моделей. Моделирование производственных задач оптимального планирования (средствами линейного программирования) в условиях принципа максимального, безрезервного использования производственных мощно-

стей, напряженности плана и дефицита на некоторые ресурсы может привести к тому, что система ограничений полученной модели будет несовместной, а сама модель — несобственной. Практика решения производственных задач планирования показывает, что это довольно обычная ситуация. Конечно, в этом случае, исходя из тех или иных эвристических соображений, можно ряд ограничений снять или ослабить, получив обычную разрешимую модель. Однако куда важнее и интереснее подход, основанный на применении объективных процедур для «развязки» такой модели, т. е. для преобразования ее к виду разрешимой задачи. Если такое преобразование адекватным образом связано с необходимостью тех или иных изменений в самом производстве, то не безразлично, во что они обойдутся. Этот принцип можно было бы, в частности, и положить в основу метода «развязки», хотя возможны подходы и на основе других критериев.

«Развязку» несобственной модели можно еще интерпретировать как развязку узких мест.

Таким образом, модель с несовместной системой ограничений содержательно может быть не менее важной (а в ряде случаев и более важной), чем с совместной. Сказанное дополним еще одним соображением. Возможно, моделируя производственную ситуацию (речь идет о задачах планирования), целесообразно идти от собственной модели к несобственной (за счет включения в модель новых ограничений, не учтенных или не учитываемых ранее), а от последней — к собственной, но уже на основе объективных процедур преобразования несобственной модели.

Поясним смысл условий (2.7.3)—(2.7.5). Первое из них означает, что если при некотором значении параметра $\eta \in \mathbb{R}^m$ (пусть при $\eta = \bar{\eta}$) система неравенств

$$Ax \leq b + \bar{\eta}, \quad x \geq 0, \quad (2.7.6)$$

совместна, задача ЛП

$$\max \{(c, x) : Ax \leq b + \bar{\eta}, x \geq 0\} \quad (2.7.7)$$

разрешима. Действительно, пусть (2.7.7) разрешима, тогда разрешима и двойственная к ней задача

$$\min \{(b + \bar{\eta}, u) : A^T u \geq c, u \geq 0\},$$

следовательно, $M^* \neq \emptyset$. Обратно, если $M^* \neq \emptyset$ и система (2.7.6) совместна, то, взяв любое ее решение \bar{x} и некоторый вектор $\bar{u} \in M^*$, будем иметь

$$(c, \bar{x}) \leq (A^T \bar{u}, \bar{x}) = (\bar{u}, A\bar{x}) \leq (b + \bar{\eta}, \bar{u}),$$

т. е.

$$\sup \{(c, x) : Ax \leq b + \bar{\eta}\} < (b + \bar{\eta}, \bar{u}),$$

что и дает (в силу теоремы о достижимости в линейном программировании) разрешимость задачи (2.7.7).

Условие (2.7.4) означает, что в задаче (2.7.1) оптимальное значение равно $+\infty$.

Наконец, условие (2.7.5) эквивалентно тому, что как только система (2.7.6) совместна, оптимальное значение задачи (2.7.7) равно $+\infty$.

Приведенные характеристики являются тривиальными следствиями теорем двойственности в линейном программировании.

В прикладном отношении наиболее интересным является случай, когда задача (2.7.1) является несобственной 1-го рода, т. е. когда реализуется условие (2.7.3). Что касается (2.7.4) и (2.7.5), то эти условия, скорее, важны в вопросах анализа корректности задач ЛП.

Что касается методов получения оптимальных планов уже откорректированной задачи, то здесь возможны два подхода. Первый из них двухэтапный. Он состоит в том, что вначале осуществляется «развязка», а затем уже решение полученной задачи. Второй подход исходит из возможности организации такой вычислительной процедуры, которая ведет к оптимальному плану исходной модели, если она является собственной, и автоматически — к оптимальному плану откорректированной модели в противном случае.

Именно второй подход был реализован в § 2 гл. II (пп. 2.3 и 2.4) применительно к решению систем неравенств, которые могут быть как совместными, так и несовместными.

7.1. Симметрическая аппроксимация. Задачам (2.7.1) и (2.7.2) поставим в соответствие

$$\max \{(c + \xi, x) : Ax \leq b + \eta, x \geq 0\}, \quad (2.7.8)$$

$$\min \{(b + \eta, u) : A^T u \geq c + \xi, u \geq 0\}, \quad (2.7.9)$$

где $R^n \ni \xi$ и $R^m \ni \eta$ — векторные параметры.

Задачи (2.7.8) и (2.7.9) являются взаимно двойственными. Обозначим через K множество векторов $[\bar{\xi}, \bar{\eta}] \in \mathbb{R}^{n+m}$ таких, что после подстановки их в (2.7.8) мы получаем разрешимую задачу. Аналогичное множество для (2.7.9) будет тем же, т. е. будет совпадать с K . Его еще можно охарактеризовать как множество векторов $[\xi, \eta]$, для которых система неравенств

$$Ax \leq b + \eta, \quad A^T u \geq c + \xi, \quad x \geq 0, \quad u \geq 0 \quad (2.7.10)$$

совместна относительно $[x, u]$. Последнее утверждение является очевидным, так как из совместности систем ограничений в паре двойственных задач следует их разрешимость, и из разрешимости одной из них следует разрешимость другой.

K есть декартово произведение выпуклых полиэдральных множеств K_1 и K_2 :

$$K_1 = \{\eta: \{x: Ax \leq b + \eta, x \geq 0\} \neq \emptyset\} \quad (2.7.11)$$

и

$$K_2 = \{\xi: \{u: A^T u \geq c + \xi, u \geq 0\} \neq \emptyset\}. \quad (2.7.12)$$

Задание каждого из них системой линейных неравенств может быть получено путем последовательного исключения x_1, \dots, x_n из системы (2.7.11) и u_1, \dots, u_m — из системы (2.7.12).

В качестве критериев выбора значений для аппроксимирующих векторных параметров ξ и η могут быть выбраны функции

$$f_1(\xi, \eta) = \sum_{i=1}^n \bar{c}_i |\xi_i| + \sum_{j=1}^m \bar{b}_j |\eta_j|,$$

$$f_2(\xi, \eta) = \sum_{i=1}^n \bar{c}_i \xi_i^2 + \sum_{j=1}^m \bar{b}_j \eta_j^2,$$

которые необходимо минимизировать на K ; здесь $\bar{c}_i > 0$, $\bar{b}_j > 0$ ($\forall i \in N_n, \forall j \in N_m$), $\xi = [\xi_1, \dots, \xi_n]^T$, $\eta = [\eta_1, \dots, \eta_m]^T$.

Первую из них можно интерпретировать следующим образом: изменение вектора b на величину η сопряжено с потерями $\sum_{j=1}^m \bar{b}_j |\eta_j|$, а изменение вектора c на вели-

чину ξ — с потерями $\sum_{i=1}^n \bar{c}_i |\xi_i|$; требуется найти $[\bar{\xi}, \bar{\eta}] \in K$,

обеспечив минимальные суммарные потери $f_1(\bar{\xi}, \bar{\eta})$.

Задача

$$\min \{f_1(\xi, \eta): [\xi, \eta] \in K\}$$

распадается на две самостоятельные задачи выпуклого кусочно-линейного программирования

$$\min \left\{ \sum_{i=1}^n \bar{c}_i |\xi_i|: \xi \in K_2 \right\}$$

и

$$\min \left\{ f_1(\eta) = \sum_{j=1}^m \bar{b}_j |\eta_j|: \eta \in K_1 \right\}. \quad (2.7.13)$$

Рассматривая задачу минимизации функции $f_2(\xi, \eta)$ на K , приходим к двум задачам выпуклого квадратичного программирования

$$\min \left\{ \sum_{i=1}^n \bar{c}_i \xi_i^2: \xi \in K_2 \right\}$$

и

$$\min \left\{ f_2(\eta) = \sum_{j=1}^m \bar{b}_j \eta_j^2: \eta \in K_1 \right\}. \quad (2.7.14)$$

7.2. Несобственные задачи ЛП 1-го рода. Если задача (2.7.1) является несобственной 1-го рода, то задача аппроксимации запишется как

$$\min \{f(\eta): \eta \in K_1\},$$

$f(\eta)$ — функция качества аппроксимации. Рассмотрим случай $f(\eta) = f_1(\eta)$. Конечная цель теперь состоит в том, чтобы найти оптимальный вектор задачи

$$\max \{(c, x): Ax \leq b + \tilde{\eta}, x \geq 0\}, \quad (2.7.15)$$

где $\tilde{\eta} \in \text{Arg} (2.7.13)$.

Если $Ax \leq b \sim l_j(x) \leq 0 (\forall j \in N_m)$, то (2.7.15) можно переписать в эквивалентном виде:

$$\max \{(c, x): l_j(x) + y_j = \tilde{\eta}_j, x \geq 0, y_j \geq 0 (\forall j \in N_m)\}.$$

Мы имеем, таким образом, двухэтапную задачу последо-

вательного программирования, допускающую эквивалентное сведение ее к одной задаче вида [24]

$$\max \left\{ (c, x) - r \sum_{j=1}^m \bar{b}_j |l_j(x) + y_j| : x \geq 0, y_j \geq 0 \ (\forall j \in N_m) \right\}, \quad (2.7.16)$$

где r — достаточно большое положительное число. Последняя может быть переписана в форме задачи ЛП:

$$\max \left\{ (c, x) - r \sum_{j=1}^m \bar{b}_j u_j : \right. \\ \left. -u_j \leq l_j(x) + y_j \leq u_j, x \geq 0, y_j \geq 0 \ (\forall j \in N_m) \right\}.$$

В качестве итоговой для рассматриваемой двухэтапной задачи можно взять также задачу

$$\max \left\{ (c, x) - r \sum_{j=1}^m \bar{b}_j (u'_j + u''_j) : l_j(x) + y_j = u'_j - u''_j, \right. \\ \left. x \geq 0, y_j \geq 0, u'_j \geq 0, u''_j \geq 0 \ (\forall j \in N_m) \right\}.$$

Такой прием редукции применяется обычно при использовании симплекс-метода в условиях неразрешимости системы ограничений в задаче ЛП.

7.3. Применение фейеровских методов. Возьмем теперь в качестве $f(\eta)$ функцию $f_2(\eta) = \sum_{j=1}^m \bar{b}_j \eta_j^2$ и рассмотрим задачу (2.7.14) в сопряжении с (2.7.15) при $\tilde{\eta} \in \text{Arg} (2.7.14)$. В этом случае рассмотренная в предыдущем пункте редукция будет носить приближенный характер, а именно: задача

$$\max \{ (c, x) - r \sum \bar{b}_j [l_j(x) + y_j]^2 : x \geq 0, y \geq 0 \} \quad (2.7.17)$$

асимптотически по r эквивалентна задаче (2.7.15), при этом (2.7.17) тем точнее приближает задачу (2.7.15), чем больше r [27]. Что касается задачи (2.7.17), то для ее решения может быть применен фейеровский итерационный метод, описанный в п. 2.3 § 2 гл. II, с учетом возможности замены в (2.7.17) функции (c, x) на $-(c, x) - \gamma^2$ при достаточно большом $\gamma > 0$ (см. п. 2.6 § 2 гл. II).

ФЕЙЕРОВСКИЕ МЕТОДЫ ДЛЯ НЕСТАЦИОНАРНЫХ ЗАДАЧ ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ С НЕГЛАДКИМИ ОГРАНИЧЕНИЯМИ

В данной главе изучаются процессы вида

$$\{x_{t+1} \in \varphi(x_t, s_t) + \alpha_t \partial f_0(x_t)\}_{\forall t},$$

поставленные в соответствие нестационарной модели выпуклого программирования

$$\max \{f_0(x): f_j(x, s_t) \leq 0, \forall j \in N_m\}_{\forall t},$$

т. е. модели, реализованной в форме дискретной последовательности ее состояний; выше $\varphi(\cdot, s_t) \in F_{M_t}$, $M_t = \{x: f_j(x, s_t) \leq 0, \forall j \in N_m\}$. Если, например, как уже отмечалось, в задаче ЛП

$$\max \{(c, x): (a_j, x) - b_j \leq 0, \forall j \in N_m\}$$

вектор исходных данных $s = [c, a_1, \dots, a_m, b]$ реализуется в форме последовательности $\{s_t\}$ его приближений, то мы можем говорить о нестационарной модели линейного программирования. На выписанный выше процесс можно смотреть как на итерационный метод, соединяющий в себе оптимизацию и одновременно идентификацию модели по системе исходных данных.

§ 1. Фейеровские отображения, зависящие от параметра, и порождаемые ими процессы

1.1. Исходные предположения. Выпишем процесс

$$\{x_{t+1} \in \varphi(x_t, s_t)\}_{\forall t}, \quad (3.1.1)$$

в котором $x \in R^n, s \in V \subset H$ — вещественное гильбертово пространство, $\varphi(\cdot, s) \in F_{M(s)}$, $M(s) = \{x: \varphi(x, s) = x\}$. При изучении процесса (3.1.1) будет предполагаться, как правило, сходимость последовательности $\{s_t\}$. Такое требование является довольно жестким, тем не менее в ряде случаев оно возникает естественным образом. Например,

если H — конечномерно и \mathbf{s} интерпретируется как вектор исходных данных, то выбор \mathbf{s}_t на шаге t процесса (3.1.1) из некоторого сужающегося от шага к шагу информационного множества S_t можно подчинить правилу нормализации:

$$\mathbf{s}_t = \arg \min \{ \|\mathbf{s}\| : \mathbf{s} \in S_t \}.$$

В предположении, что множества S_t ($\forall t \in \mathbb{N}$) выпуклы, замкнуты и $S = \bigcap_{(t)} S_t \neq \emptyset$, имеет место сходимость $\{\mathbf{s}_t\} \rightarrow \mathbf{s}' \in S$.

С другой стороны, получение точных результатов о последовательности $\{\mathbf{x}_t\}$, порожденной в силу (3.1.1), естественно связывать с регулярным (в том или ином смысле) поведении параметра \mathbf{s} . Условие сходимости последовательности $\{\mathbf{s}_t\}$ можно понимать как одну из форм проявления такой регулярности.

Введем ряд обозначений: если A и B — множества из \mathbb{R}^n , то $|A - B| = \sup_{x \in A} \inf_{y \in B} |\mathbf{x} - \mathbf{y}|$ и $|A - B|_+ = \max \{ |A - B|, |B - A| \}$ — хаусдорфово отклонение множеств A и B .

Определение 3.1.1. **Отображение**

$$\varphi(\mathbf{x}, \mathbf{s}) \in \{\mathbb{R}^n \times H \rightarrow 2^{\mathbb{R}^n}\}$$

называется *замкнутым по $\mathbf{z} = [\mathbf{x}, \mathbf{s}]$* , если из условия

$$\{[\mathbf{x}_t, \mathbf{s}_t]\} \rightarrow [\mathbf{x}', \mathbf{s}'], \varphi(\mathbf{x}_t, \mathbf{s}_t) \ni \mathbf{y}_t \rightarrow \mathbf{y}' \quad (t \rightarrow +\infty)$$

следует, что $\mathbf{y}' \in \varphi(\mathbf{x}', \mathbf{s}')$.

Пусть $\varphi(\cdot, \mathbf{s}) \in F_{M(\mathbf{s})}$, $\forall \mathbf{s} \in V$. Последовательности $\{\mathbf{s}_t\} \subset V$ поставим в соответствие множества

$$M_{\infty, t} = \bigcap_{k \geq t} M(\mathbf{s}_k), \quad M_{\infty} = \bigcup_{(t)} M_{\infty, t};$$

здесь $M(\mathbf{s}) = \{\mathbf{x} : \varphi(\mathbf{x}, \mathbf{s}) = \mathbf{x}\}$. Каждое из введенных здесь множеств является выпуклым (в случае непустоты). Введем ряд условий, которые в разных сочетаниях будут фигурировать в формулируемых ниже теоремах.

1°. $V \supset \{\mathbf{s}_t\} \rightarrow \mathbf{s}'$;

2°. $\varphi(\mathbf{x}, \mathbf{s})$ замкнуто по $\mathbf{z} = [\mathbf{x}, \mathbf{s}] \in \mathbb{R}^n \times H$;

3°. $M_{\infty} \neq \emptyset$;

4°. $|M_t - M|_+ \rightarrow 0$ ($t \rightarrow +\infty$); здесь для краткости положено $M_t = M(\mathbf{s}_t)$.

Справедливы следующие простые утверждения.

Утверждение 3.1.1. Из условий 1⁰, 2⁰ и 3⁰ следует условие

$$|M_t - M| \rightarrow 0 \quad (t \rightarrow +\infty).$$

Утверждение 3.1.2. Из условий 1⁰ и 2⁰ следует условие $M_\infty \subset M$.

Докажем, например, второе из них. Пусть $y \in M_\infty$, т. е. при некотором t' : $y \in M_t, \forall t \geq t'$. Это означает: $y \in \varphi(y, s_t)$, что в силу 2⁰ дает $y = \varphi(y, s')$, т. е. $y \in M$.

Укажем на два простых случая, когда выполнено свойство 4⁰. Если

1) $M_t \supset M_{t+1}, \forall t; \bigcap M_t = M$ не пусто и ограничено или

2) $M_t \subset M_{t+1}, \forall t; \bigcup M_t = M$ не пусто и ограничено, то $|M_t - M|_+ \rightarrow 0 \quad (t \rightarrow +\infty)$.

Эти случаи могут соответствовать ситуации, когда нестационарность модели порождается последовательностью аппроксимаций ее допустимой области (внешних — в случае 1), внутренних — в случае 2)).

1.2. Теоремы о сходимости. Повсюду ниже в теоремах о сходимости процессов предполагается (без оговорок), что начальный элемент $x_0 \in \mathbb{R}^n$ произволен.

Теорема 3.1.1. Пусть для $\varphi(\cdot, s) \in F_{M(s)}, \forall s \in V$, задан процесс (3.1.1). Если при этом выполнены условия 1⁰—3⁰, то последовательность $\{x_t\}$ ограничена и

$$\{x_t\}' \subset M(s') \quad (= M). \quad (3.1.2)$$

Доказательство. Ограниченность процесса следует из определения фейеровского отображения и условия 3⁰. Пусть теперь $x' \in \{x_t\}'$ и $\{x_t\} \rightarrow x'$; можно считать, что и $\{x_{t+1}\} \rightarrow x''$. В силу 2⁰ $x'' \in \varphi(x', s')$. Если $x' \notin M$, то $x' \notin M_\infty$, но тогда $|x'' - y| < |x' - y|, \forall y \in M_\infty$, что дает противоречие, так как все точки из $\{x_t\}'$ лежат на одинаковом расстоянии от y (это непосредственно вытекает из определения фейеровского отображения). Полученное противоречие доказывает включение $x' \in M$, а вместе с тем и (3.1.2).

Следствие 1. Если M_∞ телесно, то последовательность $\{x_t\}$ из теоремы 3.1.1 сходится к элементу из M .

Доказательство. Если $x', x'' \in \{x_t\}'$ и $x' \neq x''$, то, как уже отмечалось, $|\bar{x} - y| = \text{const}$ для $y \in M_\infty$ и

всех $\bar{x} \in \{x_t\}'$. Но тогда получаем, что \bar{x} (а потому и M_∞) принадлежит гиперплоскости, являющейся геометрическим местом точек, равноудаленных от x' и x'' . Так как при $x' \neq x''$ это противоречит телесности M_∞ , то получаем, что $\{x_t\} \rightarrow x' \in M$.

Следствие 2. Если в теореме 3.1.1 сделать дополнительное предположение $M = \bar{M}_\infty$ (\bar{M}_∞ — замыкание множества M_∞), то $\{x_t\} \rightarrow x' \in M$.

Следует из доказательства теоремы с учетом того, что в соотношении $|\bar{x} - y| = \text{const}$, $y \in M_\infty$ (а потому $|\bar{x} - y| = \text{const}$, $\forall y \in \bar{M}_\infty$), $\forall \bar{x} \in \{x_t\}'$, можно положить $y = \bar{x}$.

Замечание. В связи с предыдущим утверждением обратим внимание на то, что если $|M_t - M|_+ \rightarrow 0$ ($t \rightarrow +\infty$) и M телесно, то $M = \bar{M}_\infty$.

Теорема 3.1.2. Пусть для $\varphi(\cdot, s) \in F_{M(s)}$, $H = \mathbf{R}^l$ выполнены условия 1° , M_∞ телесно, последовательность $\{s_t\}$ ограничена. Тогда процесс (3.1.1) сходится к $x' \in \bigcap_{s \in \{s_t\}'} M(s)$.

Сформулированное утверждение вытекает из доказательств теоремы 3.1.1 и ее следствия 1.

Как уже отмечалось (утверждение 3.1.1), выполнимость условий $1^\circ - 3^\circ$ гарантирует свойство $|M_t - M| \rightarrow 0$ ($t \rightarrow +\infty$). Усиление этого свойства (в качестве исходного требования) позволяет при изучении процесса (3.1.1) освободиться от условия 3° с заменой его на 4° .

Справедлива элементарно доказываемая

Лемма 3.1.1. Если множество $Q \subset \mathbf{R}^n$ выпукло и замкнуто, то из $|p - Q| = |q - Q|$ следует существование $y \in Q$ такого, что $|p - y| = |q - y|$.

Лемма 3.1.2. Пусть $\varphi(\cdot, s) \in F_{M(s)}$, $M(s_t) \neq \emptyset$, $\forall t$. Если выполнены условия 1° , 2° и процесс (3.1.1) ограничен, то $\{x_t\}' \subset M$.

Доказательство. Пусть

$$x' = \arg \max \{ |y - M| : y \in \{x_t\}' \} \text{ и } \{x_t\} \rightarrow x'.$$

Считаем, что $\{x_{t-1}\} \rightarrow x''$. Так как $x_{t-1} \in \varphi(x_{t-1}, s_{t-1})$, то в силу условий 1° и 2° имеем: $x' \in \varphi(x'', s')$. Если $x'' \in M$, то $\varphi(x'', s') = x'' = x' \in M$, но тогда, оче-

видно, $\{x_t\}' \subset M$. Если же $x'' \notin M$, то из правила выбора x' и неравенства $|x'' - M| \leq |x' - M|$ вытекает равенство $|x' - M| = |x'' - M|$, а потому, согласно лемме 3.1.1, $|x' - y| = |x'' - y|$ при некотором $y \in M$. Это же противоречит тому, что в ситуации $x' \in \varphi(x'', s')$ и $x'' \notin M$ выполняется неравенство

$$|x' - y| < |x'' - y|, \quad \forall y \in M.$$

Лемма доказана.

Лемма 3.1.3. Пусть $\varphi(\cdot, s) \in F_{M(s)}$, $M(s) \neq \emptyset$, $\forall s \in V$ и выполнены условия 1° , 2° и 4° . Если $M(s')$ ограничено, то для $\delta > \varepsilon > 0$ существуют t' и $\theta \in (0; 1)$ такие, что

$$\begin{aligned} |y - M_t| &\leq \theta |x - M_t|, & (3.1.3) \\ \forall x \in R(\delta, \varepsilon) = M^\delta \setminus M^\varepsilon, \forall y \in \varphi(x, s_t), \quad \forall t \geq t'. \end{aligned}$$

Доказательство. Предположив противное, можно выделить $\{x_t\} \subset R(\delta, \varepsilon)$ и $y_t \in \varphi(x_t, s_t)$, $\forall t$ так, что $\{x_t\} \rightarrow \bar{x}$, $\{y_t\} \rightarrow \bar{y}$, $|y_t - M_t| = \theta_t |x_t - M_t|$ и $\{\theta_t\} \rightarrow 1 - 0$. Отсюда в силу 4° получаем: $|\bar{y} - M| = |\bar{x} - M|$. С другой стороны, из 2° вытекает, что $\bar{y} \in \varphi(\bar{x}, s')$. Но так как $\bar{x} \notin M$, то $|\bar{y} - M| < |\bar{x} - M|$, т. е. получилось противоречие. Лемма доказана.

Замечание. Условия леммы 3.1.3 обеспечивают существование $\theta_0 \in (0; 1)$ такого, что

$$|y - M| \leq \theta_0 |x - M|, \quad \forall x \in R(\delta, \varepsilon), \quad \forall y \in \varphi(x, s').$$

Действительно, взяв t таким, чтобы выполнялось условие

$$\theta_0 = \theta \left(1 + \frac{\varepsilon_t(1+\theta)}{\theta\varepsilon} \right) < 1,$$

будем иметь

$$\begin{aligned} |y - M| &\leq |y - M_t| + \varepsilon_t \leq \theta |x - M_t| + \varepsilon_t \leq \\ &\leq \theta (|x - M| + \varepsilon_t) + \varepsilon_t \leq \\ &\leq \theta \left(1 + \frac{\varepsilon_t(1+\theta)}{\varepsilon\theta} \right) |x - M| = \theta_0 |x - M|, \end{aligned}$$

что и требовалось.

Лемма 3.1.4. Пусть выполнены условия леммы 3.1.3. Тогда процесс (3.1.1) ограничен.

Доказательство. Выберем δ и ε из условий: $\delta > |x_0 - M|$, $\delta - \varepsilon \geq 2 \sup_{(t)} \varepsilon_t$, где x_0 — начальный элемент процесса, $\varepsilon_t = |M_t - M|_+$. Пусть уже доказано неравенство $|x_t - M| \leq \delta$. Докажем, что $|x_{t+1} - M| \leq \delta$. Если $|x_t - M| \geq \varepsilon$, то в силу замечания к лемме 3.1.3 имеем: $|x_{t+1} - M| \leq \theta_0 |x_t - M| < \delta$. Если же $|x_t - M| < \varepsilon$, то $|x_{t+1} - M| \leq |x_{t+1} - M_t| + \varepsilon_t \leq |x_t - M_t| + \varepsilon_t \leq |x_t - M| + 2\varepsilon_t \leq \varepsilon + 2\varepsilon_t \leq \varepsilon + (\delta - \varepsilon) = \delta$, т. е. $\{x_t\} \subset M^\delta$.

Так как M ограничено, то и последовательность $\{x_t\}$ ограничена.

Теорема 3.1.3. Пусть $\varphi(\cdot, s) \in F_{M(s)}$, $M(s) \neq \emptyset$, $\forall s \in V$. Если выполнены условия 1^o, 2^o, 4^o и M ограничено, то процесс (3.1.1) ограничен и

$$\{x_t\}' \subset M.$$

Доказательство теоремы следует из лемм 3.1.2 и 3.1.4.

§ 2. Применение общих теорем о сходимости к нестационарным системам выпуклых неравенств

В § 1 гл. II были описаны некоторые базовые конструкции M -фейеровских отображений для случая, когда множество M задается системой выпуклых неравенств. Конечно, эти конструкции остаются в силе и для параметрической системы выпуклых неравенств

$$f_j(x, s) \leq 0, \quad \forall j \in N_m. \quad (3.2.1)$$

Ниже приводятся результаты о сходимости процессов, порождаемых возникающими при этом отображениями при нестационарном поведении параметра s .

2.1. Основные теоремы о сходимости.

Лемма 3.2.1. Пусть $V \supset S$ — замкнутое множество, $f(x, s)$ — непрерывная по $z = [x, s]$ на $R^n \times S$ и выпуклая по x для всех $s \in V$ функция, $M_j(s) = \{x: f(x, s) \leq 0\} \neq \emptyset$, $\forall s \in V$. Положим

$$\mu(z, h) = \begin{cases} \frac{f(z)}{|h|^2} h, & f(z) > 0, \\ 0, & f(z) \leq 0; \end{cases}$$

здесь $h \in \partial_x f(x, s)$.

Отображение

$$H(z) = \{\mu(z, h) : h \in \partial_x f(x, s)\}$$

является замкнутым на $\mathbb{R}^n \times S$.

Доказательство. Пусть

$$\begin{aligned} \mathbb{R}^n \times S \supset \{z_k\} \rightarrow z' = [x', s'], \\ y_k = \mu(z_k, h_k) \rightarrow y' \quad (k \rightarrow +\infty); \end{aligned}$$

нужно доказать, что $y' \in H(z')$.

Заметим, во-первых, что если $\{h_k\} \rightarrow h'$, то

$$h' \in \partial_x f(x', s'). \quad (3.2.2)$$

Действительно, так как

$$(h_k, x - x_k) \leq f(x, s_k) - f(x_k, s_k), \quad \forall k, \forall x \in \mathbb{R}^n, \quad (3.2.3)$$

то

$$(h', x - x') \leq f(x, s') - f(x', s'), \quad \forall x \in \mathbb{R}^n,$$

но это и означает (3.2.2).

Во-вторых,

$$\sup_{(k)} |h_k| < +\infty. \quad (3.2.4)$$

Действительно, если в (3.2.3) положить $x = x_k + h_k/|h_k|$ для $h_k \neq 0$, то получается

$$|h_k| \leq f\left(x_k + \frac{h_k}{|h_k|}, s_k\right) - f(x_k, s_k),$$

что в силу непрерывности $f(z)$ по z дает (3.2.4).

Наконец, отметим справедливость неравенства

$$|\mu(z, h)| \leq |x - v|, \quad \forall v \in M_f(s), \quad s \in S. \quad (3.2.5)$$

При $x \notin M_f(s)$, т. е. $f(x, s) > 0$, гиперплоскость, определяемая уравнением

$$(h, \bar{x} - x) + f(x, s) = 0$$

с переменным вектором \bar{x} , отделяет точку x от $M_f(s)$. Поэтому для величины $|\mu(z, h)| = f(x, s)/|h|$, являющейся расстоянием от x до выписанной гиперплоскости, выполняется (3.2.5).

Приступим непосредственно к доказательству леммы. Рассмотрим два случая.

1. Пусть $\exists \{i_k\}: f(z_{i_k}) \leq 0 \ (\forall k \in \mathbb{N})$. Тогда $f^+(z') = 0$ и $y' = 0 \in H(z')$.

2. Пусть $f(z_k) > 0, \forall k \geq k'$. Если $f(z') > 0$, то $\lim |h_k| = \delta > 0$. Действительно, в противном случае из (3.2.3) следовало бы

$$f(z') \leq f(x, s'), \quad \forall x \in \mathbb{R}^n,$$

что в силу $f(z') > 0$ противоречит предположению $M_f(s') \neq \emptyset$. Из (3.2.4) и ограниченности $\{h_k\}$ снизу величиной $\delta > 0$ вытекает, что вместе с $\{y_k\} \rightarrow y'$ справедливо и $\{h_k\} \rightarrow h' \neq 0$, но тогда в силу (3.2.2) имеет место включение $y' \in H(z')$. Осталось рассмотреть случай, когда $f(z') = 0$. Тогда $0 \in H(z')$ и $x' \in M_f(s')$. Соотношение (3.2.5) дает $|y_k| \leq |x_k - x'| \rightarrow 0 \ (k \rightarrow +\infty)$, т. е. $y' = 0 \in H(z')$. Замкнутость отображения $H(z)$, таким образом, доказана.

Следствие 1. В предположениях леммы 3.2.1 отображение

$$\varphi_f(x, s) = x - \lambda H(x, s), \quad \lambda \in (0; 2), \quad (3.2.6)$$

является замкнутым по $[x, s]$ на $\mathbb{R}^n \times S$ и $M_f(s)$ -фейеровским.

Следствие 2. Если для каждой функции $f_j(x, s)$ системы неравенств (3.2.1) выполнены условия леммы (3.2.1) и существует \bar{x} такой, что

$$\sup_{(i), s \in S} f_j(\bar{x}, s) \leq 0, \quad (3.2.7)$$

то отображение

$$\varphi(x, s) = \left\{ x - \lambda \sum_{j=1}^m \alpha_j \frac{f_j^+(x, s)}{|h_j|^2} h_j : h_j \in \partial_x f_j(x, s) \right. \\ \left. (\forall j \in \mathbb{N}_m) \right\} \quad (3.2.8)$$

является замкнутым по $[x, s] \in \mathbb{R}^n \times S$ (и в силу результатов § 1 гл. II $M(s)$ -фейеровским); выше $\lambda \in (0; 2)$,

$\sum_{j=1}^m \alpha_j = 1, \alpha_j > 0 \ (\forall j \in \mathbb{N}_m)$. В (3.2.8) j -е слагаемое полагается равным 0 при $h_j = 0$.

Действительно, если каждой функции $f_j(\mathbf{x}, \mathbf{s})$ поставить в соответствие $H_j(\mathbf{z})$ (по аналогии с $H(\mathbf{z})$ для $f(\mathbf{x}, \mathbf{s})$) и $\varphi_j(\mathbf{x}, \mathbf{s}) = \mathbf{x} - \lambda_j H_j(\mathbf{z})$, $\lambda_j \in (0; 2)$, то $\varphi(\mathbf{x}, \mathbf{s}) = \sum_{j=1}^m \alpha_j \varphi_j(\mathbf{x})$, при этом в силу леммы 3.2.1 отображение $\varphi_j(\mathbf{x}, \mathbf{s})$ замкнуто ($\forall j \in N_m$). Отсюда замкнутость отображения $\varphi(\mathbf{x}, \mathbf{s})$ будет следовать очевидным образом, если учесть, что в ситуации

$$\begin{aligned} \{\mathbf{z}_k\} \rightarrow \mathbf{z}', \quad \mathbf{y}_k = \sum_{j=1}^m \alpha_j \mathbf{y}_k^j \rightarrow \mathbf{y}' \quad (k \rightarrow +\infty), \\ \mathbf{y}_k \in \varphi(\mathbf{x}_k, \mathbf{s}_k), \\ \mathbf{y}_k^j \in \varphi_j(\mathbf{x}_k, \mathbf{s}_k), \quad \forall k, \quad \forall j \in N_m, \end{aligned}$$

каждая из последовательностей $\{\mathbf{y}_k^j\}_k$ ($\forall j \in N_m$) будет ограниченной. Последнее вытекает из выполнимости неравенств $|\mathbf{y}_k^j - \bar{\mathbf{x}}| \leq |\mathbf{x}_k - \bar{\mathbf{x}}|$, $\forall k$; здесь $\bar{\mathbf{x}}$ — из (3.2.7).

Замечание 1. Из доказательства леммы 3.2.1 получается такой ее ослабленный вариант.

Пусть $f(\mathbf{x}, \mathbf{s})$ — непрерывная на $\mathbf{R}^n \in S$ по $\mathbf{z} = [\mathbf{x}, \mathbf{s}]$ и выпуклая по \mathbf{x} для каждого $\mathbf{s} \in V$ функция. Если $\{\mathbf{z}_k\} \rightarrow \mathbf{z}' = [\mathbf{x}', \mathbf{s}']$, $H(\mathbf{z}_k) \ni \mathbf{y}_k \rightarrow \mathbf{y}'$ ($k \rightarrow +\infty$), $M(\mathbf{s}') \neq \emptyset$, то $\mathbf{y}' \in H(\mathbf{z}')$.

Перечислим ряд условий на систему (3.1.1), которые будут фигурировать ниже в теоремах:

1'. $\bar{V} \ni \{\mathbf{s}_t\} \rightarrow \mathbf{s}'$.

2'. $f_j(\mathbf{x}, \mathbf{s})$ ($\forall j \in N_m$) непрерывны по $[\mathbf{x}, \mathbf{s}] \in \mathbf{R}^n \times V$ и выпуклы по \mathbf{x} на \mathbf{R}^n .

3'. $M_\infty = \bigcup_{(i) k \geq t} \cap M_k \neq \emptyset$; здесь $M_k = M(\mathbf{s}_k)$ — множество решений системы (3.1.1) при $\mathbf{s} = \mathbf{s}_k$.

3''. M_∞ телесно.

Пусть $g(y_1, \dots, y_m)$ — выпуклая на \mathbf{R}^m функция, обладающая свойствами:

1) $g(\mathbf{0}) = 0$;

2) $g(\mathbf{y}) > 0$, $\forall \mathbf{y} \in \mathbf{R}_+^m$, $\mathbf{y} \neq \mathbf{0}$.

Образую функцию

$$f(\mathbf{x}, \mathbf{s}) = g(f_1^+(\mathbf{x}, \mathbf{s}), \dots, f_m^+(\mathbf{x}, \mathbf{s}))$$

и отображение

$$\bar{\varphi}(\mathbf{x}, \mathbf{s}) = \left\{ \mathbf{x} - \lambda \frac{f(\mathbf{x}, \mathbf{s})}{|h|^2} \mathbf{h} : \mathbf{h} \in \partial_x f(\mathbf{x}, \mathbf{s}) \right\}. \quad (3.2.9)$$

Теорема 3.2.1. Если выполнены условия 1'—3', то процесс $\{x_{t+1} \in \tilde{\varphi}(x_t, s_t)\}_{\forall t}$, порожденный отображением (3.2.9), ограничен и]

$$\{x_t\}' \subset M(s').$$

Доказательство. Для функции $f(x, s)$ выполнены условия леммы 3.2.1 при $S = \{s_t\} \cup \{s'\}$, а потому, согласно ее следствию 1, отображение $\tilde{\varphi}(x, s)$ будет замкнутым по $z = [x, s]$ на $R^n \times S$ и $M(s)$ -фейеровским. Остается воспользоваться теоремой 3.1.1.

Теорема 3.2.2. Если выполнены условия 1', 2' и 3'', то процесс $\{x_{t+1} \in \tilde{\varphi}(x_t, s_t)\}_{\forall t}$, порожденный отображением (3.2.9), сходится к элементу из $M(s')$, т. е.

$$\{x_t\} \rightarrow x' \in M(s').$$

Справедливость теоремы вытекает из того, что для данной ситуации выполняются условия следствия 1 теоремы 3.1.1.

Теорема 3.2.3. Пусть выполнены условия 1'—3'. Тогда процесс $\{x_{t+1} \in \varphi(x_t, s_t)\}_{\forall t}$, порожденный отображением (3.2.8), ограничен и

$$\{x_t\}' \subset M(s').$$

Обоснование является почти тем же, что и для теоремы 3.2.1: нужно лишь вместо f подставить f_j , а затем учесть факт замкнутости по $z = [x, s]$ выпуклой комбинации

$\sum_{j=1}^m \alpha_j \varphi_j(x, s)$ замкнутых отображений

$$\varphi_j(x, s) = \left\{ x - \lambda \frac{f_j'(x, s)}{\|h\|^2} h : h \in \partial_x f(x, s) \right\} \quad (3.2.10)$$

$$(\forall j \in N_m), \quad \lambda \in (0; 2).$$

Образуем из отображений (3.2.10) отображение

$$\varphi'(x, s) = \varphi_1(\dots \varphi_m(x, s)\dots, s). \quad (3.2.11)$$

Так как отображение $\varphi'(x, s)$ замкнуто по $z = [x, s]$, то для него можно формулировать аналоги теорем 3.2.1—3.2.3 при выполнении, естественно, соответствующих условий. Например, аналогом теоремы 3.2.2 служит

Теорема 3.2.4. Если выполнены условия 1', 2' и 3'', то процесс

$$\{x_{t+1} \in \varphi'(x_t, s_t)\}_{\forall t}$$

сходится к элементу из $M(s')$.

Применительно к изучаемым процессам рассмотрим возможность применения теоремы 3.1.3. Цель этого — проиллюстрировать возможность обеспечения условия 4° из этой теоремы.

Теорема 3.2.5. Пусть выполнены условия 1', 2', а также: $M(s')$ ограничено и существует вектор $p \in \mathbb{R}^n$ такой, что $f_j(p, s') < 0$ ($\forall j \in N_m$). Тогда, если ψ — любое из отображений (3.2.8), (3.2.9) и (3.2.11), то процесс

$$\{x_{t+1} \in \psi(x_t, s_t)\}_{\forall t}$$

сходится к некоторому элементу из $M(s')$.

При обосновании этого утверждения действительно достаточно сослаться на теорему 3.1.3, учтя, что из принятых предположений условие 4° вытекает.

2.2. Решение бесконечных систем выпуклых неравенств. Здесь будет рассмотрен итерационный метод решения бесконечной системы выпуклых неравенств

$$f_j(x) \leq 0 \quad (\forall j \in \mathbb{N}) \quad (3.2.12)$$

с точки зрения организации нестационарного процесса.

Пусть

$$s' = [\alpha_1, \alpha_2, \dots, \alpha_t, \dots], \quad \sum_{i=1}^{+\infty} \alpha_i = 1, \quad \alpha_t > 0, \quad \forall t;$$

$$s_t = [\alpha_1, \dots, \alpha_t], \quad \delta_t = \sum_{i=1}^t \alpha_i.$$

Исходя из

$$\varphi_i(x) = \left\{ x - \lambda_i \frac{f_i^+(x)}{|h|^2} h : h \in \partial f_i(x) \right\}$$

$$(\forall i \in \mathbb{N}),$$

построим

$$\varphi(x, s_t) = \frac{1}{\delta_t} \sum_{i=1}^t \alpha_i \varphi_i(x) \quad \dots (3.2.13)$$

и

$$\varphi(\mathbf{x}, \mathbf{s}') = \sum_{i=1}^{+\infty} \alpha_i \varphi_i(\mathbf{x}); \quad (3.2.14)$$

здесь $\lambda_i \in [\varepsilon; 2 - \varepsilon]$, $0 < \varepsilon < 1$.

Сделаем замечание по поводу смысла $\varphi(\mathbf{x}, \mathbf{s}')$. Если положить

$$\mu_i(\mathbf{x}, \mathbf{h}) = \frac{f'_i(\mathbf{x})}{|\mathbf{h}^2|} \mathbf{h},$$

то по аналогии с (3.2.5) будем иметь

$$|\mu_i(\mathbf{x}, \mathbf{h})| \leq |\mathbf{x} - \mathbf{y}|,$$

$\forall \mathbf{h} \in \partial f_i(\mathbf{x}), \forall \mathbf{y} \in M = \bigcap_{i=1}^{\infty} M_i \neq \emptyset$, где $M_i = \{\mathbf{x} : f_i(\mathbf{x}) \leq 0\}$.

Отсюда вытекает условие

$$\sup \{ \|\mathbf{u}\| : \mathbf{u} \in \varphi_i(\mathbf{x}), \mathbf{x} \in S; \forall i \in \mathbb{N} \} < +\infty \quad (3.2.15)$$

при любом ограниченном $S \subset \mathbb{R}^n$. Это приводит к абсолютной сходимости ряда $\sum_{i=1}^{+\infty} \alpha_i \mathbf{u}_i$, как бы ни выбирались $\mathbf{u}_i \in \varphi_i(\mathbf{x})$. Этим определен смысл соотношения (3.2.14).

Обратим внимание на то, что множествами неподвижности для (3.2.13) и (3.2.14) являются соответственно множества M_t и M ; при этом $\varphi(\cdot, \mathbf{s}_t) \in F_{M_t}, \forall t$; $\varphi(\cdot, \mathbf{s}') \in F_M$.

Лемма 3.2.2. Из $\{\mathbf{x}_t\} \rightarrow \mathbf{x}', \{\mathbf{z}_t\} \rightarrow \mathbf{z}', \mathbf{z}_t \in \varphi(\mathbf{x}_t, \mathbf{s}_t), \forall t$, следует, что

$$\mathbf{z}' \in \varphi(\mathbf{x}', \mathbf{s}'). \quad (3.2.16)$$

Доказательство. В силу условия $\mathbf{z}_t \in \varphi(\mathbf{x}_t, \mathbf{s}_t)$ имеем

$$\mathbf{z}_t = \frac{1}{\delta_t} \sum_{i=1}^t \alpha_i \mathbf{y}_i^t, \mathbf{y}_i^t \in \varphi_i(\mathbf{x}_t), \forall t.$$

Определим последовательность $\{\mathbf{y}_i^t\}$ индуктивно: \mathbf{y}_1^t — некоторая предельная точка последовательности $\{\mathbf{y}_1^t\}$; если уже определены $\mathbf{y}_1^t, \dots, \mathbf{y}_k^t$, причем $[\mathbf{y}_1^{kt}, \dots, \mathbf{y}_k^{kt}] \rightarrow [\mathbf{y}_1^t, \dots, \mathbf{y}_k^t]$ ($t \rightarrow +\infty$), то в качестве \mathbf{y}_{k+1}^t берется предельная точка последовательности $\{\mathbf{y}_{k+1}^{kt}\}_t$. В силу

замкнутости $\varphi_i(\mathbf{x})$ справедливы включения $\mathbf{y}'_i \in \varphi_i(\mathbf{x}')$, $\forall i$. Если теперь доказать, что $\mathbf{z}' = \sum_{i=1}^{+\infty} \alpha_i \mathbf{y}'_i$, то это и будет означать справедливость (3.2.16). В силу (3.2.15) для произвольного $\varepsilon > 0$ можно указать i_0 и T такие, что

$$\left| \mathbf{z}_t - \frac{1}{\delta_t} \sum_{i=1}^{i_0} \alpha_i \mathbf{y}_i^t \right| \leq \varepsilon, \quad \forall t \geq T.$$

Отсюда $\left| \mathbf{z}' - \sum_{i=1}^{i_0} \alpha_i \mathbf{y}'_i \right| \leq \varepsilon$. Следовательно, равенство $\mathbf{z}' = \sum_{i=1}^{+\infty} \alpha_i \mathbf{y}'_i$, а вместе с тем и (3.2.16), доказано.

Из свойств $\varphi(\cdot, \mathbf{s}_t) \in F_{M_t}$, $\forall t$, (3.2.16), а также $M = \bigcap_{(t)} M_t \neq \emptyset$ вытекает

Теорема 3.2.6. Если система выпуклых неравенств (3.2.12) совместна, то процесс

$$\{\mathbf{x}_{t+1} \in \varphi(\mathbf{x}_t, \mathbf{s}_t)\}_{\forall t}$$

сходится к некоторому ее решению; здесь $\varphi(\mathbf{x}, \mathbf{s}_t)$ задано согласно (3.2.13).

По существу, сформулированный факт является конкретизацией теоремы 3.1.1, которая дает условие $\{\mathbf{x}_t\}' \subset M$. Но так как в рассматриваемом случае имеет место неравенство $|\mathbf{x}_{t+1} - \mathbf{y}| \leq |\mathbf{x}_t - \mathbf{y}|$, $\mathbf{y} \in M$, $\forall t$, то $\{\mathbf{x}_t\} \rightarrow \mathbf{x}' \in M$.

З а м е ч а н и е 2. Если вместо $\varphi(\mathbf{x}, \mathbf{s}_t)$ в 3.2.13 взять отображение

$$\varphi_0(\mathbf{x}, \mathbf{s}_t) = \left\{ \mathbf{x} - \sum_{i=1}^t \alpha_i \lambda_i \frac{f_i^+(\mathbf{x})}{|h_i|} \mathbf{h}_i : \mathbf{h}_i \in \partial f_i(\mathbf{x}), \quad \forall i \in N_t \right\},$$

то процесс из теоремы 3.2.6, отнесенной к этому отображению, будет также сходиться к решению системы (3.2.12), при этом доказательство остается прежним.

Рассмотрим теперь вместо системы (3.2.12) совместную систему

$$f_j(\mathbf{x}) \leq 0 \quad (\forall j \in N), \quad \mathbf{x} \in C, \quad (3.2.17)$$

относительно которой предположим: функции, стоящие в левых частях системы, выпуклы и дифференцируемы,

C — выпуклое и замкнутое множество. Образует для нее отображения

$$\bar{\varphi}(\mathbf{x}, \mathbf{s}_t) = \pi_C \left(\mathbf{x} - \sum_{i=1}^t \alpha_i \lambda_i \frac{f_i'(\mathbf{x})}{|\nabla f_i(\mathbf{x})|^2} \nabla f_i(\mathbf{x}) \right) \quad (3.2.18)$$

и

$$\bar{\varphi}(\mathbf{x}, \mathbf{s}) = \pi_C \left(\mathbf{x} - \sum_{i=1}^{+\infty} \alpha_i \lambda_i \frac{f_i'(\mathbf{x})}{|\nabla f_i(\mathbf{x})|^2} \nabla f_i(\mathbf{x}) \right). \quad (3.2.19)$$

Предположения относительно $\{\alpha_t\}$ и $\{\lambda_t\}$ прежние.

Теорема 3.2.7. В сделанных относительно системы (3.2.17) предположениях процесс

$$\{\mathbf{x}_{t+1} \in \bar{\varphi}(\mathbf{x}_t, \mathbf{s}_t)\}_{\forall t}$$

сходится к некоторому ее решению.

Это утверждение можно было бы доказать по аналогии с предыдущей теоремой даже при более слабых условиях, а именно, без предположения дифференцируемости функций $\{f_j(\mathbf{x})\}$ (тогда, естественно, выражение, стоящее в (3.2.18) в круглых скобках, нужно было бы заменить его точечно-множественным вариантом), однако мы его получим в качестве следствия из следующей общей теоремы.

Теорема 3.2.8. Пусть $A(\mathbf{x})$ — непрерывный M -фейеровский оператор, $A_t(\mathbf{x})$ — M_t -фейеровский ($\forall t$); при этом последовательность $\{A_t(\mathbf{x})\}$ сходится к $A(\mathbf{x})$ равномерно на любом ограниченном множестве из \mathbb{R}^n и $\bigcap_{(t)} M_t = M \neq \emptyset$.

Тогда процесс

$$\{\mathbf{x}_{t+1} = A_t(\mathbf{x}_t)\}_{\forall t}$$

сходится к элементу из M .

Доказательство. Будем считать $\{\mathbf{x}_t\} \cap M = \emptyset$. Если из $\{\mathbf{x}_t\}$ удалить возможные повторения, то получим M -фейеровскую последовательность. Пусть ею будет сама $\{\mathbf{x}_t\}$. Положим $B_t(\mathbf{x}) = A(\mathbf{x}) - A_t(\mathbf{x})$. Так как последовательность $\{\mathbf{x}_t\}$ ограничена, то условия теоремы гарантируют, что

$$|B_t(\mathbf{x}_t)| \rightarrow 0 \quad (t \rightarrow +\infty).$$

Пусть $\{\mathbf{x}_{i_t}\} \rightarrow \mathbf{x}'$, тогда из условия

$$\mathbf{x}_{i_t+1} = A(\mathbf{x}_{i_t}) - B_{i_t}(\mathbf{x}_{i_t})$$

получаем: $\{x_{t+1}\} \rightarrow x''$ и $x'' = A(x')$. Если $x' \in M$, то $\{x_t\} \rightarrow x'$, в противном случае $|x'' - y| < |x' - y|$, $y \in M$, но это противоречиво, так как $|x - y| = \text{const}$, $\forall x \in \{x_t\}'$. Теорема доказана.

Если теперь в качестве $A(x)$ взять (3.2.19), а в качестве $A_t(x)$ — (3.2.8), то можно без труда проверить, что все условия теоремы 3.2.8 выполняются. Это и дает справедливость теоремы 3.2.7. При проверке условий необходимо учесть, что дифференцируемая выпуклая функция является и непрерывно дифференцируемой (см., например, [36]).

§ 3. Сходимость основного стандартного процесса

Выпишем процесс

$$\{x_{t+1} \in \varphi(x_t, s_t) + \alpha_t \partial f_0(x_t)\}_{\forall t}, \quad (3.3.1)$$

поставленный в соответствие нестационарной модели выпуклого программирования

$$\max \{f_0(x) : f_j(x, s_t) \leq 0, \forall j \in N_m\}, \quad (3.3.2)$$

где $x \in \mathbb{R}^n$, $\{s_t\} \subset V \subset \mathbb{R}^l$, $\{-f_0(x), f_j(x), \forall j \in N_m\}$ — выпуклые на \mathbb{R}^n функции для всех $s \in V$; $\alpha_t \geq 0, \forall t$, $\partial f_0(x)$ — множество опорных линейных функционалов к функции $f_0(x)$ в точке x , $\varphi(x, s)$ — $M(s)$ -фейеровское отображение.

1 3.1. Условия ограниченности процесса. Вначале будет изучаться процесс (3.3.1), в котором конструкция отображения $\varphi(\cdot, s) \in F_{M(s)}$ непосредственно не связывается с моделью (3.3.2), функция $f_0(x)$ всюду ниже предполагается вогнутой.

При рассмотрении (3.3.1) будут в разных сочетаниях использоваться следующие условия:

1°. $V \supset \{s_t\} \rightarrow s'$;

2°. $F_{M(s)} \ni \varphi(\cdot, s)$ — замкнуто по $[x, s]$ на $\mathbb{R}^n \times V$, здесь $M(s)$ — множество неподвижности для $\varphi(x, s)$;

3°. $M(s') = M$ — ограничено (s' — из 1°);

4°. $|M_t - M|_+ = \varepsilon_t \rightarrow 0$ ($t \rightarrow +\infty$), здесь $M_t = M(s_t)$;

5°. $\sum_{(t)} \varepsilon_t < +\infty$;

6°. *Условие регулярного сжатия*: для любых $R > \delta > 0$ существует $\varepsilon > 0$ такое, что

$$|z - M| \leq |x - M| - \varepsilon, \quad \forall x \in M^R \setminus M^\delta, \quad \forall z \in \varphi(x),$$

здесь $\varphi \in \mathbf{F}_M$, $M^R = \{x : |x - M| \leq R\}$;

$$7^\circ. \{\alpha_t \geq 0\} \rightarrow 0, \quad \sum_{(t)} \alpha_t = +\infty.$$

Повсюду ниже под условиями 1°—7° будут пониматься именно перечисленные выше условия.

Лемма 3.3.1. Если выполнены условия 1°—4°, то из выполнимости для $\varphi(\cdot, s') \in \mathbf{F}_M$ условия 6° вытекает его выполнимость и для $\varphi(\cdot, s_t)$ для всех t , начиная с некоторого.

Доказательство. Предположив противное, можно указать последовательности $\{x_t\} \rightarrow x'$, $\{z_t\} \rightarrow z'$ и $\{\beta_t \geq 0\} \rightarrow 0$ такие, что

$$\begin{aligned} \{x_t\} &\subset M^R \setminus M^\delta, \quad z_t \in \varphi(x_t, s_t), \quad \forall t, \\ |z_t - M_t| &= |x_t - M_t| - \varepsilon_t. \end{aligned}$$

Перейдя в этом соотношении к пределу (а это возможно в силу сходимости M_t к M по Хаусдорфу), получим $|z' - M| = |x' - M|$. Но так как $x' \notin M^\delta$ и $z' \in \varphi(x', s')$, то $|z' - M| < |x' - M|$, т. е. получилось противоречие. Лемма доказана.

Лемма 3.3.2. Если выполнены условия 1°—3°, то для $\varphi(\cdot, s') \in \mathbf{F}_M$ и $0 < \delta < R$ существует $\varepsilon > 0$ такое, что для $x \in M^R \setminus M^\delta$ и $z \in \varphi(x, s')$ выполнено соотношение

$$|z - M| \leq |x - M| - \varepsilon.$$

Действительно, из замечания к лемме 3.1.3 следует существование $\theta_0 \in (0, 1)$ такого, что

$$|z - M| \leq \theta_0 |x - M|,$$

а потому доказываемое соотношение выполняется при $\varepsilon = \delta(1 - \theta_0)$.

Выясним условия ограниченности процесса (3.3.1).

Так как, в силу леммы 2.2.2,

$$\sup \{ \|h\| : h \in \partial f_0(x), x \in S \} < +\infty$$

при любом ограниченном $S \subset \mathbf{R}^n$, то при условии 3°

$$d = \sup \{ \|h\| : h \in \partial f_0(x), x \in M^R \} < +\infty, \quad (3.3.3) \\ R > 0.$$

Лемма 3.3.3. Пусть выполнены условия 1°, 3°, 4° и 6°. Тогда при $0 < r < R$ можно утверждать существование t' такого, что при

$$0 < \alpha < \min \left\{ \frac{R-r}{d}, \frac{\varepsilon}{d} \right\} \quad (3.3.4)$$

справедливо включение

$$\bigcup_{x \in M^R, t \geq t'} \{ \varphi(x, s_t) + \alpha \partial f_0(x) \} \subset M^R; \quad (3.3.5)$$

здесь $M = M(s')$, d — из (3.3.3), ε — из условия 6°.

Доказательство. Выберем число t' так, чтобы оно удовлетворяло лемме 3.3.1, а также условию

$$\alpha \leq \min \left\{ \frac{R-r-2\varepsilon_t}{d}, \frac{\varepsilon-2\varepsilon_t}{d} \right\}, \quad \forall t \geq t'.$$

Возможность этого вытекает из (3.3.4) и 4°. Пусть $t \geq t'$ и z принадлежит левой части соотношения (3.3.5). Рассмотрим два случая:

1) если $x \in M^r$, то

$$\begin{aligned} |z - M| &\leq |z' - M_t| + \varepsilon_t + \alpha d \leq |x - M_t| + \varepsilon_t + \alpha d \leq \\ &\leq |x - M| + 2\varepsilon_t + \alpha d \leq r + 2\varepsilon_t + d \frac{R-r-2\varepsilon_t}{d} = R, \end{aligned}$$

т. е. $z \in M^R$; выше $z' \in \varphi(x, s_t)$;

2) пусть $x \in M^R \setminus M^r$, тогда

$$\begin{aligned} |z - M| &\leq |x - M| + 2\varepsilon_t - \varepsilon + \alpha d \leq \\ &\leq R + 2\varepsilon_t - \varepsilon + d \frac{\varepsilon-2\varepsilon_t}{d} = R, \end{aligned}$$

т. е. $z \in M^R$.

Лемма доказана.

Следствие 1. Если выполнены условия леммы 3.3.3, то при достаточно малом $\bar{\alpha} = \sup_{(t)} \alpha_t > 0$ процесс (3.3.1) ограничен.

Следствие 2. Пусть выполнены условия 1°—4°. Тогда, каково бы ни было $x_0 \in \mathbb{R}^n$, при достаточно малом $\bar{\alpha} = \sup_{(t)} \alpha_t > 0$ процесс (3.3.1) ограничен.

Действительно, если в лемме 3.3.3 вместо условия 6° предположить выполненными свойства 2° и 3°, то этим, согласно лемме 3.3.2, обеспечивается выполнимость 6°

для $\mathbf{x} \in M^R \setminus M^r$, где R выбрано таким, чтобы обеспечить включение $\mathbf{x}_0 \in M^R$. Но тогда (3.3.5) выполняется в силу тех же выкладок, которые обосновывают лемму 3.3.3. Этим показана и справедливость доказываемого утверждения.

3.2. Теорема о существовании неподвижной точки итерационного оператора. Условия леммы 3.3.3 (в другом варианте: условия следствия 2 к ней) делают справедливым включение (3.3.5). В связи с этим включением рассмотрим вопросы о существовании неподвижных точек отображений

$$\Phi(\mathbf{x}, \mathbf{s}_t; \alpha) = \varphi(\mathbf{x}, \mathbf{s}_t) + \alpha \partial f_0(\mathbf{x}), \quad (3.3.6)$$

а также их близости к оптимальному множеству \tilde{M} задачи

$$\max \{f_0(\mathbf{x}) : \mathbf{x} \in M = M(\mathbf{s}')\} (= \bar{f}_0). \quad (3.3.7)$$

Пусть \tilde{M}_t — оптимальное множество задачи

$$\max \{f_0(\mathbf{x}) : \mathbf{x} \in M_t = M(\mathbf{s}_t)\}. \quad (3.3.8)$$

Лемма 3.3.4. Из выпуклости, замкнутости, ограниченности множеств $M, M_t, \forall t$, и свойства 4° вытекает, что

$$|\tilde{M}_t - \tilde{M}| \rightarrow 0 \quad (t \rightarrow +\infty),$$

где $\tilde{M} = \text{Arg}(3.3.7)$, $\tilde{M}_t = \text{Arg}(3.3.8)$.

Теорема 3.3.1. Пусть выполнены условия 1°—4° и отображения $\varphi(\mathbf{x}, \mathbf{s}_t)$ непрерывны по $\mathbf{x}, \forall t$. Тогда существуют $\bar{\alpha} > 0$ и t' такие, что для всех $t \geq t'$ и $0 < \alpha \leq \bar{\alpha}$ отображение (3.3.6) обладает неподвижной точкой $\mathbf{x}_t(\alpha)$, т. е.

$$\mathbf{x}_t(\alpha) \in \Phi(\mathbf{x}_t(\alpha), \mathbf{s}_t; \alpha); \quad (3.3.9)$$

при этом

$$\lim_{t \rightarrow +\infty} \lim_{\alpha \rightarrow 0} |\mathbf{x}_t(\alpha) - \tilde{M}| = 0. \quad (3.3.10)$$

Доказательство. Условия сформулированного утверждения гарантируют справедливость (3.3.5). В качестве $\bar{\alpha} > 0$ можно взять любое число, удовлетворяющее соотношению (3.3.4). Существование неподвижной точки $\mathbf{x}_t(\alpha)$ для $\Phi(\mathbf{x}, \mathbf{s}_t; \alpha)$ следует из теоремы Какутани

с учетом того, что отображение $x \rightarrow \partial f_0(x)$ является замкнутым, а $\partial f_0(x)$ — выпуклым и замкнутым множеством [93].

Перейдем к доказательству (3.3.9). Заметим, что любая последовательность $\{x_k = x_t(\alpha_k)\}_k$, сгенерированная в соответствии с (3.3.9) при фиксированном t и $\{\alpha_k > 0\} \rightarrow 0$, ограничена.

Докажем, что $|x_k - \tilde{M}_t| \rightarrow 0$ ($k \rightarrow +\infty$). Предположив противное, можно исходить из $\{x_k\} \rightarrow x' \notin \tilde{M}_t$, что ведет к соотношениям

$$f_0(x_k) < f_0(\tilde{x}_t) - \delta, \quad \forall k \geq k';$$

здесь $\delta > 0$, $\tilde{x}_t \in \tilde{M}_t$. Так как

$$x_k = \varphi(x_k, s_t) + \alpha_k h_k, \quad h_k \in \partial f_0(x_k),$$

то имеем

$$\begin{aligned} |x_k - \tilde{x}_t|^2 &= (\varphi(x_k, s_t) + \alpha_k h_k, x_k - \tilde{x}_t) \leq \\ &\leq |\varphi(x_k, s_t) - \tilde{x}_t| \cdot |x_k - \tilde{x}_t| + \alpha_k (h_k, x_k - \tilde{x}_t) \leq \\ &\leq |x_k - \tilde{x}_t|^2 + \alpha_k (f_0(x_k) - f_0(\tilde{x}_t)) \leq \\ &\leq |x_k - \tilde{x}_t|^2 - \alpha_k \delta, \end{aligned}$$

т. е. $0 < -\alpha_k \delta$, что дает противоречие.

Из полученного условия $|x_k - \tilde{M}_t| \rightarrow 0$ ($k \rightarrow +\infty$), леммы 3.3.4 и неравенства $|x_t(\alpha) - \tilde{M}| \leq |x_t(\alpha) - \tilde{M}_t| + |\tilde{M}_t - M|$ доказываемое соотношение (3.3.10) следует теперь очевидным образом.

Общее замечание. Варианты приводимых выше условий, гарантирующих ограниченность процесса (3.3.1), достаточны и для ограниченности процесса, заданного его модификацией

$$\{x_{t+1} \in \varphi(x_t + \alpha_t \partial f_0(x), s_t)\}_{\forall t} \quad (3.3.11)$$

для случая $f_0(x) = (c, x)$.

3.3. Теорема о сходимости основного процесса. Свойства процесса (3.3.1), т. е. процесса

$$\{x_{t+1} \in \varphi(x_t, s_t) + \alpha_t \partial f_0(x_t)\}_{\forall t}, \quad (3.3.12)$$

существенно зависят от неявного параметра ε_t из условия 4° и параметра $\alpha_t > 0$. И хотя для того и другого являются естественными условия $\{\alpha_t\} \rightarrow 0$ и $\{\varepsilon_t\} \rightarrow 0$,

однако внутренний смысл требования к процессу (3.3.1), обеспечивающего его сходимость, таков, что стремление $\{\alpha_t\} \rightarrow 0$ должно быть достаточно «медленным», а $\{\varepsilon_t\} \rightarrow 0$ — достаточно «быстрым». Это можно выразить в форме условий 5° и 7°.

Лемма 3.3.5. Пусть выполнены условия 1°—4° и $\{\alpha_t \geq 0\} \rightarrow 0$. Тогда при произвольном $x_0 \in \mathbb{R}^n$ и достаточно малом $\bar{\alpha} = \sup_{(t)} \alpha_t$ для процесса (3.3.1) справедливо утверждение

$$|x_t - M| \rightarrow 0 \quad (t \rightarrow +\infty). \quad (3.3.13)$$

Доказательство. Согласно следствию 2 к лемме 3.3.3 существование $\bar{\alpha} > 0$ и ограниченность последовательности гарантированы. Пусть

$$\{x_{j_t+1}\} \rightarrow x'' = \arg \max \{|x - M| : x \in \{x_t\}'\}.$$

Можно считать, что $\{x_{j_t}\} \rightarrow x'$. Так как отображение

$$\Phi(x, s; \alpha) = \varphi(x, s) + \alpha df_0(x)$$

замкнуто по $[x, s, \alpha]$ на $\mathbb{R}^n \times V \times [0; 1] \subset \mathbb{R}^{n+t+1}$, то $x'' \in \Phi(x', s'; 0) = \varphi(x', s')$. Если $x' \in M$, то $\varphi(x', s') = x' = x''$, следовательно, $\{x_t\}' \subset M$, т. е. (3.3.13) справедливо. Пусть $x' \notin M$. Из $|x'' - y| < |x' - M|$, $\forall y \in M$, следует неравенство $|x'' - M| \leq |x'' - y|$. В силу способа выбора x'' имеем $|x' - M| = |x'' - M|$, но тогда, согласно лемме 3.3.1, существует $\bar{y} \in M$ такой, что $|x' - \bar{y}| = |x'' - \bar{y}|$, но это противоречиво. Лемма доказана.

Теорема 3.3.2. Пусть выполнены условия 1°—3°, 5° и 7°. Тогда при произвольном $x_0 \in \mathbb{R}^n$ и достаточно малом $\bar{\alpha} = \sup_{(t)} \alpha_t$ для процесса (3.3.1) справедливо соотношение

$$|x_t - \tilde{M}| \rightarrow 0 \quad (t \rightarrow +\infty). \quad (3.3.14)$$

Доказательство. Во-первых, докажем, что $\{x_t\}' \cap \tilde{M} \neq \emptyset$. Предположив противное, можно указать $\delta' > 0$ такое, что

$$f_0(x_t) \leq f_0(\tilde{x}) - \delta', \quad \forall t \geq t'$$

(\tilde{x} — некоторый элемент из \tilde{M}). Пусть \tilde{x}_t — проекция \tilde{x} на M_t , тогда в силу 4° $|\tilde{x}_t - \tilde{x}| \leq \varepsilon_t \rightarrow 0$, поэтому при некотором $\delta_0 > 0$ будет выполняться соотношение

$$f_0(x_t) \leq f_0(\tilde{x}_t) - \delta_0, \quad \forall t \geq t_0. \quad (3.3.15)$$

Записав (3.3.12) в виде

$$x_{t+1} = y_t + \alpha_t h_t, \quad y_t \in \varphi(x_t, s_t), \quad h_t \in \partial f_0(x_t),$$

из (3.3.15) получим

$$(h_t, x_t - \tilde{x}_t) \leq f_0(x_t) - f_0(\tilde{x}_t) \leq -\delta_0. \quad (3.3.16)$$

Положим $r = \sup_{(t)} |h_t|^2 (< +\infty)$.

Ниже в оценочных выкладках будут использованы соотношения (3.3.13) и (3.3.16):

$$\begin{aligned} |x_{t+1} - \tilde{x}|^2 &= (y_t + \alpha_t h_t - \tilde{x}, y_t + \alpha_t h_t - \tilde{x}) \leq \\ &\leq |y_t - \tilde{x}|^2 + 2\alpha_t (h_t, y_t - \tilde{x}) + \alpha_t^2 r \leq (|y_t - \tilde{x}| + \varepsilon_t)^2 + \\ &\quad + 2\alpha_t (h_t, x_t - \tilde{x}_t + y_t - x_t + \tilde{x}_t - \tilde{x}) + \alpha_t^2 r \leq \\ &\leq (|x_t - \tilde{x}| + 2\varepsilon_t)^2 + \alpha_t^2 r + \\ &\quad + 2\alpha_t [-\delta_0 + \sqrt{r}(|x_t - M| + \varepsilon_t) + \sqrt{r}\varepsilon_t] \leq \\ &\leq [|x_t - \tilde{x}|^2 + 2\varepsilon_t]^2 - \alpha_t \delta, \quad \delta > 0 \end{aligned}$$

для достаточно больших t , пусть для $t \geq t'$. Отсюда следует

$$|x_{t+1} - \tilde{x}|^2 - |x_t - \tilde{x}|^2 \leq \varepsilon_t r' - \alpha_t \delta, \quad r' > 0. \quad (3.3.17)$$

Суммирование (3.3.17) очевидным образом приводит к противоречию с предположениями 5° и 7°. Следовательно, соотношение $\{x_t\}' \cap \tilde{M} \neq \emptyset$ доказано.

Докажем, что $\{x_t\}' \subset \tilde{M}$. Положим

$$\begin{aligned} P_+^t &= \{x: (h_t, x - x_t) + f_0(x_t) \geq f_0(\tilde{x})\}, \\ P_-^t &= \{x: (h_t, x - x_t) + f_0(x_t) < f_0(\tilde{x})\} \\ Q_+ &= \{x_t: x_{t+1} \in P_+^t\} (= \{x_{t_s}\}) \end{aligned}$$

и докажем $Q_+ \subset \tilde{M}$, где штрих над Q_+ означает производное множество. Пусть $x_t \in Q_+$, тогда

$$(h_t, x_{t+1} - x_t) + f_0(x_t) \geq f_0(\tilde{x}).$$

Отсюда

$$\begin{aligned} f_0(\tilde{\mathbf{x}}) &\leq f_0(\mathbf{x}_t) + \alpha_t |\mathbf{h}_t| + (\mathbf{h}_t, \mathbf{y}_t - \mathbf{x}_t) \leq \\ &\leq f_0(\mathbf{x}_t) + \alpha_t \sqrt{r} + 2\sqrt{r} [|\mathbf{x}_t - M| + \varepsilon_t]. \end{aligned}$$

Следовательно, любая предельная точка \mathbf{x}' из Q'_+ удовлетворяет неравенству $f_0(\tilde{\mathbf{x}}) \leq f_0(\mathbf{x}')$. Но так как $\mathbf{x}' \in M$, то $\mathbf{x}' \in \tilde{M}$. Отсюда, в частности, следует, что если $\{\mathbf{x}_{j_k}\} \rightarrow \mathbf{x}' \in \tilde{M}$ то, обозначив

$$\begin{aligned} \Delta_k &= \{\mathbf{x}_{j_k}, \dots, \mathbf{x}_{j_{k+1}-1}\} \\ \text{и} \quad p_k &= \max \{|\mathbf{x}_s - \tilde{M}| : \mathbf{x}_s \in \Delta_k \cap Q_+\}, \end{aligned}$$

будем иметь $p_k \rightarrow 0$ ($k \rightarrow +\infty$).

Пусть теперь $\mathbf{x}_{t+1} \in \Delta_k$ и $\mathbf{x}_{t+1} \notin P_+^t$. Так как $P_+^t \supset \{\mathbf{x} : f_0(\mathbf{x}) \geq f_0(\tilde{\mathbf{x}})\} \supset \tilde{M}$, то, очевидно, сдвиг $\mathbf{y}_t \rightarrow \mathbf{x}_{t+1}$, являясь фейеровским относительно P_+^t , будет фейеровским и относительно \tilde{M} , а потому $|\mathbf{x}_{t+1} - \mathbf{z}| \leq |\mathbf{y}_t - \mathbf{z}|$, $\forall \mathbf{z} \in \tilde{M}$. Сделаем оценку для $|\mathbf{x}_{t+1} - M|$. Пусть $\tilde{\mathbf{x}}_t$ — проекция \mathbf{x}_t на \tilde{M} , $\bar{\mathbf{x}}_t$ — проекция $\tilde{\mathbf{x}}_t$ на M_t , тогда $|\tilde{\mathbf{x}}_t - \bar{\mathbf{x}}_t| \leq \varepsilon_t$. Имеем

$$\begin{aligned} |\mathbf{x}_{t+1} - \tilde{M}| &\leq |\mathbf{x}_{t+1} - \tilde{\mathbf{x}}_t| \leq |\mathbf{y}_t - \tilde{\mathbf{x}}_t| \leq \\ &\leq |\mathbf{y}_t - \bar{\mathbf{x}}_t| + \varepsilon_t \leq |\mathbf{x}_t - \bar{\mathbf{x}}_t| + \varepsilon_t \leq \\ &\leq |\mathbf{x}_t - \tilde{\mathbf{x}}_t| + 2\varepsilon_t = |\mathbf{x}_t - \tilde{M}| + 2\varepsilon_t \leq \dots \leq \\ &\leq \max \left\{ p_k + 2 \sum_{i=s}^t \varepsilon_i, |\mathbf{x}_{j_k} - \tilde{M}| + 2 \sum_{i=j_k}^t \varepsilon_i \right\} \end{aligned}$$

где s — наименьший индекс из $[j_k, \dots, t]$, для которого $\mathbf{x}_{s+1} \in P_+^s$. Последняя оценка вместе с уже доказанным свойством $p_k \rightarrow 0$ ($k \rightarrow +\infty$) приводит к окончательному заключению:

$$\{\mathbf{x}_t\}' \subset \tilde{M},$$

что эквивалентно (3.3.14).

З а м е ч а н и е 1. Приведенное доказательство сходимости (3.3.14) показывает, что при его реализации использованы ограниченность процесса (3.3.1) и свойство (3.3.13), обеспеченные предположениями теоремы 3.3.2.

Так что, если в самом начале предположить выполненными эти два свойства (или другие свойства, влекущие выполнимость этих), то для сходимости процесса (3.3.1) к оптимальному множеству \bar{M} предельной задачи $\max \{f_0(\mathbf{x}): \mathbf{x} \in M\}$ достаточны условия 5° и 7° (естественно, при $\varphi(\cdot, \mathbf{s}_t) \in F_{M_t}$). Это замечание полезно тем, что дает возможность конструировать различные варианты исходных условий, влекущих (3.3.14). \blacksquare

З а м е ч а н и е 2. Условие 3° ограниченности $M = M(\mathbf{s}')$ можно было бы ослабить, заменив его, например, на условие ограниченности \bar{M} . Но для этого понадобилось бы реконструировать отображения $\varphi(\mathbf{x}, \mathbf{s}_t)$, заменив их на $\varphi_\pi(\mathbf{x}, \mathbf{s}_t) = \pi(\varphi(\mathbf{x}, \mathbf{s}_t))$, где π — оператор проектирования на некоторый шар S , строго содержащий \bar{M} .

Предельной задачей для реконструированной задачи будет $\max \{f_0(\mathbf{x}): \mathbf{x} \in M_0 = M \cap S\}$; при этом M_0 ограничено. Что касается остальных исходных предположений, то они будут выполненными (здесь следует иметь в виду, что множества $\{M_t\}$ перейдут в $\{M_t \cap S\}$).

З а м е ч а н и е 3. В формулировке теоремы 3.3.2 имеется условие достаточной малости числа $\bar{\alpha} = \sup_{(t)} \alpha_t$. Оно фигурировало и в теореме 2.6.2. Такое условие носит, как уже отмечалось, неконструктивный характер. Чтобы избежать связанных с этим затруднений, можно использовать прием «возврата», описанный в замечании к теореме 2.6.2.

3.4. z -нормализованный процесс. Рассмотрим далее задачу

$$\min \{ \|\mathbf{x} - \mathbf{z}\| : \mathbf{x} \in M \}. \quad (3.3.18)$$

Свяжем с ней процесс

$$\{\mathbf{x}_{t+1} \in (1 - \alpha_t) \varphi(\mathbf{x}_t, \mathbf{s}_t) + \alpha_t \mathbf{z}\}_{\forall t}; \quad (3.3.19)$$

здесь \mathbf{z} — фиксированная точка, $M = M(\mathbf{s}')$.

Ниже в качестве следствия теоремы 3.3.2 будет получена

Теорема 3.3.3. Пусть выполнены условия 1°, 2°, 5° и 7°. Тогда при произвольном $\mathbf{x}_0 \in \mathbb{R}^n$ процесс (3.3.19) сходится к проекции элемента \mathbf{z} на M , т.е. к $\mathbf{z} = \arg$ (3.3.18).

Доказательство. Убедимся, во-первых, в ограниченности процесса (3.3.19). Так как $\mathbf{x}_{t+1} = (1 - \alpha_t) \mathbf{y}_t + \alpha_t \mathbf{z}$, $\mathbf{y}_t \in \Phi(\mathbf{x}_t, \mathbf{s}_t)$, то получаем

$$\begin{aligned} |\mathbf{x}_{t+1} - \bar{\mathbf{z}}| &= |(1 - \alpha_t)(\mathbf{y}_t - \bar{\mathbf{z}}) + \alpha_t(\mathbf{z} - \bar{\mathbf{z}})| \leq \\ &\leq (1 - \alpha_t)(|\mathbf{x}_t - \bar{\mathbf{z}}| + 2\varepsilon_t) + \alpha_t|\mathbf{z} - \bar{\mathbf{z}}| \leq \\ &\leq \begin{cases} |\mathbf{z} - \bar{\mathbf{z}}| + 2\varepsilon_t, & \text{если } |\mathbf{x}_t - \bar{\mathbf{z}}| \leq |\mathbf{z} - \bar{\mathbf{z}}|, \\ |\mathbf{x}_t - \bar{\mathbf{z}}| + 2\varepsilon_t, & \text{если } |\mathbf{x}_t - \bar{\mathbf{z}}| > |\mathbf{z} - \bar{\mathbf{z}}|. \end{cases} \end{aligned}$$

Отсюда

$$|\mathbf{x}_{t+1} - \bar{\mathbf{z}}| \leq \max\{|\mathbf{x}_0 - \bar{\mathbf{z}}|, |\mathbf{z} - \bar{\mathbf{z}}|\} + \sum_{i=0}^t \varepsilon_i,$$

что в силу условия 5° и влечет ограниченность последовательности $\{\mathbf{x}_t\}$.

Условие 2° замкнутости $\Phi(\mathbf{x}, \mathbf{s})$ по $[\mathbf{x}, \mathbf{s}]$ в полном соответствии с доказательством леммы 3.3.5 приводит к $\{\mathbf{x}_t\}' \subset M$.

Перейдем к доказательству равенства $\{\mathbf{x}_t\}' = \{\bar{\mathbf{z}}\}$. С этой целью образуем отображение

$$\bar{\Phi}(\mathbf{x}, \sigma) = (1 - \alpha)\Phi(\mathbf{x}, \mathbf{s}) + \alpha\mathbf{z},$$

здесь $\sigma = [\mathbf{s}, \alpha]$, $0 \leq \alpha \leq 1$. Так как

$$\bar{\Phi}(\mathbf{x}, \sigma) - \alpha \nabla \left(\frac{1}{2} |\mathbf{x} - \mathbf{z}|^2 \right) = (1 - \alpha)\Phi(\mathbf{x}, \mathbf{s}) + \alpha\mathbf{z},$$

то процесс (3.3.19) может быть переписан в виде

$$\left\{ \mathbf{x}_{t+1} \in \bar{\Phi}(\mathbf{x}_t, \sigma_t) - \alpha_t \nabla \left(\frac{1}{2} |\mathbf{x} - \mathbf{z}|^2 \right) \Big|_{\mathbf{x} = \mathbf{x}_t} \right\}_{\forall t}; \quad (3.3.20)$$

здесь $\sigma_t = [\mathbf{s}_t, \alpha_t]$. Для $\bar{\Phi}(\mathbf{x}, \sigma_t)$ множеством неподвижности является M_t , для $\bar{\Phi}(\mathbf{x}, \sigma')$ — множество M (здесь $\sigma' = [\mathbf{s}', 0]$). Следовательно, к процессу (3.3.20), являющемуся перезаписью процесса (3.3.19), применимо замечание 1 к теореме 3.3.2, а потому

$$\{\mathbf{x}_t\}' \subset \text{Arg min} \left\{ \frac{1}{2} |\mathbf{x} - \mathbf{z}|^2 : \mathbf{x} \in M \right\} = \{\bar{\mathbf{z}}\},$$

что и требовалось.

§ 4. Применения к нестационарным задачам выпуклого программирования: общие замечания

В силу результатов §§ 1—2 данной главы построение нестационарных процессов (3.3.1) для модели (3.3.2) не представляет труда. Это связано с широкими возможностями конструирования итерационных отображений $\varphi(\cdot, \mathbf{s}_t) \in F_{M_t}$, поставленных в соответствие системе неравенств

$$f_j(\mathbf{x}, \mathbf{s}_t) \leq 0 \quad (\forall j \in N_m), \quad (3.4.1)$$

т. е. системе неравенств из задачи (3.3.2). Нуждается в обсуждении вопрос об условиях, гарантирующих сходимость процесса. Это мы сделаем применительно к модели

$$\max \{f_0(\mathbf{x}) : \mathbf{x} \in M_t\} \quad \forall t, \quad (3.4.2)$$

в которой M_t задано либо системой (3.4.1), либо в форме

$$M_t = \{\mathbf{x} : f_j(\mathbf{x}) \leq 0, \forall j \in N_t\}. \quad (3.4.3)$$

Первой из этих ситуаций посвящены теоремы из § 2 гл. III, ситуации (3.4.3)—теорема 3.2.6. Помимо уже оговоренных в § 1 условий на $\{f_0(\mathbf{x}), f_j(\mathbf{x}, \mathbf{s}), \forall j \in N_m\}$, будем еще предполагать функции $\{f_j(\mathbf{x}, \mathbf{s})\}$ непрерывными по $\mathbf{z} = [\mathbf{x}, \mathbf{s}]$ на $\mathbb{R}^n \times V$. Это в силу леммы 3.2.1 и ее следствий обеспечивает свойство замкнутости по $\mathbf{z} = [\mathbf{x}, \mathbf{s}]$ отображений $\varphi(\cdot, \mathbf{s}) \in F_{M(\mathbf{s})}$, базовые конструкции которых рассмотрены в § 1 гл. II. Так обстоит дело с условием 2°. Условия 1°, 3°, 6° и 7° по своему характеру не требуют особого обсуждения. Что касается условия 4°, то оно выполняется, как уже отмечалось, для двух хотя и простых, но важных случаев:

- 1) $M_t \supset M_{t+1}, \forall t; \bigcap M_t = M$ — не пусто и ограничено;
- 2) $M_t \subset M_{t+1}, \forall t; \bigcup_{(t)} M_t = M$ — не пусто и ограничено.

Первый из них соответствует, в частности, ситуации (3.4.3). Второй может быть связан с нестационарностью, возникающей в результате построения последовательности внутренних аппроксимаций допустимой области задачи выпуклого программирования.

Если M_t задано системой (3.4.1), то можно предложить различные варианты условий, диктующих выполнимость свойства 4°, например условие Слейтера для предельной системы $f_j(\mathbf{x}, \mathbf{s}') \leq 0$ ($\forall j \in N_m$) и ограниченность множества ее решений (естественно, в предположении выпуклости $f_j(\mathbf{x}, \mathbf{s})$ по \mathbf{x} и непрерывности по $\mathbf{z} = [\mathbf{x}, \mathbf{s}]$, $\forall j \in N_m$). Об этом уже было сказано в § 1 настоящей главы.

С выполнимостью условия 5° дело обстоит сложнее. В ряде случаев, когда, например, M_t выступает в качестве аппроксимирующего множества, выбор которого осуществляется вычислителем, можно считать, что последний контролирует и ε_t — как меру точности аппроксимации. С этой позиции условие 5° воспринимается естественным образом. Применительно к ситуации (3.4.3), связанной с решением бесконечной системы выпуклых неравенств, дело сводится к выбору подходящей подпоследовательности $\{j_t\} \subset \{t\}$ и заменой отображений (3.2.13) на

$$\psi(\mathbf{x}, \mathbf{s}_t) = \frac{1}{\delta_t} \sum_{i=1}^{j_t} \alpha_i \varphi_i(\mathbf{x}),$$

где $\mathbf{s}_t = [\alpha_1, \dots, \alpha_{j_t}]$, $\delta_t = \sum_{i=1}^{j_t} \alpha_i$.

Можно привести и другие ситуации, когда может быть обеспечена выполнимость условия 5°.

ГЛАВА IV

МАТЕМАТИЧЕСКИЕ МЕТОДЫ РАСПОЗНАВАНИЯ ОБРАЗОВ

Задача распознавания образов в общем виде есть задача классификации объектов (любой природы). Иными словами, эта задача состоит в целесообразном разбиении какого-либо заданного множества объектов на классы. При этом в каждый класс входят объекты, в определенном смысле «близкие» друг к другу с точки зрения фиксированной цели, а объекты разных классов существенно различны.

Легко привести примеры задач распознавания образов. Кроме задач узнавания (опознавания) зрительных и звуковых объектов, сюда относятся проблемы диагностики, планирования и управления технико-экономическими, биологическими и другими системами, распознавания характера ситуаций в задачах принятия решений, установления типологий и т. п.

Распознавание образов — одна из основных логических операций мышления. Поэтому математические методы решения задач классификации играют большую роль не только в системах планирования и управления, но и в решении целого ряда других проблем практически во всех отраслях науки и техники.

При построении и анализе математических моделей объектов, а также при управлении объектами методы распознавания образов позволяют:

приспосабливать теоретические модели к реальным условиям их применения;

работать с плохо структурированными системами, формализовать расплывчатые, «размытые» требования; проводить диагностику задач, выбирать метод решения;

агрегировать и обобщать информацию и т. д.

Распознавание образов является средством формализации эвристического поиска, самообучения, имитации,

упорядочения и систематизации материала исследований. Распознавание образов позволяет включить конкретные сведения в общую структуру информации. С другой стороны, с помощью методов распознавания образов можно выделять, фиксировать объекты исследования.

Классификация является также средством уточнения понятий. Определение закономерностей образования классов делает возможным установление других закономерностей.

Процессы моделирования вообще в той или иной форме содержат распознавание образов, так как создание модели уже предполагает наложение на объект некоторой макроструктуры. Реальный объект представляет собой нечто сложное и на первый взгляд хаотическое. Моделирование же по необходимости требует выявления порядка в этом кажущемся хаосе, какого-то агрегирования.

Методы распознавания образов отличаются универсальностью их применения. Однако некие «законы сохранения» требуют, чтобы достоинства уравнивались теми или иными недостатками. Так, например, большинство математических методов распознавания образов, доставляющих достаточно сложную классификацию, связаны с необходимостью оперирования большими массивами информации и соответственно с использованием значительного объема памяти ЭВМ.

Распознавание образов, точнее, математический аспект этой проблемы, является сравнительно новым направлением в прикладной математике. Своеобразная структура этой проблемы, накладываясь на ранее разработанный в близких областях математический аппарат (такой, например, как математическое программирование), приводит к некоторым новым конструкциям и моделям.

Выбор нами аппарата теории систем неравенств и математического программирования в качестве базового для исследования проблем распознавания образов основан на следующих соображениях. С одной стороны, задача математического программирования может быть интерпретирована как некоторая абстрактная задача распознавания образов в ее оптимизационной постановке, и наоборот, многие задачи распознавания образов могут быть хорошо промоделированы задачами математического

программирования. С другой стороны, методы математического программирования оказались эффективным средством решения прикладных задач исследования операций, в том числе настолько сложных, что для них налицо все признаки, обуславливающие применимость тех или иных процедур распознавания образов (плохо формализуемые факторы, неточная или недоопределенная информация и др.).

§ 1. Математические модели распознавания образов

1.1. Понятие образа. Введем некоторые определения.

Образ— множество всех объектов, сходных друг с другом в каком-либо фиксированном отношении.

Например, образами являются следующие множества: совокупность всех крупных машиностроительных предприятий с мелкосерийным производством;

множество всех людей, страдающих определенным заболеванием.

Распознать объект или *распознать образ объекта*— значит указать, к какому образу (классу похожих объектов) он относится.

Приведем следующий пример. Пусть задано некоторое множество $X \subset \mathbf{R}^n$. Элементы $x \in X$ представляют собой модели некоторых объектов, которые функционируют как элементы определенной реальной системы. Требуется найти *функционал* $x^* \in X^* \subset \{X \rightarrow \mathbf{R}\}$, *классифицирующий* элементы множества X с целью установления их специализации, оптимальной для функционирования всей системы.

1.2. Решение задачи распознавания образов человеком и ЭВМ. Человек хорошо распознает зрительные образы: предметы и их взаиморасположения, изображения. То же относится и к звуковым образам (шумы, речь, музыка и т. д.). Более абстрактны задачи распознавания ситуаций: жизненных, производственных и др. Встречаясь с новой ситуацией, человек сопоставляет ее (возможно, неосознанно) с аналогичными ситуациями, а также принятыми в них решениями.

Что касается распознавания больших массивов числового материала (показания приборов, таблицы) и приня-

тия соответствующего оперативного решения, то эта задача представляет для человека либо значительные трудности, либо просто непреодолимые.

В связи с процессом автоматизации производства возникла задача построения автоматов, способных распознавать зрительные, звуковые и числовые образы. В частности, решается задача обучения ЭВМ распознаванию образов. На основе разработанных в настоящее время алгоритмов ЭВМ узнает зрительные и звуковые образы, в общем, хуже, чем человек, но почти всегда лучше анализирует плохо структуризованные массивы чисел.

1.3. Актуальность применения методов распознавания образов в планировании и управлении. Развитие науки сопровождается возрастанием сложности изучаемых систем. Одна из причин этого: постоянное усложнение техники и организации производства и связанная с этим необходимость все более глубокого анализа сущности природных, технических, экономических и социальных явлений, при котором учитываются и самые слабые взаимные влияния элементов систем. Поэтому при прогнозировании ситуаций и характеристик объектов с помощью математических моделей приходится учитывать большое число взаимосвязанных факторов в условиях, когда доступная исследователю информация может иметь самые разнообразные формы и быть различной степени детерминированности.

Большие трудности связаны с моделированием нестационарных и плохо определенных процессов, расплывчатых целей, нечетких ограничений. В качестве примеров систем, при моделировании которых возникают принципиально новые трудности, можно назвать:

- большие системы экономики;
- сложные производственные и технические системы;
- биологические системы;
- социологические системы.

Так, например, хотя абстрактные математические модели реальных [биологических систем и могут быть построены, но их наполнение конкретным содержанием (измерение числовых параметров) связано с необходимостью применения достаточно тонкой техники, в частности требуется применение самообучающихся программ. Эффективным средством учета плохо формализуемых соот-

ношений при моделировании сложных систем является распознавание образов.

1.4. Сфера приложений распознавания образов. Задачи распознавания образов как задачи автоматизации получения и обработки данных, анализа закономерностей в массивах информации, могут быть разбиты на три группы:

1) выделение, формирование и оценка существенных признаков (факторов), преобразование пространства признаков;

2) анализ закономерностей размещения элементов конечного множества в многомерном фазовом пространстве;

3) разбиение пространства на классы эквивалентности, отнесение элементов пространства к классам.

Область приложений методов решения указанных задач может быть описана в общем виде по следующей схеме.

I. Оценка ситуаций: интерпретация обстановки, ориентация, диагностика, идентификация, узнавание, обнаружение, индикация, сравнение.

II. Структуризация информации, организация материала наблюдений.

II.1. Классификация, разбиение, агрегирование, сортировка исходных данных, сжатие, упорядочение практического опыта.

II.2. Типология, унификация, стандартизация, построение параметрических рядов, типопредставителей, эталонов; группировка.

II.3. Запоминание информации, заполнение пробелов в ней, обобщение опыта.

III. Моделирование.

III.1. Решение обратных задач, идентификация моделей.

III.2. Построение разделяющих поверхностей, поверхностей уровня, областей безразличия, проведение границ (например, между средой и системой), построение эмпирических зависимостей, целей и предпочтений, моделей «вход → выход», экстраполяция решения частных задач на более широкий класс задач и перенос на практические ситуации.

III.3. Решение задач поиска и выбора: прогнозирование, управление, планирование, проектирование, кон-

струирование, оценка параметров систем, определение области эффективности применения систем (технических, организационных и др.).

1.5. Модель объекта. *Моделью объекта* служит упорядоченный набор n вещественных чисел — параметров, характеризующих объект, т. е. вектор

$$\mathbf{x} = [x_1, x_2, \dots, x_n].$$

Пример. Объект — предприятие, x_1 — годовая мощность предприятия, x_2 — стоимость основных фондов, x_3 — себестоимость продукции, x_4 — количество рабочих и служащих на предприятии, x_5 — количество внешних поставщиков сырья и полуфабрикатов. Моделью предприятия в данном случае является 5-мерный вектор $\mathbf{x} = [x_1, x_2, x_3, x_4, x_5]$.

1.6. Модель дискриминантного анализа. Предположим, что заданы два конечных множества n -мерных векторов: множество X_1 векторов, изображающих объекты одного рода (первого образа), множество X_2 векторов, изображающих объекты другого рода (второго образа).

Множества X_1 и X_2 вместе составляют *материал обучения*. Это означает, что, проанализировав вид векторов множеств X_1 и X_2 , мы должны выработать правило, согласно которому каждый новый объект, представляемый n -мерным вектором, будет отнесен к первому, либо ко второму образу. Правило можно построить так. Пусть $f(\mathbf{x}) = f(x_1, \dots, x_n)$ — функция от n переменных, для которой: если \mathbf{x} — вектор из множества X_1 , то $f(\mathbf{x}) > 0$; если же \mathbf{x} — элемент множества X_2 , то $f(\mathbf{x}) < 0$.

Тогда *правило распознавания* состоит в следующем: произвольный вектор \mathbf{x} пространства \mathbf{R}^n относится к первому образу, если $f(\mathbf{x}) > 0$, и ко второму, если $f(\mathbf{x}) < 0$. Если же $f(\mathbf{x}) = 0$, то элемент \mathbf{x} можно по произволу отнести либо к первому, либо ко второму образу.

Итак, задача дискриминантного анализа имеет следующий вид: требуется найти функцию f , удовлетворяющую системе неравенств

$$f(\mathbf{x}) > 0 \quad (\forall \mathbf{x} \in X_1), \quad f(\mathbf{x}) < 0 \quad (\forall \mathbf{x} \in X_2), \quad f \in F. \quad (4.1.1)$$

Здесь множество F — выделенный класс функций f . Функция f , решающая систему (4.1.1), называется *разделяющей* или *дискриминантной*.

Идея задачи дискриминантного анализа состоит в продолжении функций, причем на конечных множествах значения этих функций или их знаки должны совпадать с некоторыми заданными. В более общем виде можно говорить о построении модели, реакция которой на некоторое конечное множество векторов переменных должна совпадать с заданной. Очевидна связь этой задачи с задачей идентификации.

Приведем некоторые конкретные постановки.

1. Заданы множества $X_1 \subset X, \dots, X_k \subset X$. Требуется найти

$$\{x_1^*, \dots, x_{k-1}^*\} \subset X^* \subset \{X \rightarrow \mathbf{R}\}$$

такие, что

$$\begin{aligned} \forall j \in \mathbf{N}_{k-1}: x_j^*(\mathbf{x}) > 0 & \quad (\forall \mathbf{x} \in X_j), \\ x_j^*(\mathbf{x}) < 0 & \quad (\forall \mathbf{x} \in \bigcup_{i>j} X_i). \end{aligned}$$

2. Заданы множества $X_1^*, \dots, X_k^* \subset X^* \subset \{X \rightarrow \mathbf{R}\}$. Требуется найти $\{\mathbf{x}_1, \dots, \mathbf{x}_{k-1}\} \subset X$ такие, что

$$\begin{aligned} \forall j \in \mathbf{N}_{k-1}: x^*(\mathbf{x}_j) > 0 & \quad (\forall x^* \in X_j^*), \\ x^*(\mathbf{x}_j) < 0 & \quad (\forall x^* \in \bigcup_{i>j} X_i^*). \end{aligned}$$

1.7. Модель таксономии. Дано конечное множество X n -мерных векторов. Требуется разбить множество X на непересекающиеся подмножества (*таксоны*), причем объекты, входящие в один таксон, должны быть достаточно близки друг к другу с точки зрения выбранного *критерия близости*, а элементы из разных таксонов должны быть достаточно далеки друг от друга.

Наиболее употребляемый критерий близости друг к другу точек

$$\mathbf{x} = [x_1, \dots, x_n], \quad \mathbf{y} = [y_1, \dots, y_n]$$

есть евклидово расстояние:

$$\rho(\mathbf{x}, \mathbf{y}) = \left[\sum_{i=1}^n (x_i - y_i)^2 \right]^{1/2}.$$

В качестве примеров задачи таксономии можно привести задачи типологии объектов.

Задача таксономии может быть сформулирована и в таком виде, что постановки задачи дискриминантного анализа будут являться ее частными случаями:

Требуется найти разбиение заданного множества X на подмножества X_1, \dots, X_k , удовлетворяющее тем или иным условиям, в частности:

$$1) \forall i \in \overline{1, k}, \exists x_i^* \in X^*: X_i = \{x \in X: x_i^*(x) > 0\};$$

$$2) X_i \cap X_j = \emptyset \quad (i \neq j), \quad X = \bigcup_{i=1}^k X_i;$$

3) $\rho(X_i, X_j) \geq \alpha \quad (i \neq j); \quad \forall i, \forall x, y \in X_i: \rho(x, y) \leq \beta$, где ρ — тем или иным способом введенная мера близости, α и β — заданные числа;

4) $\forall i: X_i \supset X'_i$, где X'_i — заданные множества;

5) $\forall i: |X_i| \geq p|X|$, p — заданное число, $p \in (0, 1)$;

6) x_i^* — решение i -й максимальной совместной подсистемы (кратко: μ -подсистемы) системы

$$x^*(x) > 0 \quad (\forall x \in X), \quad x^* \in X^*.$$

1.8. Задача оценивания информативности подсистем признаков. Как уже было сказано, математической моделью объекта является n -мерный вектор $x = [x_1, \dots, x_n]$. Параметры, значениями которых являются x_1, \dots, x_n , называются *признаками объекта*. Для характеристики объектов можно выбирать различные системы признаков, и задача оценки информативности как раз состоит в нахождении *оптимальной* (с точки зрения того или иного критерия оптимальности) *системы признаков*.

Например, может быть поставлена задача: найти систему признаков, измерение которых связано с наименьшими возможными затратами и при использовании которых возможно разделение заданных конечных множеств объектов с помощью функций заданного класса. Задача с другим критерием оптимальности: выбрать подсистему, включающую m признаков из общего их числа n ($n > m$), обеспечивающую наилучшее качество разделения.

Более общая постановка: найти преобразование $\varphi \in \{\mathbb{R}^n \rightarrow \mathbb{R}^m\}$, которое дает соответствующее качество классификации объектов $x \in \mathbb{R}^n$ при условии, что классифицируются объекты $\varphi(x) \in \mathbb{R}^m$. Один из возможных подходов: агрегирование признаков x_i , т. е. решение задачи таксономии признаков в пространстве объектов.

Приведем примеры конкретных постановок. Пусть

$$\begin{aligned} I &= \{1, \dots, n\}, \quad J \subset I, \quad J = \{i_1, \dots, i_k\}, \\ \mathbf{x} &= \mathbf{x}(I) = [x_1, \dots, x_n], \quad \mathbf{x}(J) = [x_{i_1}, \dots, x_{i_k}], \\ X(J) &= \{\mathbf{x}(J) : \mathbf{x} \in X\}; \quad X_1 \subset \mathbb{R}^n, \quad X_2 \subset \mathbb{R}^n, \\ \omega(J) &= \inf \{ \|\mathbf{x} - \mathbf{y}\| : \mathbf{x} \in X_1(J), \quad \mathbf{y} \in X_2(J) \}. \end{aligned}$$

Тогда решается одна из задач:

$$\max \{ \omega(J) : |J| \leq p \} \quad \text{или} \quad \min \{ |J| : \omega(J) \geq \alpha \},$$

где p и α — заранее фиксированные числа.

Эти задачи могут решаться и при других видах функционала качества $\omega(J)$.

1.9. Задача разбиения и коррекции материала наблюдений. Предположим, что задано некоторое конечное множество $X \subset \mathbb{R}^n$, представляющее собой материал наблюдений. Это множество необходимо разбить на три подмножества: $X = Y \cup Z \cup V$, где Y и Z состоят из всех тех элементов множества X , для которых принадлежность их к образам известна; V — материал распознавания — состоит из элементов множества X с неизвестной принадлежностью к образам; Y — материал обучения — состоит из тех элементов, которые мы выделяем для идентификации решающего правила (нахождения разделяющего функционала); Z — материал контроля — подмножество а priori классифицированных объектов для проверки качества идентификации решающего правила. *Задача разбиения материала наблюдений* состоит в нахождении разбиения множества $X \setminus V$ на множества Y и Z , оптимального по тому или иному критерию, оценивающему качество найденного решающего правила.

Задача коррекции материала наблюдений в общих чертах состоит в следующем: по результатам решения задачи дискриминантного анализа необходимо указать элементы множества $X_1 \cup X_2$, которые могли быть неправильно классифицированы «учителем».

Приведем одну из конкретизаций этой постановки. Пусть а priori известно, что множества X_1 и X_2 разделимы классом X^* , но в результате ошибок априорной классификации (ошибок «учителя») они оказались неразделимыми, т. е. система

$$x^*(\mathbf{x}) > 0, \quad (\forall \mathbf{x} \in X_1), \quad x^*(\mathbf{x}) < 0 \quad (\forall \mathbf{x} \in X_2), \quad x^* \in X^*$$

несовместна. Требуется переклассифицировать наименьшее число точек множества $X_1 \cup X_2$, так, чтобы система стала совместной.

1.10. Конструкции для систем неравенств и задачи распознавания образов. Отметим некоторые возможные направления исследований в математической теории распознавания образов, связанных с анализом систем неравенств:

условия совместности соответствующих систем ограничений;

обоснованный выбор класса разделяющих функционалов;

правило принятия решений в случае несовместности систем ограничений;

итерационные алгоритмы решения систем неравенств.

Рассмотрим некоторые способы преодоления несовместности системы ограничений:

$$y \in D_j, \quad (\forall j \in J), \quad y \in Y. \quad (4.1.2)$$

Здесь множества D_j , J и Y пока произвольны, так же как и способ их задания. Такого рода системы возникают при формализации многих задач, в том числе задач математического программирования и распознавания образов.

Задача (4.1.2), состоящая в нахождении элемента y , удовлетворяющего всем заданным условиям, может быть неразрешимой (т. е. система (4.1.2) — несовместной). В этом случае имеются по крайней мере следующие возможности обобщения понятия решения (реализуемые в конкретных задачах различными способами):

1) расширение множеств D_j (на этом пути можно, например, прийти к чебышевским приближениям);

2) расширение класса Y ; это расширение может быть проведено с помощью каких-либо стандартных операторов, применимых к классам Y различного вида; в качестве примера можно привести взятие выпуклой оболочки множества Y (если Y — множество в линейном пространстве);

3) введение «коллективного» решения $K \subset Y$ вместо обычного $y \in Y$ или рандомизация решения (примеры: смешанные стратегии в матричных играх, комитеты разделяющих функционалов в распознавании образов);

4) сужение множества J ; например, рассмотрение максимальных совместных подсистем системы (4.1.2).

Решения максимальных совместных подсистем (μ -подсистем) можно использовать, в частности, в качестве членов коллективного решения. Так, например, пусть x_j^* — решение j -й μ -подсистемы обобщенной задачи распознавания образов, которая определяется соответствующим максимальным подмножеством X_j множества X . Пусть $x \in \mathbb{R}^n$. Для оценки x находим ближайшее к x множество X_j и полагаем в качестве классифицирующей оценки для x число $x_j^*(x)$.

Если разделяющий множества X_1 и X_2 функционал отыскивается в классе аффинных функционалов

$$y(x) - y_{n+1} = (y, x) - y_{n+1},$$

то он должен удовлетворять системе линейных неравенств

$$\left. \begin{aligned} y(x) &> y_{n+1} \quad (\forall x \in X_1), \\ y(x) &< y_{n+1} \quad (\forall x \in X_2). \end{aligned} \right\} \quad (4.1.3)$$

В случае совместности системы (4.1.3) для ее решения применимы методы: свертывания (С. Н. Черникова), линейного программирования, методы линейной коррекции, методы фейеровских отображений и др. Если же система (4.1.3) несовместна, разделение множеств X_1 и X_2 может быть осуществлено с помощью нескольких линейных функций; это и делается, например, в методах комитетов, кусочно-линейных функций, секущих плоскостей и др.

§ 2. Методы аффинного дискриминантного анализа

Рассмотрим задачу дискриминантного анализа для двух множеств: найти элемент $x^* \in X^* \subset \{\mathbb{R}^n \rightarrow \mathbb{R}\}$ такой, чтобы выполнялись неравенства при $X_1 \subset \mathbb{R}^n$, $X_2 \subset \mathbb{R}^n$:

$$x^*(x) > 0 \quad (\forall x \in X_1), \quad x^*(x) < 0 \quad (\forall x \in X_2). \quad (4.2.1)$$

Если X^* — класс аффинных функционалов, т. е. $x^*(x) = (y^*, x) + z^*$, где $y^* \in \mathbb{R}^n$, $z^* \in \mathbb{R}$, то задача нахождения хотя бы одного решения системы линейных неравенств (4.2.1) называется задачей *аффинного дискриминантного анализа*.

Класс аффинных функционалов — наиболее важный класс дискриминантных функционалов.

Использование аффинных функционалов в качестве разделяющих оправдывается их простой структурой и технической реализуемостью, а также тем, что линейный случай дает большие возможности для математического анализа и строгого обоснования вычислительных процедур. Кроме того, многие нелинейные методы дискриминантного анализа в качестве основного блока содержат модуль разделения конечных множеств с помощью аффинного функционала (например, метод комитетов, метод потенциалов и др.).

Рассматриваемые ниже методы относятся к числу *непараметрических*, т. е. считается неизвестным какое бы то ни было статистическое описание образов. Поэтому за *критерий качества классификации*, осуществляемой с помощью найденной разделяющей функции, принимается число безошибочных классификаций элементов множеств X_1 и X_2 .

Решение задачи аффинного дискриминантного анализа в принципе может быть осуществлено с помощью любого метода решения системы линейных неравенств (симплекс-метода, методов сверток С. Н. Черникова, фейеровских отображений, линейной коррекции и т. д.). Однако задачи распознавания образов обладают определенной спецификой, которая влечет свои требования к организации вычислительных процедур. Эта специфика состоит в том, что разделяемые множества X_1 и X_2 , элементы которых формируют решаемую систему линейных неравенств, могут либо быть конечными, но состоящими из очень большого числа элементов, либо представлять собой бесконечные последовательности, элементы которых появляются на входе распознающей системы в дискретном времени. Поэтому предпочтительны итерационные методы (например, фейеровского типа), адаптирующиеся к условиям пополнения материала обучения в процессе его реализации.

Пополнение может осуществляться путем генерирования последовательности $\{x_i\} \subset \mathbf{R}^n$ в потенциально бесконечном процессе наблюдений и безошибочной диагностики «экспертом» реализаций состояния x . В этом случае может рассматриваться задача моделирования «эксперта». Генерирование последовательности $\{x_i\}$ может быть вызвано, например, следующими обстоятельствами.

1. Вектор $x \in R^n$ — модель состояния объекта; происходит бесконечный процесс наблюдений и диагностики все новых реализаций состояний $x \in R^n$.

2. Последовательность $\{x_t\}$ может быть последовательностью приближений к решению задачи

$$\sup \{f(x) : (x \in M) \& Q(x)\},$$

где $Q(\cdot)$ — плохо формализуемый предикат. Эта последовательность может возникать в итерационном вычислительном процессе решения указанной задачи либо в связи с нестационарностью элементов f и M . В этом случае разделяемые множества таковы: X_1 состоит из точек, удовлетворяющих плохо формализуемому ограничению, а X_2 — из точек, не удовлетворяющих ему. Следовательно, обучение дискриминации — процесс все более точного моделирования таких ограничений.

2.1. Методы линейной коррекции. Методы линейной коррекции имеют простейшую структуру. Рассмотрим задачу аффинного дискриминантного анализа: найти хотя бы одно решение $[x^*, y^*]$ системы

$$(x^*, x) + y^* > 0 \quad (\forall x \in X_1), \quad (x^*, x) + y^* < 0 \quad (\forall x \in X_2). \quad (4.2.2)$$

Введем обозначения:

$$[x^*, y^*] = z^* \in R^{n+1}, \quad [x, 1] = z \in R^{n+1}, \\ Z_i = \{z = [x, 1] : x \in X_i\} \quad (i = 1, 2).$$

Тогда система примет вид

$$(z^*, z) > 0 \quad (\forall z \in Z_1), \quad (z^*, z) < 0 \quad (\forall z \in Z_2). \quad (4.2.3)$$

Из элементов множества $Z_1 \cup Z_2$ составляется последовательность $\{z_t\}$, в которой каждый элемент встречается бесконечное число раз, например циклическая последовательность. Затем строится $\{z_t^*\}$ по формуле

$$z_{t+1}^* = z_t^* + \alpha_t z_t,$$

где α_t — коэффициент коррекции. В зависимости от правила выбора α_t получается тот или иной вид метода коррекции:

а) алгоритм с постоянным коэффициентом коррекции: в нем $\alpha_t = 0$, если $(z_t^*, z_t) > 0$, $z_t \in Z_1$, либо

$$(z_t^*, z_t) < 0, \quad z_t \in Z_2; \quad (4.2.4)$$

в противном случае

$$\alpha_t = \begin{cases} \alpha, & \text{если } z_t \in Z_1, \\ -\alpha, & \text{если } z_t \in Z_2, \end{cases}$$

где $\alpha > 0$;

б) алгоритм абсолютной коррекции: полагаем $\alpha_t = 0$, если выполняется соотношение (4.2.4); в противном случае

$$\alpha_t = \min \left\{ \alpha : \alpha > \frac{|(z_i^*, z_t)|}{(z_t, z_t)}, \alpha \text{ — целое} \right\};$$

в) алгоритм частичной коррекции: Полагаем $\alpha_t = 0$, если выполняется (4.2.4); в противном случае α_t выбирается так, чтобы выполнялось равенство

$$|(z_i^*, z_t) - (z_{i+1}^*, z_t)| = \lambda |(z_i^*, z_t)|,$$

где $\lambda \in (0; 2)$.

Теорема 4.2.1. Если система (4.2.3) совместна, то при использовании алгоритма а) существует натуральное число m , при котором $z_m^* = z_{m+1}^* = \dots$, где z_m^* — решение системы (4.2.3).

2.2. Методы фейеровских отображений. Системе (4.2.2) можно поставить в соответствие систему нестрогих линейных неравенств

$$(z^*, z) \geq 1 \quad (\forall z \in Z_1), \quad (z^*, z) \leq -1 \quad (\forall z \in Z_2). \quad (4.2.5)$$

Если \bar{z}^* — решение системы (4.2.3), то $\lambda \bar{z}^*$ — решение системы (4.2.5) при достаточно большом $\lambda > 0$; с другой стороны, всякое решение системы (4.2.5) удовлетворяет системе (4.2.3). Более того, решениями системы (4.2.3) могут быть приближенные решения системы (4.2.5): если $(z^*, z) \geq 1 - \varepsilon \quad (\forall z \in Z_1), \quad (z^*, z) \leq -1 + \varepsilon \quad (\forall z \in Z_2)$,

то \bar{z}^* — решение системы (4.2.3) при $\varepsilon < 1$. Следовательно, итерационный метод, дающий последовательность, которая сходится к решению системы (4.2.5), обеспечивает решение системы (4.2.3) за конечное число шагов. В качестве итерационного оператора может выступать фейеровское отображение.

2.3. Фейеровские методы в задаче обучения дискриминации бесконечных множеств. Рассмотрим задачу аффинного дискриминантного анализа: построить аффинный

функционал $z(\mathbf{x})$, разделяющий некоторые множества X_1 и X_2 , которые выступают в роли обучающей выборки образов $\tilde{X}_1 \subset \mathbb{R}^n$, $\tilde{X}_2 \subset \mathbb{R}^n$: $X_1 \subset \tilde{X}_1$, $X_2 \subset \tilde{X}_2$. Посредством $z(\mathbf{x})$ любой элемент $\mathbf{x} \in \tilde{X}_1 \cup \tilde{X}_2$ классифицируется по правилу

$$\mathbf{x} \in \begin{cases} \tilde{X}_1, & \text{если } \text{sign } z(\mathbf{x}) = -1, \\ \tilde{X}_2, & \text{если } \text{sign } z(\mathbf{x}) = +1. \end{cases}$$

В основу построения $z(\mathbf{x})$ могут быть положены те или иные итерационные процессы, обладающие свойством адаптации к режиму нестационарности состояний множеств X_1 и X_2 , например к режиму, когда множества X_1 и X_2 (материал обучения) пополняются. Пополнение может осуществляться путем генерирования последовательности элементов $\mathbf{x}_t \in \tilde{X}_1 \cup \tilde{X}_2$, подвергающихся безошибочной классификации «экспертом» σ :

$$\sigma(\mathbf{x}_t) = \begin{cases} +1, & \text{если } \mathbf{x}_t \in \tilde{X}_1, \\ -1, & \text{если } \mathbf{x}_t \in \tilde{X}_2. \end{cases} \quad (4.2.6)$$

Поставим задачу поиска решающего правила (в форме линейного функционала), моделирующего «эксперта» σ . С этой целью построим итерационный процесс, ведущий в пределе к функционалу, безошибочно дискриминирующему бесконечные множества со $(X_1 \cup \{\mathbf{x}_t\}_-)$ и со $(X_2 \cup \{\mathbf{x}_t\}_+)$; здесь

$$\{\mathbf{x}_t\}_- = \{\mathbf{x}_t: \sigma(\mathbf{x}_t) = -1\}, \quad \{\mathbf{x}_t\}_+ = \{\mathbf{x}_t: \sigma(\mathbf{x}_t) = +1\}.$$

Последовательность $\{\mathbf{x}_t\}$ выступает, таким образом, в качестве обучающей.

В основу анализа сходимости процесса обучения дискриминации, использующего информацию «эксперта» (4.2.6), положим теорему 3.2.7.

Рассмотрим совместную систему линейных неравенств над \mathbb{R}^n :

$$z_j(\mathbf{x}) = (\mathbf{a}_j, \mathbf{x}) - b_j \leq 0 \quad (\forall j \in \mathbb{N}), \quad \mathbf{x} \in C, \quad (4.2.7)$$

где C — непустое выпуклое множество. Будем предполагать, что $|\mathbf{a}_j| \neq 0$ ($\forall j$). Пусть $\{\alpha_j\}$ — числовая последовательность, для которой $\sum_{j=1}^{+\infty} \alpha_j = 1$, $\alpha_j > 0$ ($\forall j \in \mathbb{N}$).

Положим

$$M_t = \{x: z_j(x) \leq 0 \quad (\forall j \in N_t)\} \cap C, \quad M = \bigcap_{(t)} M_t;$$

$$\varphi_t(x) = \pi_C \left(x - \lambda \sum_{j=1}^t \alpha_j \frac{z_j^+(x)}{|a_j|^2} a_j \right), \quad (4.2.8)$$

где $\lambda \in (0, 2)$, $\pi_C(\cdot)$ — оператор метрического проектирования на C ;

$$\varphi(x) = \pi_C \left(x - \lambda \sum_{j=1}^{+\infty} \alpha_j \frac{z_j^+(x)}{|a_j|^2} a_j \right). \quad (4.2.9)$$

Запись (4.2.9) носит формальный характер. Однако если $M \neq \emptyset$, то, как легко видеть, функция

$$d(x) = \sup_{(j)} \frac{z_j^+(x)}{|a_j|}$$

конечна на всем \mathbf{R}^n , откуда усматривается сходимость ряда

$$\sum_{j=1}^{+\infty} \alpha_j z_j^+(x) |a_j|^{-2} a_j,$$

а потому и существование определенного значения для $\varphi(x)$.

В силу теоремы 3.2.7 справедливо следующее утверждение.

Теорема 4.2.2. *Если система (4.2.7) совместна и оператор φ_t определен формулой (4.2.8), то процесс*

$$\{x_{t+1} = \varphi_t(x_t)\}_{\forall t}$$

сходится к элементу из M .

Перейдем к задаче обучения дискриминации. Будем предполагать строгую линейную разделимость образов \tilde{X}_1 и \tilde{X}_2 . Можно считать, что множества X_1 и X_2 поглощаются множествами $\{x_t\}_-$ и $\{x_t\}_+$ соответственно. Требуется найти вектор $z = [c, \beta] \in \mathbf{R}^{n+1}$ такой, чтобы выполнялись условия

$$\begin{aligned} (x, c) + \beta &> 0 \quad (\forall x \in \{x_t\}_-), \\ (x, c) + \beta &< 0 \quad (\forall x \in \{x_t\}_+). \end{aligned} \quad (4.2.10)$$

Эта задача эквивалентна отысканию $z = [c, \beta]$ из системы

$$\begin{aligned} (x, c) + \beta - 1 &\geq 0 \quad (\forall x \in \{x_t\}_-) \\ (x, c) + \beta + 1 &\leq 0 \quad (\forall x \in \{x_t\}_+) \end{aligned}$$

или из системы

$$L_j(z) = \sigma_j [(x_j, c) + \beta + \sigma_j] \leq 0 \quad (\forall j \in N);$$

здесь $\sigma_j = \sigma(x_j)$.

В соответствии с (4.2.8) при $C = \mathbf{R}^{n+1}$ (тогда $\pi_C(z) = z$) образуем оператор

$$\Phi_t(z) = z - \lambda \sum_{j=1}^t \sigma_j \alpha_j \frac{L_j^+(z)}{|x_j|^2 + 1} [x_j, +1]$$

при прежних условиях на λ и $\{\alpha_j\}$.

Теорема 4.2.3. *В предположении строгой линейной делимости образов \tilde{X}_1 и \tilde{X}_2 последовательность $\{z_t\}$, порожденная соотношением*

$$z_{t+1} = \Phi_t(z_t) \quad (4.2.11)$$

при произвольном $z_0 \in \mathbf{R}^{n+1}$, сходится к некоторому вектору $\tilde{z} = [\tilde{c}, \tilde{\beta}]$, удовлетворяющему системе (4.2.10). Следовательно, функция $(\tilde{c}, x) + \tilde{\beta}$ безошибочно дискриминирует элементы множеств со $\{x_t\}_-$ и со $\{x_t\}_+$, где черта сверху означает топологическое замыкание.

Это утверждение непосредственно следует из теоремы 4.2.2.

Замечание 1. На искомый функционал $(c, x) + \beta$ могут быть наложены дополнительные требования в форме $z = [c, \beta] \in C$. Если C — выпуклое замкнутое множество, соотношение (4.2.11) следует заменить на

$$z_{t+1} = \pi_C[\Phi_t(z_t)].$$

Возникающая при этом последовательность $\{z_t\}$ решает задачу нахождения классифицирующего функционала с дополнительным условием.

Замечание 2. Если \tilde{X}_1 и \tilde{X}_2 линейно делимы, но не обязательно строго, то поиск $z = [c, \beta]$ из условий (4.2.10), а потому и из (4.2.11) невозможен. В этом случае \tilde{z} можно искать, решая систему

$$l_j(z) = \sigma_j [(x_j, c) + \beta] \leq 0 \quad (\forall j \in N) \quad (4.2.12)$$

путем организации процесса

$$z_{t+1} = \Phi_t(z_t)$$

при

$$\varphi_t(z) = z - \lambda \sum_{j=1}^t \sigma_j \alpha_j \frac{l_j^+(z)}{|x_j|^2 + 1} [x_j, 1]. \quad (4.2.13)$$

Здесь хотя и гарантируется, что $z_t \rightarrow \tilde{z}$ ($t \rightarrow +\infty$), но нельзя утверждать, что $\tilde{z} \neq 0$. Для отыскания ненулевого разделяющего функционала необходимо наложить то или иное условие нормализации на выбор z . Если $c = [c_1, \dots, c_n]$, то заведомо ведет к цели одна из нормализаций: $\sum_{i=1}^n c_i + \beta = \pm 1$, однако заранее нельзя сказать, какая именно.

Эту трудность можно преодолеть следующим образом. Положим

$$C_1 = \left\{ z = [c, \beta]: \sum_{i=1}^n c_i + \beta = 1 \right\},$$

$$C_2 = \left\{ z = [c, \beta]: \sum_{i=1}^n c_i + \beta = -1 \right\}.$$

Образум две последовательности:

$$z'_{t+1} = \pi_{C_1} [\varphi_t(z'_t)], \quad z''_{t+1} = \pi_{C_2} [\varphi_t(z''_t)],$$

где φ_t — из (4.2.13). По крайней мере одна из этих последовательностей сходится к решению системы (4.2.12). Какая именно, можно установить, отслеживая невязки

$$\varepsilon'_t = \max_{j \in N_t} l_j^+(z'_t) \quad \text{и} \quad \varepsilon''_t = \max_{j \in N_t} l_j^+(z''_t):$$

если, например, к решению системы (4.2.12) сходится $\{z'_t\}$, то $\varepsilon'_t \rightarrow 0$ ($t \rightarrow +\infty$).

Предлагаемый процесс может быть применен при решении практических задач классификации и диагностики объектов различной природы (техничко-экономических, природных и других); процесс позволяет проводить диагностику в условиях малого информационного обеспечения и постепенно улучшать качество диагностики при увеличении информации.

Кроме того, он может быть интерпретирован как процесс решения задачи планирования с потенциально бесконечным числом ограничений, причем первоначальное

конечное число учитываемых ограничений растет с течением времени. Наконец, рассматриваемый процесс может быть истолкован как все более точная идентификация модели, имеющей вид аффинного функционала.

2.4. Методы фейеровского типа в случае линейной неразделимости множеств. Уточним классификацию разделимости множеств.

Определение 4.2.1. Множества $X_1 \subset \mathbf{R}^n$, $X_2 \subset \mathbf{R}^n$ назовем *аффинно разделимыми*, если система

$$(\mathbf{x}^*, \mathbf{x}) + y^* \leq 0 \quad (\forall \mathbf{x} \in X_1), \quad (\mathbf{x}^*, \mathbf{x}) + y^* \geq 0 \quad (\forall \mathbf{x} \in X_2) \quad (4.2.14)$$

имеет ненулевое решение.

Определение 4.2.2. Множества X_1 , X_2 назовем *линейно разделимыми*; если система

$$(\mathbf{x}^*, \mathbf{x}) \leq 0 \quad (\forall \mathbf{x} \in X_1), \quad (\mathbf{x}^*, \mathbf{x}) \geq 0 \quad (\forall \mathbf{x} \in X_2) \quad (4.2.15)$$

имеет ненулевое решение.

Определение 4.2.3. Множества X_1 , X_2 назовем *устойчиво (строго) аффинно разделимыми*, если совместна система

$$(\mathbf{x}^*, \mathbf{x}) + y^* < 0 \quad (\forall \mathbf{x} \in X_1), \quad (\mathbf{x}^*, \mathbf{x}) + y^* > 0 \quad (\forall \mathbf{x} \in X_2). \quad (4.2.16)$$

Определение 4.2.4. Множества X_1 , X_2 назовем *устойчиво (строго) линейно разделимыми*, если совместна система

$$(\mathbf{x}^*, \mathbf{x}) < 0 \quad (\forall \mathbf{x} \in X_1), \quad (\mathbf{x}^*, \mathbf{x}) > 0 \quad (\forall \mathbf{x} \in X_2). \quad (4.2.17)$$

Определение 4.2.5. Множества X_1 , X_2 назовем *неустойчиво аффинно разделимыми*, если система (4.2.14) имеет ненулевое решение, а система (4.2.16) несовместна.

Определение 4.2.6. Множества X_1 , X_2 назовем *неустойчиво линейно разделимыми*, если система (4.2.15) имеет ненулевое решение, а система (4.2.17) несовместна.

Задачу (4.2.2), где множества X_1 и X_2 конечны, можно свести к одной из следующих трех. Если предположить, что существует решение, в котором $y^* = 0$, то нужно найти решение системы

$$(\mathbf{x}^*, \mathbf{x}) > 0 \quad (\forall \mathbf{x} \in X_1), \quad (\mathbf{x}^*, \mathbf{x}) < 0 \quad (\forall \mathbf{x} \in X_2).$$

Если $y^* > 0$, то неравенства (4.2.2) можно поделить на y^* . Обозначив $x^*/y^* = v^*$, получим

$$(v^*, x) + 1 > 0 \quad (\forall x \in X_1), \quad (v^*, x) + 1 < 0 \quad (\forall x \in X_2).$$

Если же $y^* < 0$, то получаем систему

$$(v^*, x) + 1 < 0 \quad (\forall x \in X_1), \quad (v^*, x) + 1 > 0 \quad (\forall x \in X_2). \quad (4.2.18)$$

Предположим теперь, что множества X_1 и X_2 не разделяются гиперплоскостью. Тогда рассмотренные задачи можно заменить задачей нахождения чебышевского приближения: найти

$$\min \{t \geq 0: - (v^*, x) - 1 \leq t \quad (\forall x \in X_1), \\ (v^*, x) + 1 \leq t \quad (\forall x \in X_2)\}$$

либо

$$E = \min \{t \geq 0: (v^*, x) + 1 \leq t \quad (\forall x \in X_1), \\ - (v^*, x) - 1 \leq t \quad (\forall x \in X_2)\}. \quad (4.2.19)$$

Действительно, пусть, например, система (4.2.18) несовместна. Кроме того, будем предполагать, что множества X_1, X_2 не являются неустойчиво аффинно разделимыми. Пусть $[\bar{v}^*, \bar{t}]$ — какое-либо решение задачи (4.2.19). Тогда $\bar{t} > 0$, так как иначе была бы совместной либо система (4.2.18), либо система

$$\begin{aligned} (v^*, x) + 1 &\leq 0 \quad (\forall x \in X_1), \\ - (v^*, x) - 1 &\leq 0 \quad (\forall x \in X_2). \end{aligned} \quad (4.2.20)$$

Если же система (4.2.18) несовместна, а система (4.2.20) совместна, то множества X_1, X_2 неустойчиво аффинно разделимы, что противоречит предположению. Итак, $\bar{t} > 0$.

Кроме того, $\bar{t} \leq 1$, так как при $t = 1$ допустимое множество задачи (4.2.19) не пусто: в него входит $v^* = [0, \dots, 0]$.

Так как $[\bar{v}^*, \bar{t}]$ — решение задачи (4.2.19), то

$$\begin{aligned} (\bar{v}^*, x) + (1 - \bar{t}) &\leq 0 \quad (\forall x \in X_1), \\ (\bar{v}^*, x) + (1 + \bar{t}) &\geq 0 \quad (\forall x \in X_2). \end{aligned}$$

Таким образом, смысл задачи (4.2.19) таков: гиперплоскость

$$H = \{x: (\bar{v}^*, x) + 1 = 0\}$$

приближенно разделяет X_1 и X_2 , причем неправильно классифицированы могут быть только точки *плоского слоя*

$$\{x: -\bar{t} \leq (\bar{v}^*, x) + 1 \leq \bar{t}\}.$$

Предположим, что наряду с условиями (4.2.19) имеется линейный критерий качества разделяющего функционала: $x(v^*) = (c, v^*)$. Тогда задача дискриминантного анализа примет вид

$$\min \{(c, v^*): v^* \in M^*\},$$

где $M^* = \{v^*: (v^*, x) + 1 \leq E \ (\forall x \in X_1), \quad -(v^*, x) - 1 \leq E \ (\forall x \in X_2)\}$, E — из (4.2.19).

Приведем алгоритм решения этой задачи, не предполагающий предварительного определения числа E . Для всякого $v^* \in R^n$ определим:

$$\begin{aligned} \alpha(v^*) &= \max \{(v^*, x) + 1: x \in X_1\}, \\ \beta(v^*) &= \max \{-(v^*, x) - 1: x \in X_2\}, \\ \gamma(v^*) &= \max \{0, \alpha(v^*), \beta(v^*)\}, \\ a(v^*) &= \arg \max \{(v^*, x) + 1: x \in X_1\}, \\ b(v^*) &= \arg \max \{-(v^*, x) - 1: x \in X_2\}. \end{aligned}$$

Построим последовательность $\{v_i^*\} \subset R^n$:

$$v_0^* \text{ — произвольная точка,}$$

$$v_{i+1}^* = \begin{cases} v_i^* - \lambda_i \gamma(v_i^*) | a(v_i^*) |^{-2} - \varepsilon_i c, & \text{если } \gamma(v_i^*) = \alpha(v_i^*), \\ v_i^* - \lambda_i \gamma(v_i^*) | b(v_i^*) |^{-2} - \varepsilon_i c, & \text{если } \gamma(v_i^*) = \beta(v_i^*), \\ v_i^* - \varepsilon_i c, & \text{если } \gamma(v_i^*) = 0. \end{cases}$$

Здесь $\lambda_t \in (0, 2)$; ε_t удовлетворяет одному из условий:

$$\varepsilon_t \rightarrow 0, \quad \sum_{t=0}^{\infty} \varepsilon_t = +\infty, \quad \lambda_t = \lambda \in (0, 2), \quad (4.2.21)$$

$$\varepsilon_t \rightarrow 0, \quad \sum_{t=0}^{\infty} \varepsilon_t = +\infty, \quad \lambda_t = \varepsilon_t, \quad (4.2.22)$$

$$\varepsilon_t = \varepsilon > 0; \quad \lambda_t = \lambda \in (0, 2). \quad (4.2.23)$$

Теорема 4.2.4. Если $E=0$, то при условиях (4.2.21) и $|\mathbf{x}|=1$ ($\forall \mathbf{x} \in X_1 \cup X_2$) выполняется соотношение

$$|\mathbf{v}_t^* - \text{Arg min } \{(\mathbf{v}^*, \mathbf{c}): \mathbf{v}^* \in M^*\}| \rightarrow 0 \quad (t \rightarrow +\infty).$$

Доказательство этой теоремы ввиду его громоздкости мы не приводим, оно использует технику гл. III (полное доказательство см. в [84]).

Пусть теперь $E > 0$. Рассмотрим процесс, определяемый соотношениями (4.2.22). Положим

$$\begin{aligned} \gamma(\mathbf{v}_t^*) &= \gamma_t, \\ \mathbf{c}_t &= \begin{cases} a(\mathbf{v}_t^*), & \text{если } \gamma_t = \alpha(\mathbf{v}_t^*), \\ -b(\mathbf{v}_t^*), & \text{если } \gamma_t = \beta(\mathbf{v}_t^*), \end{cases} \\ \bar{\mathbf{v}}_{t+1}^* &= \mathbf{v}_t^* - \lambda_t \gamma_t |\mathbf{c}_t|^{-2} \mathbf{c}_t. \end{aligned}$$

В силу замечания к лемме 2.4.4 справедливо следующее утверждение.

Лемма 4.2.1. Пусть

$$M = \bigcap_{j=1}^m M_j \neq \emptyset, \quad M_j = \{\mathbf{x}: (\mathbf{d}_j, \mathbf{x}) - b_j \leq 0\}.$$

Тогда существует константа $D > 0$ такая, что $|\mathbf{x} - M| \leq D |\mathbf{x} - M_{j(\mathbf{x})}|$, где

$$j(\mathbf{x}) = \arg \max \{(\mathbf{d}_j, \mathbf{x}) - b_j: j=1, \dots, m\}.$$

Лемма 4.2.2. При условии $E > 0$ существует $\theta \in (0, 1)$ такое, что для любого $r > 0$ и некоторого $\varepsilon > 0$ справедливо утверждение: если $\lambda_t \leq \varepsilon$ и $|\mathbf{v}_t^* - M| \geq r$, то

$$|\mathbf{v}_{t+1}^* - M^*| \leq |\mathbf{v}_t^* - M^*| - \lambda_t E \theta |\mathbf{c}_t|^{-1}.$$

Доказательство. Мы имеем соотношение

$$|\mathbf{v}_t^* - \bar{\mathbf{v}}_{t+1}^*| = \lambda_t |\gamma_t| |\mathbf{c}_t|^{-1} \geq \lambda_t E |\mathbf{c}_t|^{-1}. \quad (4.2.24)$$

Пусть \mathbf{y}_t^* — проекция \mathbf{v}_t^* на M^* . Покажем, что

$$\exists \mu > 0: (\mathbf{y}_t^* - \mathbf{v}_t^*, \bar{\mathbf{v}}_{t+1}^* - \mathbf{v}_t^*) |\mathbf{y}_t^* - \mathbf{v}_t^*|^{-1} |\mathbf{v}_{t+1}^* - \mathbf{v}_t^*|^{-1} \geq \mu. \quad (4.2.25)$$

Обозначим через H_t гиперплоскость

$$H_t = \left\{ \mathbf{x}: \begin{cases} (\mathbf{c}_t, \mathbf{x}) + 1 = E, & \text{если } \mathbf{c}_t = a(\mathbf{v}_t^*), \\ (\mathbf{c}_t, \mathbf{x}) - 1 = E, & \text{если } \mathbf{c}_t = -b(\mathbf{v}_t^*). \end{cases} \right\}$$

Пусть z_i^* — проекция точки v_i^* на H_t . По лемме 4.2.1

$$|v_i^* - y_i^*| \leq D |v_i^* - z_i^*|, \quad D > 0.$$

Гиперплоскость H_t разделяет точки v_i^* и y_i^* ; пусть $\{v_t\} = H_t \cap \overline{v_i^*, y_i^*}$. Тогда

$$\begin{aligned} |y_i^* - v_i^*| &\geq |v_t - v_i^*| = \\ &= |v_i^* - z_i^*| (|z_i^* - v_i^*|^{-1} (z_i^* - v_i^*), |v_t - v_i^*|^{-1} (v_t - v_i^*))^{-1} = \\ &= |z_i^* - v_i^*| (|z_i^* - v_i^*|^{-1} (z_i^* - v_i^*), |y_i^* - v_i^*|^{-1} (y_i^* - v_i^*))^{-1}. \end{aligned}$$

Отсюда

$$\begin{aligned} (y_i^* - v_i^*, z_i^* - v_i^*) |y_i^* - v_i^*|^{-1} |z_i^* - v_i^*|^{-1} &\geq \\ &\geq |z_i^* - v_i^*| |y_i^* - v_i^*|^{-1}. \end{aligned} \quad (4.2.26)$$

Из соотношения (4.2.26) и определения D следует неравенство

$$(y_i^* - v_i^*, z_i^* - v_i^*) |y_i^* - v_i^*|^{-1} |z_i^* - v_i^*|^{-1} \geq D^{-1}.$$

Так как

$$|z_i^* - v_i^*|^{-1} (z_i^* - v_i^*) = |\bar{v}_{i+1}^* - v_i^*|^{-1} (\bar{v}_{i+1}^* - v_i^*),$$

то неравенство (4.2.25) выполняется при $\mu = D^{-1}$.

Положим $r_t = |v_i^* - M^*|$, $\bar{r}_t = |\bar{v}_{i+1}^* - M^*|$. Используя (4.2.24) и (4.2.25), получаем

$$\bar{r}_t \leq |\bar{v}_{i+1}^* - y_i^*| \leq (r_t^2 + \lambda_t^2 E^2 |c_t|^{-2} - 2r_t \lambda_t E |c_t|^{-1} \mu)^{1/2}.$$

Легко показать, что число

$$[r_t - (r_t^2 + \lambda_t^2 E^2 |c_t|^{-2} - 2r_t \lambda_t E |c_t|^{-1} \mu)^{1/2}] \lambda_t^{-1} |c_t| E^{-1}$$

увеличивается при уменьшении λ_t . Следовательно, при $\lambda_t \geq \varepsilon$

$$\begin{aligned} (r_t - \bar{r}_t) \lambda_t^{-1} E^{-1} |c_t| &\geq [r_t - (r_t^2 + \lambda_t^2 E^2 |c_t|^{-2} - \\ &\quad - 2r_t \lambda_t E |c_t|^{-1} \mu)^{1/2}] \lambda_t^{-1} E^{-1} |c_t| \geq \\ &\geq [r_t - (r_t^2 + \varepsilon^2 E^2 |c_t|^{-2} - 2r_t \varepsilon E |c_t|^{-1} \mu)^{1/2}] \varepsilon^{-1} E^{-1} |c_t| \geq \\ &\geq [r_t - (r_t^2 + \varepsilon^2 E^2 |c'|^{-2} - 2r_t \varepsilon E |c'|^{-1} \mu)^{1/2}] \varepsilon^{-1} E^{-1} |c'|^{-1}, \end{aligned}$$

где $|c'| = \max_t |c_t|$.

Нам нужно подобрать ε так, чтобы при $r_t \geq r$ выполнялось неравенство

$$[r_t - (r_t^2 + \varepsilon^2 E^2 |c'|^2 - 2r_t \varepsilon E \mu |c'|^{-1})^{1/2}] \varepsilon^{-1} E^{-1} |c'| \geq \geq \theta > 0.$$

Очевидно, последнее неравенство выполняется, если

$$\varepsilon \leq 2r (\mu - \theta |c'|) (1 - \theta^2)^{-1} E^{-1}. \quad (4.2.27)$$

Лемма доказана.

Теорема 4.2.5. Если $|c| < E \mu |c'|^{-1}$, то

$$|v_t^* - M^*| \rightarrow 0 \quad (t \rightarrow +\infty).$$

Доказательство. Положим

$$\theta = \mu - \delta, \quad 0 < \delta < \mu - |c| \cdot |c'| E^{-1}.$$

При выполнении условия (4.2.27) справедливо утверждение леммы 4.2.1, т. е. при любом $r > 0$ и при $\lambda_t \leq \varepsilon \leq 2r (\mu - \theta |c'|) (1 - \theta^2)^{-1} E^{-1}$ для всех v_t^* таких, что $|v_t^* - M^*| \geq r$:

$$r_{t+1} \leq r_t - \lambda_t E \theta |c_t|^{-1} + \lambda_t |c| = r_t - \lambda_t (E \theta |c_t|^{-1} - |c|).$$

Так как $E \theta |c'|^{-1} - |c| = E (\mu - \delta) (c') - |c|$, то

$$E \theta |c_t|^{-1} - |c| \geq E \theta |c'|^{-1} - |c| = E (\mu - \delta) |c'|^{-1} - |c|.$$

Ввиду выбора δ

$$E (\mu - \delta) |c'|^{-1} - |c| > E \mu |c'|^{-1} - E \mu |c'|^{-1} + |c| - |c| = 0.$$

Таким образом,

$$r_{t+1} \leq r_t - \lambda_t (E \theta |c_t|^{-1} - |c|),$$

где

$$E \theta |c_t|^{-1} - |c| \geq E \theta |c'|^{-1} - |c| > 0.$$

Так как $\sum_{t=1}^{\infty} \lambda_t = +\infty$, $\lambda_t \rightarrow 0$, то отсюда следует, что $r_t \leq r$, начиная с некоторого достаточно большого t . Устремляя r к нулю, получаем требуемое утверждение.

Теорема 4.2.5 показывает, что процесс (4.2.22) позволяет найти аффинный разделяющий функционал, оптимальный по критерию (4.2.19).

§ 3. Комитетные методы распознавания образов

Понятие комитета несовместной системы однородных строгих линейных неравенств было введено К. Эйблэу и Д. Кэйлором [112].

Комитетом системы

$$(c_j, x) > 0 \quad (\forall j \in N_m),$$

где $x, c_j \in \mathbb{R}^n$ называется такое конечное множество $K \subset \mathbb{R}^n$, что каждому ее неравенству удовлетворяют более половины элементов множества K .

3.1. Определения и обозначения. Рассмотрим систему множеств $F = \{D_j: j \in J\}$, множество Y (произвольной природы) и поставленную им в соответствие систему соотношений

$$y \in D_j \quad (\forall j \in J), \quad y \in Y. \quad (4.3.1)$$

Для $F' \subset F$ положим $J' = \{j: D_j \in F'\}$.

Определение 4.3.1. Подсистему $F' \subset F$ назовем *совместной*, если $\bigcap_{j \in J'} D_j \neq \emptyset$. В противном случае систему F' будем называть *несовместной*.

В соответствии с этим определением можно рассматривать *максимальные* (по включению) *совместные подсистемы* и *минимальные* (по включению) *несовместные подсистемы* системы F .

Определение 4.3.2. *Системой представителей* для системы F , лежащей в Y , будем называть всякое множество K , обладающее свойством

$$K \subset Y, \quad K \cap D_j \neq \emptyset \quad (\forall j \in J).$$

Можно рассматривать минимальные (по включению) системы представителей.

Определение 4.3.3. *Системой различных представителей* для системы F , лежащей в Y , называется множество

$$K = \{y_j: j \in J\} \subset Y: y_j \in D_j \quad (\forall j \in J), \quad y_i \neq y_j \quad (i \neq j).$$

Будем использовать *характеристическую функцию* множества D_j :

$$\varphi_j(y) = \begin{cases} 1, & y \in D_j; \\ -1, & y \notin D_j. \end{cases}$$

Введем обозначения:

$$P_n = \{x = [x_1, \dots, x_n] \in \mathbb{R}^n: x_i \in \{1, -1\} \quad (\forall i \in \mathbb{N}_n)\};$$

f_j^n — функция над P_n :

$$[f_j^n(a_1, \dots, a_n) > f_j^n(b_1, \dots, b_n)] \Leftrightarrow \\ \Leftrightarrow |\{i: a_i = 1\}| > |\{i: b_i = 1\}|.$$

Определение 4.3.4. Коллективным решением для системы F в классе Y (или для системы (4.3.1)) назовем такое множество K , для которого

$$K = \{y_1, \dots, y_q\} \subset Y: f_j^q(\varphi_j(y_1), \dots, \varphi_j(y_q)) > 0 \quad (\forall j \in J).$$

Определение 4.3.5. Комитетной конструкцией в классе Y для системы (4.3.1) (системы F) назовем множество $K \subset Y$ со свойством

$$\psi_j(K \cap D_j) > \psi'_j(K \setminus D_j) \quad (\forall j \in J),$$

где ψ_j, ψ'_j — некоторые функционалы над подмножествами множества Y . Например:

а) $\psi_j(A) = \psi'_j(A) = |A|$ (комитет);

б) $\psi_j(A) = \psi'_j(A) = \int_{x \in A} f_j(x) dx$, f_j — некоторый функ-

ционал;

в) $\psi_j(A) \neq \psi'_j(A)$; $\psi_j(A) = (1 - \alpha_j) |A|$; $\psi'_j(A) = \alpha_j |A|$, $\alpha_j > 0$.

Определение 4.3.6. Обобщенным решением в классе Y для системы F (системы (4.3.1)) называется множество $K \subset Y$ со свойством

$$|K \cap D_j| \geq \alpha_j |K| \quad (\forall j \in J), \quad \alpha_j > 0.$$

Определение 4.3.7. Пусть $z = [z_1, \dots, z_q] \in \mathbb{R}^q$, $p \in [0, 1]$. (z, p) -решением в классе Y для системы F (системы (4.3.1)) называется множество $K \subset Y$ со свойством

$$K = \{y_1, \dots, y_q\}, \\ \sum_{i=1}^q z_i \varphi_j(y_i) > (2p - 1) \sum_{i=1}^q |z_i|.$$

При $p = 0,5$ будем (z, p) -решение называть z -решением; при $z \in \mathbb{R}_+^q$ — неотрицательным (z, p) -решением.

Определение 4.3.8. (z, p) -комитетом системы F (системы (4.3.1)) будем называть (z, p) -решение при це-

лых z_i . *Неотрицательный p -комитет* есть множество K со свойством

$$K \subset Y, |K \cap D_j| > p|K|, p \in [0; 1].$$

Комитетом называется неотрицательный p -комитет при $p = 0,5$. (z, p) -решение и p -комитет называются *слабыми*, если в определении знак $>$ заменен на \geq .

Определение 4.3.9. *Вероятностный p -комитет* есть множество $K \subset Y$ и случайный элемент $y \in K$ с некоторым законом распределения, причем

$$p \{y \in D_j\} > p_j \quad (\forall j \in J).$$

Частные случаи: $p_j = p \ (\forall j)$; $p_j = 0,5 \ (\forall j)$.

Вероятностный p -комитет может быть дискретным, если $K = \{y_1, \dots, y_s\} \subset Y$; при этом каждому y_j соответствует вероятность q_j и $p \{y \in D_j\} > p_j \ (\forall j \in J)$. Частные случаи: $p_j = p \ (\forall j)$, $q_j = 1/s$, $p = 0,5$.

Замечание. Вообще говоря, не обязательно считать, что если $K = \{x_1, \dots, x_s\}$, то $x_i \neq x_j$ при $i \neq j$, т. е. вместо $K \subset Y$ точнее писать $x_i \in Y \ (\forall i \in N_s)$.

Применим приведенные определения к системам неравенств *обычного* и *сопряженного вида*, а также к задачам распознавания образов. Пусть $X \subset R^n$, $X^* \subset \{R^n \rightarrow R\}$. Рассмотрим системы неравенств

$$x^*(x) \rho 0 \quad (\forall x^* \in X^*), \quad x \in X; \quad (4.3.2)$$

$$x^*(x) \rho 0 \quad (\forall x \in X), \quad x^* \in X^*, \quad (4.3.3)$$

где $\rho \in \{<, \leq, =, >, \geq\}$. Система (4.3.3) относительно неизвестного функционала x^* называется системой *сопряженного вида*.

Определение 4.3.10. *Коллективным решением (обобщенным решением, (z, p) -решением)* системы (4.3.2) (соответственно системы (4.3.3)) называется соответствующее решение системы

$$x \in D_{x^*} = \{x: x^*(x) \rho 0\} \quad (\forall x^* \in X^*), \quad x \in X$$

(соответственно системы

$$x^* \in D_x = \{x^*: x^*(x) \rho 0\} \quad (\forall x \in X), \quad x^* \in X^*).$$

Определение 4.3.11. *Разделяющим* множества $X_1, X_2 \subset R^n$ *коллективным решением (обобщенным решением,*

(z, p) -решением) называется соответствующее решение системы

$$x^*(x) > 0 \quad (\forall x \in X_1), \quad x^*(x) < 0 \quad (\forall x \in X_2), \quad x^* \in X^*.$$

3.2. Некоторые примеры.

1. *Комитет системы оптимальных множеств.* Рассмотрим задачу оптимизации с несколькими целевыми функциями. Пусть допустимое множество $M \subset \mathbf{R}^n$ задается системой неравенств

$$x_1 + \dots + x_n \leq 1, \quad x_1 \geq 0, \dots, x_n \geq 0.$$

Максимизируются функции

$$f_1(x) = x_1 + \dots + x_n, \quad f_2(x) = -x_1, \dots, f_{n+1}(x) = -x_n.$$

Через \bar{M}_j обозначим $\text{Arg max } \{f_j(x) : x \in M\}$. Множества \bar{M}_j — это грани многогранника M . Комитет системы $\{\bar{M}_1, \dots, \bar{M}_{n+1}\}$ — множество всех вершин симплекса M .

2. *Стохастическое программирование.* Одна из постановок задачи стохастического программирования: найти $x \in X$ такое, что

$$p_\omega \{f_j(x, \omega) \leq 0\} \geq p_i \quad (\forall j \in N_m),$$

где ω — вероятностный параметр, p_ω — вероятность выполнения неравенства $f_j(x, \omega) \leq 0$. Если эта задача неразрешима, то можно провести рандомизацию по $x \in X$, отыскивая комитетную конструкцию, а именно — смешанную стратегию.

3. *Идентификация точки.* Пусть $\{x_s\} \rightarrow x'$, тогда множество $\{x_s\}$ является комитетом любой системы окрестностей точки x' .

4. *Обобщенное решение задачи оптимизации.* Рассмотрим задачу

$$\sup \{f(x) : f_j(x) \leq 0 \quad (\forall j \in N_m), \quad x \in C\} = \bar{f}.$$

Обобщенным решением этой задачи здесь назовем последовательность $K = \{x_s\} \subset C$ такую, что $\lim_{(s)} f(x_s) = \bar{f}$ и $\lim_{(s)} f_j(x_s) \leq 0 \quad (\forall j \in N_m)$. Очевидно, обобщенное решение является комитетом системы неравенств

$$f(x) \geq \bar{f} - \varepsilon, \quad f_j(x) \leq \varepsilon \quad (\forall j \in N_m), \quad \varepsilon > 0.$$

5. *Статистические показатели.* Пусть выступающая в статистическом обследовании популяция M каким-либо естественным образом разбита на m подпопуляций:

$$M = \bigcup_{j=1}^m M_j, \quad M_i \cap M_j = \emptyset \quad (i \neq j).$$

Для некоторого предиката P популяция M разбивается на два класса: класс M , для членов x которого предикат P принимает значение «истина»:

$$M' = \{x \in M: P(x) = 1\},$$

и класс

$$M'' = \{x \in M: P(x) = 0\},$$

состоящий из всех остальных членов популяции. Если выполняется ограничение на процентное содержание в каждой подпопуляции элементов, удовлетворяющих предикату P :

$$|\{x \in M_j: P(x) = 1\}| / |M_j| \geq \alpha \quad (\forall j \in N_m),$$

то множество M' является слабым α -комитетом для системы множеств $\{M_j: j \in N_m\}$.

6. *Комитетные конструкции для одной системы нелинейных уравнений.* Пусть

$$x \in \mathbb{R}^2, \quad a = [1, 0] \in \mathbb{R}^2, \quad b = [1, 1] \in \mathbb{R}^2.$$

Рассмотрим несовместную систему уравнений

$$|x|^2 = 1, \quad |x - a|^2 = 1, \quad |x - b|^2 = 1.$$

Выпишем все решения максимальных совместных подсистем этой системы, число которых равно трем:

$$\left\{ \begin{aligned} x_1 &= \left[\frac{1}{2}, \frac{\sqrt{3}}{2} \right], & x_2 &= \left[\frac{1}{2}, -\frac{\sqrt{3}}{2} \right] \end{aligned} \right\}, \\ \left\{ \begin{aligned} y_1 &= \left[1 - \frac{\sqrt{3}}{2}, \frac{1}{2} \right], & y_2 &= \left[1 + \frac{\sqrt{3}}{2}, \frac{1}{2} \right] \end{aligned} \right\}, \\ \{z_1 = [1, 0], \quad z_2 = [0, 1]\}.$$

Минимальные комитеты имеют вид: $K = \{x, y, z\}$, где $x \in \{x_1, x_2\}$, $y \in \{y_1, y_2\}$, $z \in \{z_1, z_2\}$. Один из неминимальных комитетов: $K' = \{x_1, x_2, y_1, y_2, z_1, z_2\}$. Множества вида K и K' являются слабыми $2/3$ -комитетами.

3.3. Существование комитетных конструкций.

Теорема 4.3.1. Если $\psi_j(A)$ и $\psi'_j(A)$ — возрастающие функции от $|A|$, то из существования комитетной конструкции в классе Y для системы F (см. определение 4.3.5) следует существование такой комитетной конструкции, члены которой являются решениями некоторых максимальных совместных подсистем (μ -подсистем) системы (4.3.1).

Доказательство. Пусть $K = \{y_1, \dots, y_q\} \subset Y$ — комитетная конструкция для системы F . Положим $I_i = \{j: y_i \in D_j\}$, I'_i — индекс μ -подсистемы, включающей I_i , z_i — решение этой подсистемы, $K' = \{z_1, \dots, z_q\}$. Тогда

$$\psi_j(K' \cap D_j) \geq \psi_j(K \cap D_j) > \psi'_j(K \setminus D_j) \geq \psi'_j(K' \setminus D_j) \quad (\forall j),$$

что и требовалось.

Эта теорема индуцирует алгоритм отыскания минимальной комитетной конструкции. Пусть z_1, \dots, z_q — решения μ -подсистем, взятые для каждой такой подсистемы по одному. Нужно найти натуральные числа m_1, \dots, m_q (m_j обозначает, сколько экземпляров элемента z_j должны войти в множество K), удовлетворяющие условиям

$$\psi_j(K \cap D_j) > \psi'_j(K \setminus D_j) \quad (\forall j \in J), \quad m_i \in \mathbf{N} \quad (\forall i \in \mathbf{N}_q).$$

В частности, для нахождения обобщенного решения следует решить задачу дискретного программирования:

$$\min \left\{ \sum_{i=1}^q m_i : \sum_{i=1}^q m_i \psi_j(y_i) \geq \right. \\ \left. \geq \alpha_j \sum_{i=1}^q m_i \quad (\forall j \in J), \quad m_i \in \mathbf{N} \quad (\forall i \in \mathbf{N}_q) \right\}.$$

Перейдем к теоремам существования, в которых из некоторых свойств пересечений множеств D_j выводится существование комитета. С этой целью вначале сделаем следующее очевидное замечание.

Для совместности системы

$$y \in D_j \quad (\forall j \in J)$$

необходимо и достаточно, чтобы для всякого $j \in J$ существовала система представителей из одного члена в D_j .

В связи с этим докажем следующую теорему.

Теорема 4.3.2. Если для всякого $j \in N_m$ существует система представителей из r членов ($r > 1$), лежащая в D_j , то система (4.3.1) обладает p -комитетом при $p = r^{-1}$. В частности, при $r = 2$ существует p -комитет при $p = 0,5$.

Доказательство. По условию

$$\forall j \in N_m, \exists M_j \subset D_j: |M_j| = r, r > 1, \\ M_j \cap D_i \neq \emptyset \quad (\forall j, \forall i).$$

Положим $K = \bigcup_{j=1}^m M_j$. Будем считать, что если некоторый элемент x входит в несколько множеств M_j , то во множестве K он повторяется соответствующее число раз. Тогда $|K| = rm$. В то же время $\forall i$

$$|D_i \cap K| \geq |M_i| + \left| \bigcup_{j \neq i} (D_i \cap M_j) \right| \geq r + m - 1.$$

Поэтому $\forall i: |D_i \cap K| / |K| \geq (m + r - 1) / mr > r^{-1}$, что и требовалось.

Очевидно, условие теоремы 4.3.2 не является необходимым для существования p -комитета системы (4.3.1) (пример: система из одного одноэлементного множества).

Теорема 4.3.3. Если всякие k множеств системы F имеют непустое пересечение, то при $k/m > p$ система F обладает p -комитетом.

Доказательство. Пусть $A = \{A_1, \dots, A_r\}$ — система всех пересечений вида $D_{i_1} \cap \dots \cap D_{i_k}, i_t \neq i_u (t \neq u)$; K — система различных представителей для набора A . Тогда K — p -комитет системы F , так как $\forall j$

$$|K \cap D_j| / |K| \geq C_{m-1}^{k-1} / C_m^k = k/m > p.$$

Теорема доказана.

Замечание. Если $\{c_1, \dots, c_m\} \subset \mathbb{R}^n$, причем $\alpha \geq 0 \Rightarrow \alpha c_i + c_j \neq 0$, то при любых $b_j \in \mathbb{R} (j = 1, \dots, m)$ для системы множеств

$$D_j = \{x: (c_j, x) - b_j > 0\} \quad (\forall j \in N_m)$$

выполняется условие теоремы 4.3.2 при $r = 2$.

Доказательство. Так как $\alpha \geq 0 \Rightarrow \alpha c_i + c_j \neq 0$, то

$$\forall j, \exists x_j \neq 0: (c_j, x_j) = 0, \\ \forall i, i \neq j: (c_i, x_j) \neq 0, \\ \exists y_j: (c_j, y_j) = b_j.$$

Пусть $u_j(\lambda, \varepsilon) = y_j + \lambda x_j + \varepsilon c_j$, $v_j(\lambda, \varepsilon) = y_j - \lambda x_j + \varepsilon c_j$. Тогда при достаточно большом $\lambda > 0$ и достаточно малом $\varepsilon > 0$

$$u_j(\lambda, \varepsilon) \in D_j, \quad v_j(\lambda, \varepsilon) \in D_j;$$

$\forall i \neq j$: $\text{sgn}[(c_i, u_j(\lambda, \varepsilon)) - b_i] = -\text{sgn}[(c_i, v_j(\lambda, \varepsilon)) - b_i]$. Поэтому множества $M_j = \{u_j(\lambda, \varepsilon), v_j(\lambda, \varepsilon)\}$ удовлетворяют условию теоремы 4.3.2, что и требовалось.

3.4. Коллективные решения в линейном пространстве. Пусть L — вещественное линейное пространство. Рассмотрим систему линейных неравенств над L :

$$x_j^*(x) > b_j \quad (\forall j \in N_m), \quad (4.3.4)$$

где x_j^* — линейные функционалы, b_j — вещественные числа. Рассмотрим вопрос о существовании коллективного решения системы (4.3.4).

Лемма 4.3.1. Если существует коллективное решение системы

$$x_j^*(x) > 0 \quad (\forall j \in N_m), \quad (4.3.5)$$

то существует и коллективное решение системы (4.3.4).

Доказательство. Пусть $K = \{y_1, \dots, y_q\} \subset L$ есть коллективное решение системы (4.3.5), т. е. при

$$\varphi_j(y) = \begin{cases} 1, & \text{если } x_j^*(y) > 0, \\ -1, & \text{если } x_j^*(y) \leq 0, \end{cases}$$

$$f_j(\varphi_j(y_1), \dots, \varphi_j(y_q)) > 0 \quad (\forall j \in N_m).$$

Подберем $\lambda > 0$ такое, что $[x_j^*(y_i) > 0] \Rightarrow [x_j^*(\lambda y_i) = \lambda x_j^*(y_i) > b_j]$. Тогда, если

$$\varphi_j'(y) = \begin{cases} 1, & \text{если } x_j^*(y) > b_j, \\ -1, & \text{если } x_j^*(y) \leq b_j, \end{cases}$$

то $[\varphi_j(y_i) = 1] \Rightarrow [\varphi_j'(\lambda y_i) = 1]$. Следовательно,

$$[f_j^q(\varphi_j(y_1), \dots, \varphi_j(y_q)) > 0 \quad (\forall j)] \Rightarrow$$

$$\Rightarrow [f_j^q(\varphi_j'(\lambda y_1), \dots, \varphi_j'(\lambda y_q)) > 0 \quad (\forall j)],$$

т. е. $\{\lambda y_1, \dots, \lambda y_q\}$ — коллективное решение для (4.3.4). Лемма доказана.

Пусть n — ранг системы линейных функционалов $\{x_j^* : \forall j \in N_m\}$, тогда $L = L_n + \Gamma$, где L_n — пространство размерности n , Γ — ядро системы $\{x_j^*\}$.

Тогда имеет место следующее простое утверждение.

Лемма 4.3.2. Для существования коллективного решения системы (4.3.5) необходимо и достаточно, чтобы существовало такое решение для системы

$$x_j^*(y) > 0 \quad (\forall j \in N_m), \quad (4.3.6)$$

где $y \in L_n$.

Согласно леммам 4.3.1 и 4.3.2 вопрос о существовании коллективных решений системы (4.3.4) сводится к случаю системы однородных строгих линейных неравенств типа (4.3.6)

$$(c_j, x) > 0 \quad (\forall j \in N_m), \quad (4.3.7)$$

где $x \in R^n$.

Далее рассмотрим вопрос о существовании некоторых частных видов коллективных решений для этой системы. Будем говорить, что совокупность векторов $\{c_j: j \in N_m\}$ удовлетворяет условию (*), если $c_j \neq 0$ ($\forall j$) и $c_j + \alpha c_i \neq 0$ ($\forall \alpha \geq 0, \forall j, i \in N_m$).

Теорема 4.3.4. При $n=2$ и условии () число q всех μ -подсистем системы (4.3.7) нечетно: $q = 2k + 1, q \leq m$. При этом для любого j число тех μ -подсистем, множества решений которых лежат в полуплоскости решений j -го неравенства, равно $k + 1$.*

Доказательство. Не ограничивая общности, предположим, что нумерация полуплоскостей соответствует возрастанию угла от граничной прямой первой полуплоскости. Будем обозначать ориентацию каждой граничной прямой знаком \oplus или \ominus в зависимости от того, с какой стороны от нее находится полуплоскость решений. Ориентация первой прямой соответствует \oplus , следующей прямой (идя по часовой стрелке) — \ominus , если соответствующая полуплоскость расположена по ту же сторону, и \oplus в противном случае.

Тогда последовательности полуплоскостей будет поставлена в соответствие последовательность

$$\underbrace{\oplus, \dots, \oplus}_{m_1}, \quad \underbrace{\ominus, \dots, \ominus}_{m_2}, \quad \underbrace{\oplus, \dots, \oplus}_{m_3}, \dots$$

Возможны два случая: когда последовательность заканчивается знаком \ominus , и когда она заканчивается знаком \oplus . Индукцией по числу перемен знака легко доказать, что число всех μ -подсистем равно числу перемен

знака в первом случае, и больше на 1 во втором. Отсюда и следует утверждение теоремы.

Доказанная теорема устанавливает *свойство дуальности* для систем однородных линейных неравенств над пространством \mathbf{R}^n . Оно состоит в том, что решения μ -подсистем, взятые по одному для каждой из них, составляют комитет. Это свойство справедливо и для некоторых классов систем над \mathbf{R}^n ; в то же время существуют примеры систем однородных линейных неравенств, удовлетворяющих условию (*), но не обладающих свойством дуальности.

Будем говорить, что система (4.3.7) *удовлетворяет условию (**)*, если $\forall k \in \mathbf{N}_m, \forall \lambda > 0, \exists \bar{\mathbf{x}} = \bar{\mathbf{x}}(k, \lambda)$:

$$(c_j, \bar{\mathbf{x}}) > 0 \quad (\forall j \neq k), \quad (c_k, \bar{\mathbf{x}}) < 0,$$

$$\min_{j \neq k} \frac{|(c_k, \bar{\mathbf{x}})|}{(c_j, \bar{\mathbf{x}})} > \lambda.$$

Условию (**), удовлетворяет, например, система

$$x_i > 0 \quad (\forall i \in \mathbf{N}_n).$$

Рассмотрим несовместную систему

$$y_j(x_1, \dots, x_n) > 0 \quad (\forall j \in \mathbf{N}_m), \quad (4.3.8)$$

где $y_j(\cdot)$ — линейные функционалы.

Теорема 4.3.5. Пусть μ -подсистема системы (4.3.8) с наибольшим числом неравенств удовлетворяет условию (**). Тогда число всех μ -подсистем системы (4.3.8), в которые входит фиксированное неравенство, больше числа всех остальных μ -подсистем.

Доказательство. Не уменьшая общности, предположим, что μ -подсистема системы (4.3.8) с наибольшим числом неравенств (пусть этим числом будет p) имеет вид

$$y_j(\mathbf{x}) > 0 \quad (\forall j \in \mathbf{N}_p). \quad (4.3.9)$$

Докажем, что при $k > p$ справедливо тождество

$$y_k(\mathbf{x}) \equiv \sum_{i=1}^p \lambda_{ki} y_i(\mathbf{x}) \quad (\lambda_{ji} \leq 0). \quad (4.3.10)$$

Действительно, при $k > p$ неравенство $y_k(\mathbf{x}) \leq 0$ есть следствие системы (4.3.9), а следовательно, и системы

$$y_j(\mathbf{x}) \geq 0 \quad (\forall j \in \mathbf{N}_p).$$

Тогда тождество (4.3.10) вытекает из теоремы Фаркаша [111].

Пусть $I_k = \{i: \lambda_{ki} < 0\}$. Ясно, что $|I_k| \geq 1$, ибо иначе нарушалось бы свойство (*). Тогда при $k > p$

$$y_k(x) \equiv \sum_{i \in I_k} \lambda_{ki} y_i(x) \quad (\lambda_{ki} < 0).$$

Докажем, что $k \neq l$ влечет $I_k \cap I_l = \emptyset$. Действительно, пусть, от противного, $k \neq l$, но $\exists s \in I_k \cap I_l$. Положим

$$\lambda = p \max_{i \neq s} \left\{ \frac{\lambda_{ki}}{\lambda_{ks}}, \frac{\lambda_{li}}{\lambda_{ls}} \right\}.$$

По условию (**)

$$\begin{aligned} \exists \bar{x} = \bar{x}(s, \lambda): y_j(\bar{x}) > 0 \\ (j = 1, \dots, s-1, s+1, \dots, p), \end{aligned}$$

$y_s(\bar{x}) < 0$. Тогда $\min_{j \neq s} |y_s(\bar{x})| |y_j^{-1}(\bar{x})| > \lambda$ и

$$\begin{aligned} y_k(\bar{x}) &= \sum_{j \in I_k} \lambda_{kj} y_j(\bar{x}) = \sum_{j \in I_k \setminus \{s\}} \lambda_{kj} y_j(\bar{x}) + \lambda_{ks} y_s(\bar{x}) = \\ &= |\lambda_{ks}| \cdot |y_s(\bar{x})| - \sum_{j \in I_k \setminus \{s\}} |\lambda_{kj}| |y_j(\bar{x})| \geq \\ &\geq |\lambda_{ks}| \cdot |y_s(\bar{x})| - p \max_{i \neq s} |\lambda_{ki}| \max_{j \neq s} |y_j(\bar{x})| \geq \\ &\geq |\lambda_{ks}| \cdot |y_s(\bar{x})| - \lambda |\lambda_{ks}| \max_{j \neq s} |y_j(\bar{x})| > 0. \end{aligned}$$

Точно так же доказывается, что $y_l(\bar{x}) > 0$. Следовательно, система

$$y_j(x) > 0 \quad (j = 1, \dots, s-1, s+1, \dots, p; k, l)$$

совместна. Но это противоречит условию, что число неравенств в μ -подсистеме не больше p .

Таким образом, можно считать, не ограничивая общности, что система (4.3.8) имеет вид

$$x^{(1)} > 0, (c_1, x^{(1)}) > 0, \dots, x^{(r)} > 0, (c_r, x^{(r)}) > 0, \quad (4.3.11)$$

где $x^{(i)} \in \mathbb{R}^{k_i}$, $c_i < 0$ ($\forall i \in N_r$).

Найдем все μ -подсистемы системы (4.3.11). Каждая подсистема

$$x^{(i)} > 0, \quad (c_i, x^{(i)}) > 0 \quad (4.3.12)$$

имеет $k_i + 1$ подсистем: одна из них —

$$x^{(i)} > 0; \quad (S_i)$$

остальные имеют вид

$$\left. \begin{aligned} x_j^{(k)} > 0 \quad (j \neq k), \\ (c_i, x^{(k)} > 0, \end{aligned} \right\} \quad (T_{ik})$$

$k = 1, \dots, k_i$; здесь $x_j^{(k)}$ — j -я координата вектора $x^{(k)}$.

То, что каждая из подсистем T_{ik} совместна, гарантируется условием (**).

Итак, все μ -подсистемы системы (4.3.11) имеют структуру

$$R_1 \cup \dots \cup R_r,$$

где $R_i \in \{S_i, T_{i1}, \dots, T_{ik_i}\}$.

Тогда число ν всех μ -подсистем системы (4.3.11) равно

$$\prod_{i=1}^r (k_i + 1),$$

а число всех μ -подсистем системы (4.3.12), не содержащих неравенства $x_s^{(i)} > 0$, равно 1. Следовательно, число всех таких подсистем равно

$$\prod_{\substack{s=1, \\ s \neq i}}^r (k_s + 1),$$

что меньше $\nu/2$.

Так как число всех μ -подсистем системы (4.3.12), не содержащих неравенства $(c_s, x^{(s)}) > 0$, равно 1, то число всех μ -подсистем системы (4.3.11), не содержащих этого неравенства, меньше чем $\nu/2$. Следовательно, теорема доказана.

Теорема 4.3.6. Если система (4.3.7) удовлетворяет условию (*), то она обладает p -комитетом при $p > \frac{m+1}{2m}$, в частности и комитетом.

Доказательство. Не уменьшая общности, можно считать, что среди векторов c_j нет одинаково направленных, так как для всякой пары неравенств с одинаково направленными c_j можно удалить одно из этих неравенств из системы, не меняя существа дела.

Вместе с условием (*) это предположение означает, что среди векторов c_j нет коллинеарных. Тогда $\forall j \in N_m, \exists x_j (c_j, x_j) = 0, (c_i, x_j) \neq 0 (i \neq j)$. Легко проверить, что

при достаточно малом $\varepsilon > 0$ множество

$$K = \{x_1 + \varepsilon c_1, -x_1 + \varepsilon c_1, \dots, x_m + \varepsilon c_m, -x_m + \varepsilon c_m\}$$

является комитетом системы (4.3.7). При этом всякому неравенству удовлетворяют не менее $m+1$ членов этого $2m$ -членного комитета, т. е. K есть p -комитет при $p > \frac{(m+1)}{2m}$, что и требовалось.

Теорема 4.3.7. *z -решение системы (4.3.7) существует тогда и только тогда, когда существует ее комитет.*

Доказательство. Всякий комитет является, очевидно, некоторым z -решением. Докажем обратное утверждение. Пусть множество

$$\{x_1, \dots, x_q, z\} \quad (x_i \in \mathbb{R}^n, z \in \mathbb{R}^q)$$

есть z -решение системы (4.3.7). Тогда существует ее z -решение, для которого $z \geq 0$. Действительно, если $z = [z_1, \dots, z_q]$, причем некоторое $z_k < 0$, то можно положить $z'_k = -z_k$, $x'_k = -x_k$, и новое множество

$$\{x_1, \dots, x_{k-1}, x'_k, x_{k+1}, \dots, x_q, z' = [z_1, \dots, z_{k-1}, z'_k, z_{k+1}, \dots, z_q]\}$$

будет z -решением.

Далее, если в общем случае z' — нецелочисленный вектор, то его можно с какой угодно степенью точности приблизить вектором z'' с рациональными компонентами, причем множество

$$\{x_1, \dots, x_q, z''\}$$

будет z -решением. Но тогда при любом $\lambda > 0$ множество

$$\{x_1, \dots, x_q, \lambda z''\}$$

будет z -решением. Поэтому остается только подобрать $\lambda > 0$ такое, чтобы вектор $\lambda z''$ был целочисленным. Теорема доказана.

Теорема 4.3.8. *При $p < \frac{m+1}{2m+1}$ p -комитет системы (4.3.5) (т. е., в частности, и комитет) существует тогда и только тогда, когда для всех i и j из множества N_m и для всех $\alpha \leq 0$ выполняется условие*

$$x_i^*(x) \neq \alpha x_j^*(x). \quad (4.3.13)$$

Доказательство. Необходимость очевидна. Докажем достаточность. Согласно лемме 4.3.1 достаточно доказать, что существует комитет системы (4.3.6). Согласно лемме 4.3.2 достаточно доказать, что существует комитет системы (4.3.7). Легко проверить, что из того, что система (4.3.6) удовлетворяет условию (4.3.13), следует, что система (4.3.7) обладает свойством (*). Тогда комитет системы (4.3.7) существует согласно теореме 4.3.6, что и требовалось.

Замечание. Для системы (4.3.7) условие (4.3.13) запишется в форме

$$[\alpha \geq 0] \Rightarrow [c_i + \alpha c_j \neq 0], \quad \forall i, j \in N_m. \quad (4.3.14)$$

Теорема 4.3.9. Если система (4.3.5), т. е. система

$$x_j^*(x) > 0 \quad (\forall j \in N_m), \quad (4.3.15)$$

удовлетворяет необходимому и достаточному условию (4.3.13) существования комитета, то число q членов минимального комитета не больше числа неравенств, т. е. т.

Доказательство. Согласно лемме 4.3.2 достаточно доказать теорему для случая системы (4.3.7) над конечномерным пространством R^n .

Тогда выполняется условие (4.3.14). Не ограничивая общности, можно предполагать, что

$$[\alpha \geq 0, i \neq j] \Rightarrow [c_i - \alpha c_j \neq 0],$$

т. е., таким образом,

$$\forall \alpha \in R: [i \neq j] \Rightarrow [c_i + \alpha c_j \neq 0].$$

Покажем, что существует матрица $A = A_{n \times 2}$ такая, что

$$\forall \alpha \in R: [i \neq j] \Rightarrow [c_i A + \alpha c_j A \neq 0 \in R^2].$$

Обозначим столбцы искомой матрицы A через $u \in R^n$, $v \in R^n$:

$$A = [u, v]; \quad c_j A = [(c_j, u), (c_j, v)] \in R^2.$$

В качестве u выберем такой вектор, что $\forall j: (c_j, u) \neq 0$. Введем обозначение: $\alpha_{ij} = (c_i, u)(c_j, u)^{-1}$. По условию $[i \neq j] \Rightarrow [c_i - \alpha_{ij} c_j \neq 0]$. Тогда существует вектор $v \in R^n$ такой, что

$$(c_i - \alpha_{ij} c_j, v) \neq 0 \quad (i \neq j).$$

Найденные векторы u , v и образуют столбцы искомой матрицы A .

Рассмотрим систему над \mathbb{R}^2 :

$$(c_j A, h) > 0 \quad (\forall j \in N_m), \quad h \in \mathbb{R}^2.$$

Согласно теореме 4.3.4 существует комитет этой системы: $K' = \{h_1, \dots, h_q\}$ с числом членов q , не большим числа m . Тогда множество

$$K = \{h_1 A^T, \dots, h_q A^T\}$$

является комитетом системы (4.3.7), а потому и (4.3.15), так как

$$(c_j A, h_i) = (c_j, h_i A^T).$$

Теорема доказана.

Из теорем 4.3.4 и 4.3.9 вытекает следующее утверждение.

Теорема 4.3.10. Пусть $\{D_j: \forall j \in N_m\}$ — система собственных замкнутых или открытых полупространств пространства \mathbb{R}^n , т. е. $D_j = \{x: (c_j, x) - b_j \leq 0\}$ либо $D_j = \{x: (c_j, x) - b_j < 0\}$; $G = \{T_1, \dots, T_k\}$ — совокупность множеств решений μ -подсистем системы

$$x \in D_j \quad (\forall j \in N_m); \quad (4.3.16)$$

$R = \{x_1, \dots, x_k\}$ — система различных представителей для G . Если выполнено условие (4.3.14), то существуют целые числа $l_i \geq 0$ ($\forall i \in N_k$) такие, что набор векторов

$$\underbrace{\{x_1, \dots, x_1\}}_{l_1}; \dots; \underbrace{\{x_k, \dots, x_k\}}_{l_k}$$

является комитетом системы (4.3.16). При $n=2$ само множество R есть комитет рассматриваемой системы (4.3.16).

Рассмотрим вопрос о структуре множества всех обобщенных решений системы (4.3.1), предполагая множество J конечным: $J = N_m$.

Назовем сигнатурой элемента $y \in Y$ вектор

$$\sigma(y) = [\varphi_1(y), \dots, \varphi_m(y)] \in \mathbb{R}^m,$$

где $\varphi_j(\cdot)$ — характеристическая функция множества D_j . Множество всех возможных сигнатур $\{\sigma(y): y \in Y\}$ содержит лишь конечное число различных элементов; обо-

значим их $\sigma_1, \dots, \sigma_q$. Пусть $\alpha = [\alpha_1, \dots, \alpha_m]$. Рассмотрим систему

$$\left. \begin{aligned} \sum_{i=1}^q z_i \sigma_i > (2\alpha - \vartheta) |z|_1, \\ z_i \in \mathbb{N} \quad (\forall i \in N_q), \end{aligned} \right\} \quad (4.3.17)$$

где $|z|_1 = \sum_{i=1}^q |z_i|$, $\vartheta = [1, \dots, 1]$.

Каждому решению $z^0 = [z_1^0, \dots, z_q^0]$ этой системы соответствует множество обобщенных решений

$$\underbrace{K_1 \times \dots \times K_1}_{z_1^0} \times \dots \times \underbrace{K_q \times \dots \times K_q}_{z_q^0} = K(z^0),$$

где

$$K_i = \{y \in Y: \sigma(y) = \sigma_i\}.$$

Все множество обобщенных решений имеет вид

$$K = \bigcup_{z \in Z} K(z);$$

здесь Z — множество всех решений системы (4.3.17).

3.5. Теоремы существования разделяющих комитетов. Пусть $X_1, X_2 \subset \mathbb{R}^n$, $|X_1| < +\infty$, $|X_2| < +\infty$.

Теорема 4.3.11. *Для существования комитета линейных функционалов, разделяющего конечные множества X_1 и X_2 , необходимо и достаточно, чтобы среди элементов множества $X_1 \cup (-X_2)$ не было нулевых и противоположно направленных. Минимальный разделяющий комитет состоит не более чем из $|X_1 \cup X_2|$ членов.*

Справедливость этого утверждения следует из теоремы 4.3.8.

Теорема 4.3.12. *Для существования комитета аффинных функционалов, разделяющего конечные множества $X_1, X_2 \subset \mathbb{R}^n$, необходимо и достаточно, чтобы $X_1 \cap X_2 = \emptyset$. Минимальный разделяющий комитет состоит не более чем из $|X_1 \cup X_2|$ членов.*

Доказательство. Необходимость очевидна. Докажем достаточность. Пусть $X_1 \cap X_2 = \emptyset$. Рассмотрим систему относительно $[x^*, y^*] \in \mathbb{R}^{n+1}$:

$$(x^*, x) > y^* \quad (\forall x \in X_1), \quad (x^*, x) < y^* \quad (\forall x \in X_2).$$

Очевидно, среди векторов $[x, -1]$ ($x \in X_1$) и $[-x, +1]$ ($x \in X_2$) нет нулевых и противоположно направленных. Поэтому утверждение теоремы следует из теоремы 4.3.8.

3.6. Алгоритмы построения комитетов (некоторые замечания). Заметим, во-первых, что доказательства теорем существования комитета 4.3.4, 4.3.6, 4.3.9 порождают и сами способы их построения. Можно предложить и другие способы построения комитетов, в том числе эвристические. Отметим, например, два из них: метод, использующий решение задачи безусловной оптимизации, и метод коррекции.

Первый состоит в следующем: решения μ -подсистем системы (4.3.7) отыскиваются как локальные максимумы

функции $f(x) = \sum_{j=1}^m \text{sign}(c_j, x)$, которая может быть заменена некоторым гладким приближением, например:

$g(x) = \sum_{j=1}^m [1 + \exp(-\alpha(c_j, x))]^{-1}$, где α — достаточно большое число.

Другой алгоритм состоит в использовании обычной схемы метода коррекции.

Из векторов c_1, \dots, c_m составляется бесконечная последовательность с циклическим их повторением:

$$d_1 = c_1, \dots, d_m = c_m, d_{m+1} = c_1, \dots \\ \dots, d_{2m} = c_m, d_{2m+1} = c_1, \dots$$

Пусть $x_1(0), \dots, x_q(0)$ — произвольные векторы из \mathbb{R}^n . Если $x_1(t), \dots, x_q(t)$ уже построены, то вычисляем

$$N_t = \sum_{j=1}^q \text{sign}(x_j(t), d_t).$$

Если $N_t > 0$, то $x_j(t+1) = x_j(t)$ ($\forall j \in N_q$). Если же $N_t \leq 0$, то выбираем $|N_t|/2 + 1$, если $|N_t|$ четно, и $(|N_t| + 1)/2$, если $|N_t|$ нечетно, векторов $x_j(t)$, имеющих наименьшие величины $|(x_j(t), d_t)|$. Каждый из этих векторов изменяется по правилу: $x_j(t+1) = x_j(t) + d_t$.

Этот метод имеет ряд достоинств (простота реализации, хорошая интерпретируемость).

Вместе с тем, помимо его эвристичности, имеется трудность, связанная с выбором числа q .

3.7. Алгоритм построения минимального комитета в \mathbb{R}^n . Применим следующий метод свертывания С. Н. Черникова [105]. Рассмотрим систему

$$\begin{aligned} a_{j_1}x_1 + \dots + a_{j_n}x_n - a_j &> 0 \quad (\forall j \in N_{m'}), \\ a_{j_1}x_1 + \dots + a_{j_n}x_n - a_j &\geq 0 \quad (\forall j \in N_m \setminus N_{m'}), \end{aligned} \quad (4.3.18)$$

где $m \geq m'$.

Пусть U — какое-либо подпространство из \mathbb{R}^n , $l_j(x) = \sum_{i=1}^n a_{ji}x_i$,

$$C(U) = \left\{ y = [y_1, \dots, y_m] \geq 0: \sum_{j=1}^m y_j l_j(x) = 0 \quad (\forall x \in U) \right\},$$

и пусть

$$\{d_i = [\delta_{1i}, \dots, \delta_{mi}]: \forall i \in N_k\} \quad (4.3.19)$$

— некоторая конечная система ненулевых порождающих элементов конуса $C(U)$. Элементу d_i поставим в соответствие неравенство

$$\sum_{j=1}^m \delta_{ji} l_j(x) - \sum_{j=1}^m \delta_{ji} a_j \geq 0,$$

если $m' = 0$ или $\delta_{1i} = \dots = \delta_{m'i} = 0$ при $m' > 0$, и неравенство

$$\sum_{j=1}^m \delta_{ji} l_j(x) - \sum_{j=1}^m \delta_{ji} a_j > 0,$$

если $m' > 0$ и хотя бы один из элементов $\delta_{1i}, \dots, \delta_{m'i}$ отличен от нуля.

Система всех получаемых так неравенств называется U -сверткой системы (4.3.18). Если конус $C(U)$ нулевой, то говорят, что U -свертка пуста. Если система (4.3.19) есть базис конуса $C(U)$, то соответствующая U -свертка называется фундаментальной. При $U = \mathbb{R}^n$ U -свертка называется полной.

Индексом ненулевого элемента конуса $C(U)$ называется совокупность номеров ненулевых координат этого элемента. Индексами неравенств U -свертки называются индексы элементов конуса $C(U)$, которые ее определяют. Нам понадобится следующее утверждение [112].¹

Теорема А. Если система (4.3.18) совместна, то совместна или пуста ее U -свертка по всякому подпрост-

ранству $U \subset \mathbb{R}^n$. Система (4.3.18) совместна, если совместна или пуста ее U -свертка хотя бы для одного подпространства $U \subset \mathbb{R}^n$.

Пусть $B = \{b_1, \dots, b_m\} \subset \mathbb{R}$. Каждой паре элементов $b_p > 0$ и $b_q < 0$ соотнесем неравенство

$$b_p(l_q(\mathbf{x}) - a_q) - b_q(l_p(\mathbf{x}) - a_p) > 0,$$

если $m' > 0$ и хотя бы один из индексов p, q не больше m' , или неравенство

$$b_p(l_q(\mathbf{x}) - a_q) - b_q(l_p(\mathbf{x}) - a_p) \geq 0$$

— в противном случае. Каждому нулевому элементу $b_s \in B$ отнесем неравенство $l_s(\mathbf{x}) - a_s > 0$, если $m' > 0$ и $s \leq m'$, или неравенство $l_s(\mathbf{x}) - a_s \geq 0$ в противном случае. Говорят, что совокупность S полученных при этом неравенств (она, очевидно, может быть и пустой) получена в результате B -деформации системы (4.3.18).

Связь этого определения с вопросом об U -свертках системы (4.3.18) вытекает из следующего очевидного утверждения.

Если $\mathbf{x}_0 \in \mathbb{R}^n$, $U(\mathbf{x}_0)$ — одномерное подпространство $\{\gamma \mathbf{x}_0 : \gamma \in \mathbb{R}\}$, $B(\mathbf{x}_0) = \{b_j : b_j = l_j(\mathbf{x}_0), \forall j \in \mathbb{N}_m\}$, то $B(\mathbf{x}_0)$ -деформация системы (4.3.18) дает некоторую ее $U(\mathbf{x}_0)$ -свертку.

Если \mathbf{e}_i — единичный орт, то обозначим $U(\mathbf{e}_i) = U_i$.

Опишем алгоритм фундаментального свертывания [105].

1. Пусть \mathbf{p}_1 — какой-либо столбец коэффициентов системы (4.3.18), например столбец коэффициентов при x_1 . Проведя \mathbf{p}_1 -деформацию системы (4.3.18), получим систему неравенств S_1 , которая является U_1 -сверткой (очевидно, фундаментальной). Каждому неравенству системы S_1 отнесем его индекс; он совпадает с множеством номеров тех неравенств системы (4.3.18), комбинированием которых оно получено.

2. Пусть уже получена непустая фундаментальная $(U_1 + \dots + U_k)$ -свертка S_k системы (4.3.18), и \mathbf{p}_{k+1} — некоторый столбец ее коэффициентов, например столбец коэффициентов при x_{k+1} . Проведем \mathbf{p}_{k+1} -деформацию системы S_k , не комбинируя при этом те пары неравенств, объединение индексов которых содержит индекс какого-нибудь третьего неравенства системы S_k . В результате полу-

чим систему S_{k+1} — фундаментальную $(U_1 + \dots + U_{k+1})$ -свертку системы (4.3.18). Каждому ее неравенству отнесем его индекс.

3. Если $S_k = \emptyset$, то полагаем $S_{k+1} = \emptyset$. Так как

$$U_1 + \dots + U_n = \mathbb{R}^n,$$

то через n шагов процесс дает полную фундаментальную свертку системы (4.3.18).

Теорема В [105]. Номера неравенств несовместной системы (4.3.18), которые входят в произвольную ее минимально несовместную подсистему, образуют индекс одного из несовместных неравенств полной фундаментальной свертки системы (4.3.18). Обратно, номера неравенств системы (4.3.18), которые входят в индекс произвольного несовместного неравенства ее полной фундаментальной свертки, суть номера неравенств некоторой минимальной несовместной подсистемы системы (4.3.18).

Заметим, что выше под несовместным неравенством понимается либо $a \geq 0$ с $a < 0$, либо $a > 0$ с $a \leq 0$.

Следовательно, подсистема T несовместной системы (4.3.18) тогда и только тогда совместна, когда совокупность номеров ее неравенств, входящих в подсистему T , не включает индекса ни одного из несовместных неравенств полной фундаментальной свертки.

Отсюда вытекает, что все μ -подсистемы системы (4.3.18) могут быть найдены с помощью следующего алгоритма:

1) находим индексы несовместных неравенств полной фундаментальной свертки, применяя алгоритм фундаментального свертывания;

2) отыскиваем максимальные подсистемы, не включающие этих индексов несовместных неравенств.

Отсюда вытекает, что минимальный комитет системы (4.3.18) может быть построен с помощью следующего алгоритма:

1) находим индексы несовместных неравенств полной фундаментальной свертки;

2) отыскиваем μ -подсистемы как максимальные подсистемы, не включающие индексов несовместных неравенств свертки;

3) строим систему различных представителей $G = \{x_1, \dots, x_q\}$ для совокупности множеств решений μ -подсистем;

4) строим минимальный комитет из системы G , решая задачу дискретного относительно $z_i \geq 0$ программирования

$$\min \left\{ \sum_{i=1}^q z_i : \sum_{i=1}^q z_i \sigma(x_i) \geq v, z_i \in \mathbb{N} (\forall i) \right\},$$

а именно: если $\{\bar{z}_i\}$ — решение этой задачи, то минимальным комитетом будет

$$\underbrace{\{x_1, \dots, x_1\}}_{z_1}; \dots; \underbrace{\{x_q, \dots, x_q\}}_{z_q}.$$

§ 4. Комитетное решение задачи многокритериальной оптимизации

4.1. Некоторые направления приложений комитетных конструкций. Комитетные конструкции могут применяться при решении задач математического программирования с противоречивой системой ограничений и с векторным критерием, выступая в качестве аналога смешанных стратегий использования планов, удовлетворяющих подсистемам условий. Кроме того, комитетные конструкции могут быть применены для моделирования сложных зависимостей, связывающих переменные [в задаче оптимизации].

Отметим также эффективность комитетных конструкций как решающих правил в задачах принятия решений, в том числе в таких областях, как классификация, исследование механизмов групповых решений и др.

4.2. Комитетные решения вариационных неравенств. Задача нахождения

$$\arg \inf \{(c, x) : x \in X\} \quad (4.4.1)$$

эквивалентна задаче решения вариационного неравенства $x \in X, (c, x - y) \leq 0 \quad (\forall y \in X)$.

Комитетные конструкции при решении этой задачи могут применяться в следующих случаях.

1. Имеются несколько целей, т. е. рассматривается система

$$x \in X, (c_j, x - y) \leq 0 \quad (\forall y \in X, \forall j \in N_m), \quad (4.4.2)$$

которая, как правило, несовместна.

2. Множество $X = X_1 \cap \dots \cap X_m$, задаваемое в форме пересечения, пусто. Тогда можно применять комитетное решение несовместной системы

$$x \in X_1, \dots, x \in X_m, (c, x - y) \leq 0 \quad (\forall y \in X).$$

3. X незамкнуто и \inf в (4.4.1) не достигается. В этом случае можно использовать понятие бесконечного комитета в качестве аналога обобщенного решения (сходящейся подпоследовательности).

Рассмотрим случай 1, т. е. задачу (4.4.2). Пусть $K = \{x_1, \dots, x_q\} \subset X$ — комитет системы (4.4.2). Тогда для всех $y \in X$ и $j \in N_m$ неравенство

$$(c_j, x_i) \leq (c_j, y)$$

выполняется более чем для половины номеров $i \in N_q$.

Предположим, что X — выпуклый многогранник: $X = \text{co} \{y_1, \dots, y_p\}$ и y_1, \dots, y_p — все его вершины. Положим $Y = \{y_1, \dots, y_p\}$. Задача (4.4.1) в разрешимом случае эквивалентна задаче нахождения

$$\arg \inf \{(c, y) : y \in Y\}$$

или задаче решения вариационного неравенства

$$x \in Y, (c, x - y) \leq 0 \quad (\forall y \in Y).$$

Задача (4.4.2) в этом случае принимает вид

$$x \in Y, (c_j, x - y) \leq 0 \quad (\forall y \in Y, \forall j \in N_m). \quad (4.4.3)$$

К этой же постановке приводятся некоторые задачи дискретного программирования.

Итак, рассмотрим задачу (4.4.3) с конечным множеством Y .

Найдем вначале условия, при которых $K = Y$, т. е. при всех $j \in N_m$ и $i \in N_p$ неравенство

$$(c_j, y_s) \leq (c_i, y_i)$$

выполняется более чем для половины номеров $s \in N_p$.

Рассмотрим матрицу

$$A = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1p} \\ \cdot & \cdot & \cdot \\ \alpha_{m1} & \dots & \alpha_{mp} \end{pmatrix},$$

где $\alpha_{jt} = (c_j, y_t)$.

Матрица A должна обладать следующим свойством (*): для каждого ее элемента α_{ji} более половины элементов j -й строки не превосходят α_{ji} .

Теорема 4.4.1. *Матрица A обладает свойством (*) тогда и только тогда, когда множество минимальных элементов каждой строки содержит более $r/2$ элементов.*

Доказательство. Упорядочим элементы j -й строки:

$$\alpha_1 \geq \dots \geq \alpha_k \geq \alpha_{k+1} \geq \dots \geq \alpha_{2k+1}$$

(предполагая r нечетным). Тогда свойство (*), очевидно, выполняется для $\alpha_1, \dots, \alpha_{k+1}$. Это свойство выполняется и для $\alpha_{k+2}, \dots, \alpha_{2k+1}$, если и только если $\alpha_{k+1} = \alpha_{k+2} = \dots = \alpha_{2k+1}$.

Если же r четно, то пусть j -я строка имеет вид

$$\alpha_1 \geq \dots \geq \alpha_k \geq \alpha_{k+1} \geq \dots \geq \alpha_{2k}.$$

Тогда свойство (*) очевидным образом выполняется для $\alpha_1, \dots, \alpha_k$. Оно выполняется и для $\alpha_{k+1}, \dots, \alpha_{2k}$, если и только если $\alpha_k = \dots = \alpha_{2k}$. Теорема доказана.

Найдем условия, при которых существует комитет $K \subset Y$. В матрице A должна существовать система столбцов с номерами i_1, \dots, i_q такая, что при всех $j \in N_m$ и $i \in N_p$ неравенство $\alpha_{j i_s} \leq \alpha_{ji}$ выполняется более чем для половины номеров $i_s \in \{i_1, \dots, i_q\}$. Такую систему столбцов назовем *комитетом столбцов матрицы A* . *Оптимальным множеством j -й строки* будем называть $J_j = \text{Arg min}_i \{\alpha_{ji} : i \in N_p\}$.

Теорема 4.4.2. *Если для некоторого $k > m/2$ всякие k множеств из J_1, \dots, J_m имеют непустое пересечение, то существует комитет столбцов матрицы A .*

Доказательство. Во всяком пересечении k множеств из J_1, \dots, J_m возьмем по одному элементу; получим множество $\{i_1, \dots, i_q\}$, где $q = C_m^k$ (некоторые номера при этом могут повторяться). Рассмотрим j -ю строку. Число тех из элементов $\alpha_{ji_1}, \dots, \alpha_{ji_q}$, которые принадлежат J_j , равно C_{m-1}^{k-1} . Тогда $\alpha_{ji} \geq \alpha_{j i_s} (\forall i_s \in J_j)$, причем

$$|\{i_s \in J_j\}| = C_{m-1}^{k-1} / C_m^k = k/m > 1/2.$$

Теорема доказана.

Условие этой теоремы не является необходимым, это видно на примере следующей матрицы:

$$A = \begin{pmatrix} 0 & 9 & \boxed{12} & \boxed{0} & 0 & 3 & \boxed{0} \\ -2 & -3 & \boxed{-4} & \boxed{-4} & -2 & -4 & \boxed{0} \\ 4 & 0 & \boxed{0} & \boxed{8} & 4 & 6 & \boxed{0} \end{pmatrix}.$$

Здесь $J_1 = \{1, 4, 5, 7\}$, $J_2 = \{3, 4, 6\}$, $J_3 = \{2, 3, 7\}$, и условие теоремы не выполняется. Однако комитет столбцов существует, он состоит из выделенных столбцов.

Если множество X бесконечно, но выпукло, замкнуто и ограничено, то систему (4.4.2) можно представить в виде

$$x \in X, \quad (c_j, x) \leq \alpha_j \quad (\forall \alpha_j \in \overline{\underline{\alpha}_j}, \forall j \in N_m), \quad (4.4.4)$$

где

$$\underline{\alpha}_j = \min \{(c_j, y) : y \in X\}, \quad \overline{\alpha}_j = \max \{(c_j, y) : y \in X\}.$$

Комитет системы (4.4.4) существует тогда и только тогда, когда существует комитет ее подсистемы

$$x \in X, \quad (c_j, x) \leq \underline{\alpha}_j \quad (\forall j \in N_m).$$

Однако комитет этой системы, лежащий в множестве X , существует лишь в исключительных случаях. Поэтому необходимо ослабление требований в формализации многокритериальной оптимизации.

Рассмотрим (пока чисто формально) задачу: найти $\sup \{u(x) = [u_1(x), \dots, u_m(x)] \in R^m : x \in X \subset R^n\}$, (4.4.5)

где $u_j(x) = (c_j, x)$. Смысл этой задачи может быть уточнен следующими конструкциями. Определим:

$$\begin{aligned} \alpha_j &= \max \{(c_j, x) : x \in X\}, \\ E_j(\varepsilon) &= \{x : (c_j, x) \geq \alpha_j - \varepsilon\}, \\ D_j(\varepsilon) &= E_j(\varepsilon) \cap X, \\ D(\varepsilon) &= \{D_1(\varepsilon), \dots, D_m(\varepsilon)\}, \\ \delta &= \min \{\varepsilon \geq 0 : D(\varepsilon) \text{ обладает комитетом}\}. \end{aligned}$$

Тогда предлагается следующая формализация понятия решения задачи (4.4.5): под последним понимать смешанную стратегию использования δ -членов комитета системы $D(\delta)$.

При этом легко видеть, что если X ограничено, то $\delta < +\infty$. Действительно, если X ограничено, то для всех j существует $\delta_j: E(\delta_j) \supset X$. Пусть $\delta = \max \delta_j$. Тогда

$\bigcap_{j \in N_m} D_j(\delta) \neq \emptyset$ и, следовательно, система $\bar{D}(\delta)$ обладает комитетом из одного члена.

Численная реализация этой конструкции состоит в следующем. Находим

$$[\bar{x}; \bar{t}] = \arg \min \{t \geq 0: (c_j, x) \geq \alpha_j - t (\forall j \in N_m)\}.$$

Будем предполагать, что $\bar{x} \in \text{int } X$. Далее выбираем достаточно малое $\varepsilon > 0$ и строим комитет $K = \{x_1, \dots, x_q\}$ системы

$$(c_j, x) \geq \alpha_j - \bar{t} + \varepsilon \quad (\forall j \in N_m).$$

В качестве решения проблемы согласования целей выступает множество

$$\{\bar{x}_1 + \lambda_1 x_1, \dots, \bar{x} + \lambda_q x_q\},$$

где $\lambda_i = \max \{\lambda: \bar{x} + \lambda x_i \in X\}$.

Что касается случая 2, то ему соответствует возможность использования понятия комитета системы множеств и обобщений этого понятия.

Рассмотрим, наконец, вариационное неравенство при условиях, отвечающих случаю 3, т. е. когда множество X является открытым, \inf в задаче (4.4.1) не достигается и, таким образом, для нее не существует решения в обычном смысле.

Тогда в качестве решения вариационного неравенства

$$x \in X, \quad (c, x) \leq (c, y) \quad (\forall y \in X)$$

(несовместного, если X — открытое множество) может выступить его комитет $K = \{x_s\}$, где

$$\{x_s\} \rightarrow \arg \inf \{(c, y): y \in X\}, \quad \{x_s\} \subset X.$$

4.3. Разбор одного числового примера. Покажем на примере, как работают приведенные выше комитетные конструкции при решении практических задач распознавания образов.

Рассмотрим условную задачу управления технологическим процессом в химическом производстве. Человек-оператор наблюдает показания приборов и по этим данным принимает одно из трех решений: 1) повысить содержание добавки первого вида; 2) повысить содержание добавки второго вида; 3) не делать никакого управляющего воздействия. Задача состоит в том, чтобы на основе опыта работы оператора выработать решающее правило, ставящее в соответствие всякой производственной ситуации (отображенной приборами) вид соответствующего управляющего воздействия.

Введем обозначения:

x — вектор состояния технологического процесса, $x = [x_1, x_2]$, где x_i — показание i -го прибора,

$X_j \subset \mathbb{R}^2$ — множество всех векторов состояний, когда необходимо принять j -е решение ($j = 1, 2, 3$).

Пусть имеется материал наблюдений:

$$X = \{[3, 11], [4, 11], [-5, 10], [-3, 10], [3, 10], [-4, 9], [6, 8], [8, 9], [7, 8], [-5, 6], [-4, 5], [-5, 4], [-2, 0], [0, 0], [-1, -1], [-7, -3], [-5, -3], [-6, -4], [-12, -6], [-10, -6], [-11, -7]\}.$$

Материал обучения и контроля, т. е. элементы из X , для которых известны принадлежности образам:

$$V_1 = \{x_1^1 = [3, 11], x_2^1 = [4, 11], x_3^1 = [-7, -3], x_4^1 = [-6, -4]\} \subset X_1,$$

$$V_2 = \{x_1^2 = [-5, 10], x_2^2 = [-4, 9], x_3^2 = [6, 8], x_4^2 = [8, 9], x_5^2 = [-5, 6], x_6^2 = [-4, 5]\} \subset X_2,$$

$$V_3 = \{x_1^3 = [-2, 0], x_2^3 = [0, 0], x_3^3 = [-12, -6], x_4^3 = [-10, -6]\} \subset X_3.$$

Материал распознавания:

$$Z = \{z_1 = [-3, 10] \in X_2, z_2 = [3, 10] \in X_1, z_3 = [7, 8] \in X_2, z_4 = [-5, 4] \in X_2, z_5 = [-5, -3] \in X_1, z_6 = [-11, -7] \in X_3, z_7 = [-1, -1] \in X_3\}.$$

Принадлежность элементов из Z множествам X_i считается неизвестной для «распознавателя».

Множества V_i разбиваются на материал обучения Y_i и материал контроля W_i .

Эту задачу разбиения материала наблюдений решим следующим образом. Выберем класс функций

$$T^* = \{t^*(x) = (x_1 - a_1)^2 + (x_2 - a_2)^2 - 4: a_1 \in \mathbb{R}, a_2 \in \mathbb{R}\}.$$

Затем для выбранного V_i находим все μ -подсистемы системы

$$t^*(x) < 0 \quad (\forall x \in V_i). \tag{4.4.6}$$

Из каждой μ -подсистемы одну точку выделяем в множество W_i , а остальные составляют множество Y_i .

Система (4.4.6) для V_1 имеет вид:

$$\begin{aligned} (3 - a_1)^2 + (11 - a_2)^2 < 4, \quad \text{т. е. } t^*(x_1^1) < 0, \\ (4 - a_1)^2 + (11 - a_2)^2 < 4, \quad \text{т. е. } t^*(x_2^1) < 0, \\ (-7 - a_1)^2 + (-3 - a_2)^2 < 4, \quad \text{т. е. } t^*(x_3^1) < 0, \\ (-6 - a_1)^2 + (-4 - a_2)^2 < 4, \quad \text{т. е. } t^*(x_4^1) < 0. \end{aligned}$$

Это система неравенств относительно неизвестной функции, т. е. относительно коэффициентов a_1, a_2 .

Любое из неравенств $t^*(x_j^i) < 0$ будем помечать символом x_j^i , так что фраза «подсистема $\{x_1^1, x_2^1\}$ » означает подсистему из первых двух неравенств выписанной системы. μ -подсистемы этой системы: $\{x_1^1, x_2^1\}, \{x_3^1, x_4^1\}$. Полагаем $Y_1 = \{x_1^1, x_3^1\}, W_1 = \{x_2^1, x_4^1\}$.

Аналогично μ -подсистемы системы (4.4.6) для V_2 : $\{x_1^2, x_2^2\}, \{x_3^2, x_4^2\}, \{x_5^2, x_6^2\}$. Полагаем $Y_2 = \{x_1^2, x_3^2, x_5^2\}, W_2 = \{x_2^2, x_4^2, x_6^2\}$.

μ -подсистемы системы (4.4.6) для V_3 : $\{x_1^3, x_2^3\}, \{x_3^3, x_4^3\}$. Полагаем $Y_3 = \{x_1^3, x_3^3\}, W_3 = \{x_2^3, x_4^3\}$.

Теперь решим задачу дискриминантного анализа: найти x_1^*, x_2^* , удовлетворяющие системе

$$\left. \begin{aligned} x_1^*(x) > 0 \quad (\forall x \in Y_1), \quad x_1^*(x) < 0 \quad (\forall x \in Y_2 \cup Y_3), \\ x_2^*(x) > 0 \quad (\forall x \in Y_2), \quad x_2^*(x) < 0 \quad (\forall x \in Y_3), \end{aligned} \right\} \begin{aligned} x_1^* \in X^*, \\ x_2^* \in X^*. \end{aligned} \tag{4.4.7}$$

Найденное решение проконтролируем на множествах W_1, W_2, W_3 , а затем проведем распознавание элементов множества Z .

Пусть X^* — класс линейных функционалов, т. е. $x_i^*(x) = (x_i^*, x)$, где $x_i^* = [x_1^{i*}, x_2^{i*}]$. Система (4.4.7) имеет

ВИД:

$$\left. \begin{array}{l} 3x_1^{1*} + 11x_2^{1*} > 0, & -5x_1^{1*} + 6x_2^{1*} < 0, \\ -7x_1^{1*} + (-3)x_2^{1*} > 0, & -2x_1^{1*} < 0, \\ -5x_1^{1*} + 10x_2^{1*} < 0, & -12x_1^{1*} - 6x_2^{1*} < 0; \\ 6x_1^{1*} + 8x_2^{1*} < 0, & \end{array} \right\} (4.4.8)$$

$$\left. \begin{array}{l} -5x_1^{2*} + 10x_2^{2*} > 0, & -2x_1^{2*} < 0, \\ 6x_1^{2*} + 8x_2^{2*} > 0, & -12x_1^{2*} - 6x_2^{2*} < 0. \\ -5x_1^{2*} + 6x_2^{2*} > 0, & \end{array} \right\} (4.4.9)$$

Система (4.4.8) несовместна. Поэтому воспользуемся понятием обобщенного решения.

Пусть α_i удовлетворяет ограничению: $\alpha_i(|K^*|) \geq \beta_i |K^*|$ ($i = 1, \dots, 7$), где β_i — некоторые числа из отрезка $[0; 1]$. Если обобщенное решение существует, то существует и такое, членами которого являются решения μ -подсистем системы (4.4.8):

$$\begin{aligned} y_1^* &= [0, -1], & y_2^* &= [-2, 1], & y_3^* &= [-6, 13], \\ y_4^* &= [2, 1], & y_5^* &= [1, -1]. \end{aligned}$$

Введем векторы $\sigma_j = [\sigma_{j1}, \dots, \sigma_{j7}]$ ($j = 1, \dots, 5$), где

$$\sigma_{ji} = \begin{cases} 1, & \text{если } y_j^* \text{ удовлетворяет } i\text{-му} \\ & \text{неравенству из (4.4.8),} \\ 0 & \text{в противном случае.} \end{cases}$$

Обозначим через z_j — число экземпляров элемента y_j^* , которые должны войти в обобщенное решение. Тогда должны выполняться соотношения

$$\sum_{j=1}^5 z_j \sigma_{ji} \geq \beta_i \sum_{j=1}^5 z_j \quad (i = 1, \dots, 7), \quad (4.4.10)$$

$$z_j \text{ — целые, } z_j \geq 0 \quad (\forall j).$$

В нашем случае

$$\begin{aligned} \sigma_1 &= [0, 1, 1, 1, 1, 0, 0]; & \sigma_2 &= [1, 1, 0, 1, 0, 1, 0]; \\ \sigma_3 &= [1, 1, 0, 0, 0, 1, 1]; & \sigma_4 &= [1, 0, 1, 0, 1, 1, 1]; \\ \sigma_5 &= [0, 0, 1, 1, 1, 0, 1]. \end{aligned}$$

Система (4.4.10) примет вид:

$$\begin{aligned} z_2 + z_3 + z_4 &\geq \beta_1 (z_1 + \dots + z_5), \\ z_1 + z_2 + z_3 &\geq \beta_2 (z_1 + \dots + z_5), \\ z_1 + z_4 + z_5 &\geq \beta_3 (z_1 + \dots + z_5), \\ z_1 + z_2 + z_5 &\geq \beta_4 (z_1 + \dots + z_5), \\ z_1 + z_4 + z_5 &\geq \beta_5 (z_1 + \dots + z_5), \\ z_2 + z_3 + z_4 &\geq \beta_6 (z_1 + \dots + z_5), \\ z_3 + z_4 + z_5 &\geq \beta_7 (z_1 + \dots + z_5), \\ z_i &\geq 0, \quad z_i - \text{целые числа } (i = 1, \dots, 5). \end{aligned}$$

Пусть $\beta_1 = \beta_3 = \beta_5 = \beta_7 = 2/3$; $\beta_2 = \beta_4 = \beta_6 = 1/2$. Тогда одно из обобщенных решений соответствует следующим величинам z_i :

$$z_1 = z_2 = z_3 = z_5 = 1, \quad z_4 = 2,$$

т. е. $K = \{y_1^*, y_2^*, y_3^*, y_4^*, y_5^*\}$.

Если положить $\beta_i = 3/5$ ($\forall i$), то получаем слабый 3/5-комитет при $z_i = 1$ ($\forall i$).

Если убрать условие, что z_i — целые, получаем понятие z -решения.

Аналогично можно подойти к решению системы (4.4.9), но мы это опустим. Перейдем к этапам контроля и распознавания.

Если x — некоторый новый элемент, то

$$x \text{ относится к } \begin{cases} 1\text{-му образу, если} \\ |\{y_i^*: y_i^*(x) > 0\}| \geq \min_{(i=1,2)} \beta_i |K^*|, \\ 2\text{-му или 3-му образам, если} \\ |\{y_i^*: y_i^*(x) < 0\}| \geq \min_{(i=3, \dots, 7)} \beta_i |K^*|. \end{cases}$$

Проведем распознавание для элементов множеств W_1 , W_2 и W_3 . Отметим, что $\min_{(i=1,2)} \beta_i = 1/2$, $\min_{(i=3, \dots, 7)} \beta_i = 1/2$.

В клетках таблицы (рис. 5) знаки чисел $y_j^*(x_k^*)$.

Результаты контроля для обобщенного решения неудовлетворительны. Подвергая контролю комитет, получим ту же таблицу, но 6-й столбец (соответствующий y_4^*) выбрасывается. Кроме того, решение принимается «боль-

W_k	$x_j \backslash y_i^*$	y_1^*	y_2^*	y_3^*	y_4^*	y_4^*	y_5^*	К какому образу отнесено
W_1	x_2^1	-	+	+	+	+	-	1
W_1	x_4^1	+	+	-	-	-	-	2 или 3
W_2	x_2^2	-	+	+	+	+	-	1
W_2	x_4^2	-	-	+	+	+	-	нет решения
W_2	x_6^2	-	+	+	-	-	-	2 или 3
W_3	x_2^3	0	0	0	0	0	0	нет решения
W_3	x_4^3	+	+	-	-	-	-	2 или 3

Рис. 5.

шинством голосов». Столбец «к какому образу отнесено» выглядит так:

1
2 или 3
1
2 или 3
нет решения
2 или 3

Результаты комитетного распознавания лучше. Поэтому выбираем в качестве решающего правила: слабый (3/5)-комитет.

Результаты распознавания показаны на рис. 6.

X_i	y_i^* z_j	y_1^*	y_2^*	y_3^*	y_4^*	y_5^*	Соотношение знаков	К какому образу отнесено
X_2	z_1	-	+	+	+	-	3:2	1
X_1	z_2	-	+	+	+	-	3:2	1
X_2	z_3	-	-	+	+	-	2:3	2 или 3
X_2	z_4	-	+	+	-	-	2:3	2 или 3
X_1	z_5	+	+	+	-	-	3:2	1
X_3	z_6	+	+	-	-	-	2:3	2 или 3
X_3	z_7	+	+	-	-	0	2:2	нет решения

Рис. 6.

НЕСТАЦИОНАРНЫЕ ПРОЦЕССЫ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ, ВКЛЮЧАЮЩИЕ АЛГОРИТМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ

В практике технико-экономического планирования, по-видимому, весьма редки примеры решения сложных задач на основе процедур и методов, полностью формализованных от начала и до конца вместе со средствами связи между отдельными частями этой задачи. Как правило, возникает необходимость диалога (человек—машина, человек—эксперимент—машина и т. д.), базирующегося, в частности, на экспертизах. В основе последних лежит накопленный опыт, описываемый большими массивами плохо структуризованной информации. Это составляет естественную основу для применения методов распознавания образов, позволяющих вносить упорядочение в такие массивы в соответствии с теми или иными целями. При этом возникают комплексные процедуры принятия решений, выражающиеся в организации определенного типа нестационарных процессов. Некоторые аспекты таких процедур могут быть проанализированы чисто формально, что и составляет содержание настоящей главы.

§ 1. Синтез нестационарных процессов, включающих распознавание образов

1.1. Основные обозначения и предположения. Рассмотрим задачу математического программирования: найти элемент множества

$$\tilde{M} = \text{Arg sup } \{f_0(x) : x \in M \cap D\}, \quad (5.1.1)$$

где

$$x \in R^n, \quad M = \bigcap_{j \in I} M_j, \quad D = \bigcap_{j \in J} D_j, \\ M_j = \{x : f_j(x) \geq 0\}, \quad D_j = \{x : g_j(x) \geq 0\}.$$

Будем предполагать функции g_j известными, а функции f_j , соответствующие плохо формализуемым ограничениям, неизвестными. Однако о зависимостях f_j можно получить некоторые сведения с помощью тех или иных экспертиз.

1.2. Операторы, с помощью которых формируются алгоритмы. Введем ряд операторов.

1. *Операторы экспертизы* E_i ($i = 1, 2, 3$).

(i) Пусть $R^n = \mathfrak{X}_1 \cup \dots \cup \mathfrak{X}_k$, $\mathfrak{X}_i \cap \mathfrak{X}_j = \emptyset$ ($i \neq j$), $A \subset R^n$. Интерпретация: \mathfrak{X}_i — образы, A — материал для экспертизы. Тогда

$$\forall a \in A: [E_1(a) = j] \Leftrightarrow [a \in \mathfrak{X}_j].$$

(ii) Пусть $\varepsilon > 0$ и $\delta > 0$ фиксированы. Тогда $\forall x \in R^n$, $\forall y \in R^n$:

$$E_2(x, y) = \begin{cases} 1, & \text{если } x + \varepsilon y \in M \cap D \text{ и} \\ & f_0(x + \varepsilon y) \geq f_0(x) + \delta, \\ 0 & \text{в противном случае.} \end{cases}$$

(iii) Пусть фиксировано $\varepsilon > 0$. Тогда

$$E_3(x) = \bar{f}_j(x) \in [f_j(x) - \varepsilon f_j(x); f_j(x) + \varepsilon f_j(x)],$$

где $\text{sign } \bar{f}_j(x) = \text{sign } f_j(x)$.

2. *Экспертные отношения порядка.*

$$x >_j y \Leftrightarrow f_j(x) \geq f_j(y) \quad (\forall j \in \{0\} \cup I).$$

3. *Экспертные операторы упорядочения* χ_j . Пусть $A \subset R^n$, $A = \{a_1, \dots, a_k\}$. Тогда

$$\chi_j(A) = [a_{i_1}, \dots, a_{i_k}]: a_{i_1} >_j \dots >_j a_{i_k} \quad (\forall j \in \{0\} \cup I).$$

4. *Оператор формирования материала обучения* Φ . Пусть $R^n = \mathfrak{X}_1 \cup \dots \cup \mathfrak{X}_k$, $\mathfrak{X}_i \cap \mathfrak{X}_j = \emptyset$ ($i \neq j$), $A_i \subset \mathfrak{X}_i$ ($\forall i \in N_k$). Интерпретация: A_i — материал обучения, $A \subset R^n$ — множество новых объектов. Тогда

$$\Phi(A_1, \dots, A_k, A) = \{A'_1, \dots, A'_k\},$$

где

$$A'_i = A_i \cup \{a \in A: E_1(a) = i\}.$$

5. *Операторы дискриминации* φ и $\bar{\varphi}$. Пусть \mathcal{N} — выделенный класс подмножеств пространства R^n , $A \subset R^n$,

$B \subset \mathbb{R}^n$. Тогда

$$\begin{aligned}\tilde{\varphi}(A, B) &= \{N \in \mathcal{N}: N \supset A, N \cap B = \emptyset\}, \\ \varphi(A, B) &= \{N_t \in \mathcal{N}\},\end{aligned}$$

причем $N_t \rightarrow \tilde{\varphi}(A, B)$ при $t \rightarrow +\infty$, где сходимость может пониматься в разных смыслах, например в смысле метрики Хаусдорфа. Если

$$\begin{aligned}A &= \{a_1, \dots, a_k\}, \quad B = \{b_1, \dots, b_l\}, \\ a_1 &>_j \dots >_j a_k >_j b_1 >_j \dots >_j b_l,\end{aligned}$$

то

$$\begin{aligned}\kappa_j(A, B) &= \{P = \{x: (c, x) \geq \beta\}: (c, a_1) \geq \dots \\ &\dots \geq (c, a_k) \geq \beta \geq (c, b_1) \geq \dots \geq (c, b_l)\}.\end{aligned}$$

6. Операторы оптимизации $\tilde{\psi}_1, \psi_1, \psi_2$.

$$\tilde{\psi}_1(f_0, M, D, x_0) = \arg \sup \{f_0(x): x \in M \cap D\}$$

— этот оператор реализует отображение всех данных о задаче (вместе с начальным приближением) на некоторый элемент из ее оптимального множества. Этот элемент может быть однозначно идентифицирован либо в силу некоторого итерационного метода с заданным начальным приближением, либо охарактеризован неалгоритмически, например как наименьший по норме элемент из $\text{Arg}(5.1.1)$.

$\psi_1[f_0, M, D](x)$ — итерационный оператор для задачи (5.1.1), природа которого не фиксируется. Заметим, что ψ_1 может быть связан с $\tilde{\psi}_1$ требованием

$$\{x_{k+1} = \psi_1[f_0, M, D](x_k)\}_{\forall k} \rightarrow \tilde{\psi}_1(f_0, M, D, x_0).$$

$\psi_2[c, \varepsilon](x) = x + \varepsilon c$ — конкретный оператор, в котором $\varepsilon > 0$ и c при реализации процесса могут меняться, т. е.

$$\{x_{t+1} = \psi_2[c_t, \varepsilon_t](x_t)\}_{\forall t}.$$

Такой оператор лежит в основе, например, методов коррекции, случайного поиска, субградиентного спуска и т. д. Управление последовательностями $\{c_t\} \subset \mathbb{R}^n$ и $\{\varepsilon_t \geq 0\} \subset \mathbb{R}$ должно быть подчинено целям эффективной сходимости разрешающей последовательности.

С помощью введенных операторов нестационарный процесс математического программирования, включающий распознавание образов, может быть представлен в виде схемы, изображенной на рис. 7.

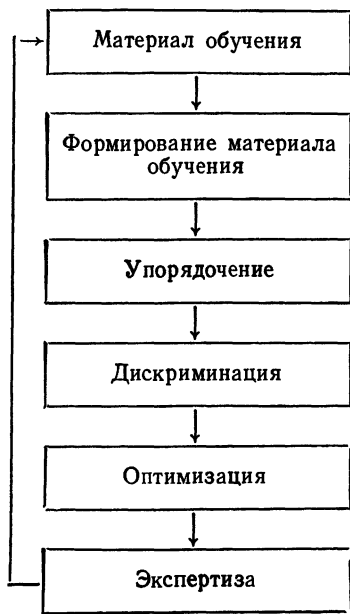


Рис. 7.

Эта схема имеет множество конкретных реализаций.

Так, например, могут быть рассмотрены различные варианты формирования материала обучения: в него могут включаться или [не включаться случайно выбираемые точки; могут указываться точки, значения функций f_j , в которых принадлежат некоторому интервалу; для зависимостей могут указываться точки надграфика и точки вне надграфика. Дискриминация может проводиться в силу различных алгоритмов дискриминантного анализа: линейной коррекции, метода фейеровского типа, метода комитетов и т. д.), с разными критериями качества разделения множеств; задача дискриминантного анализа может решаться либо до конца, либо в соответствии с заданным числом итераций.

Точно так же оптимизация может осуществляться различными алгоритмами, либо до получения решения с заданной точностью, либо с фиксированным числом шагов. Экспертизы также могут быть различными. Эксперт может указывать: оптимальный план или его окрестность; величины, производные от оптимального плана; точные зависимости f_j ; зависимости как отображения $x \rightarrow \Delta(x)$; упорядочение планов; примеры допустимых и недопустимых планов; данные для выбора вида модели и системы связей. В процессе работа эксперта может быть не задана в явной форме; в этом случае считается заданной лишь последовательность множеств, сходящаяся к допустимому множеству.

1.3. Варианты алгоритмов.

1. Алгоритм Z_0 .

Начальный шаг. Находим $x_0 \in \tilde{\Psi}_1(f_0, \mathbb{R}^n, D, x)$. Если $x_0 \in M$, то СТОП. Иначе определяем $x_j \in D \cap M_j$, $y_j \in D \setminus M_j$ ($\forall j \in I$). Полагаем $A_j^0 = \{x_j\}$, $B_j^0 = \{y_j\}$, $\{A_j^1, B_j^1\} = \Phi_j(A_j^0, B_j^0, x_0)$ ($j \in I$), где оператору Φ_j соответствует оператор экспертизы E_j^1 , действующий относительно образов $\mathfrak{A}_1^1 = M_j$ и $\mathfrak{A}_2^1 = \mathbb{R}^n \setminus M_j$.

Общий шаг. Пусть уже построены множества A_j^k, B_j^k ($\forall j \in I$). Полагаем

$$N_j^k \in \tilde{\Phi}(A_j^k, B_j^k) \quad (\forall j \in I) \quad (\text{другой вариант: } N_j^k \in \Phi(A_j^k, B_j^k)),$$

$$M_k = \bigcap_{j \in I} N_j^k,$$

$$x_k = \tilde{\Psi}_1(f_0, M^k, D, x) \quad (\text{другой вариант:}$$

$$x_k = \Psi_1^{m(k)}[f_0, M^k, D](x)),$$

$$\{A_j^{k+1}, B_j^{k+1}\} = \Phi_j(A_j^k, B_j^k, x_k) \quad (\forall j \in I).$$

Поясним работу алгоритма Z_0 . На начальном шаге находим x_0 как некоторый оптимальный вектор задачи

$$\sup \{f_0(x) : x \in D\},$$

в которой плохо формализуемое ограничение $x \in M$ не учитывается. Если $x_0 \in M$ (этот факт устанавливается экспертизой на выполнение соотношений $f_j(x_0) \geq 0$ ($\forall j \in I$)), то, очевидно, процесс решения закончен: x_0 — оптимальный вектор задачи (5.1.1). В противном случае

для всякого плохо формализуемого ограничения $f_j(\mathbf{x}) \geq 0$ находим точку $\mathbf{x}_j \in D$, удовлетворяющую ему, и точку $\mathbf{y}_j \in D$, не удовлетворяющую. Совокупность $\{\mathbf{x}_0, \mathbf{x}_j, \mathbf{y}_j\}$ разбиваем на два множества: A_j^1 включает допустимые по j -му плохо формализуемому ограничению точки, а B_j^1 — недопустимые.

Затем находим множества $N_j^1 \in \mathcal{N}^0$ такие, что $N_j^1 \supset A_j^1$, $N_j^1 \cap B_j^1 = \emptyset$, и полагаем $M^1 = \bigcap_{j \in I} N_j^1$. Таким образом, со-

отношение $\mathbf{x} \in M^1$ выступает в качестве первого приближения к ограничению $\mathbf{x} \in M$ (в дальнейшем предполагается, что $M^1 \neq \emptyset$). Далее находим \mathbf{x}_1 как один из оптимальных векторов задачи $\sup \{f_0(\mathbf{x}) : \mathbf{x} \in D \cap M^1\}$, подвергая \mathbf{x}_1 экспертизе на предмет его допустимости по каждому плохо формализуемому ограничению, пополняем множества A_j^1 и B_j^1 до множеств A_j^2 и B_j^2 ($\forall j \in I$); находим $N_j^2 \in \mathcal{N}^0$: $N_j^2 \supset A_j^2$, $N_j^2 \cap B_j^2 = \emptyset$; полагаем $M^2 = \bigcap_{j \in I} N_j^2$; затем вся процедура повторяется для M^2 вместо M^1 и т. д. Предполагается, что $M^k \neq \emptyset$ ($\forall k$).

Приводимый далее алгоритм Z_1 отличается от алгоритма Z_0 только тем, что эксперт дополнительно упорядочивает точки материала обучения по значениям каждой функции f_j , а множества A_j^k и B_j^k разделяются линейной функцией (c, \mathbf{x}) такой, что если $\mathbf{x}, \mathbf{y} \in A_j^k \cup B_j^k$ и $f_j(\mathbf{x}) \geq f_j(\mathbf{y})$, то $(c, \mathbf{x}) \geq (c, \mathbf{y})$. Такая дополнительная информация эксперта улучшает (хотя и не во всех случаях) свойства последовательности $\{\mathbf{x}_k\}$ по отношению к цели получения приближения к оптимальному вектору задачи (5.1.1).

2. Алгоритм Z_1 . Начальный шаг как в алгоритме Z_0 .

Общий шаг. Пусть уже построены множества A_j^k, B_j^k ($\forall j \in I$). Полагаем

$$P_j^k = \chi_j(\chi_j(A_j^k), \chi_j(B_j^k)) \cap \bar{\varphi}(A_j^k, B_j^k),$$

$$M_k = \bigcap_{j \in I} P_j^k,$$

$\mathbf{x}_k = \psi_1(f_0, M, D, \mathbf{x})$ (другой вариант:

$$\mathbf{x}_k = \psi_1^{m(k)}[f_0, M, D](\mathbf{x}),$$

$$\{A_j^{k+1}, B_j^{k+1}\} = \Phi_j(A_j^k, B_j^k, \mathbf{x}_k).$$

3. Алгоритм Z_2 . Предполагается, что критерий f_0 плохо формализуем, а множество M задано.

Общий шаг. Пусть уже построено множество $A^k \subset D \cap M$. Полагаем

$$y_k \in x_0(\chi_0(A^k)), \alpha_{k+1} = \tilde{\psi}_1((y_k, \cdot), M, D, x)$$

(другой вариант: ψ_1 вместо $\tilde{\psi}_1$).

4. Алгоритм Z_3 . Этот алгоритм относится к ситуации, когда плохо формализуемы как часть ограничений, задающих множество M (можно говорить о плохой формализуемости включения $x \in M$), так и целевой функции $f_0(x)$. В некоторых случаях оказывается достаточным промоделировать зависимость $f_0(x)$ функцией вида

$$F(x) \in \begin{cases} \Delta_1 \subset \mathbb{R}, & \text{если } x \in X_1, \\ \Delta_p \subset \mathbb{R}, & \text{если } x \in X_p, \end{cases}$$

где $\{X_1, \dots, X_p\}$ — некоторое разбиение пространства \mathbb{R}^n .

В алгоритме Z_3 мы примем: $f_0(x) = F(x)$ при $\Delta_i = i$, $X_1 = \{x: (a, x) \leq 1\}$, $X_i = \{x: i-1 < (a, x) \leq i\}$ ($i = 2, \dots, p-1$), $X_p = \{x: (a, x) > p-1\}$.

Эти предположения оформляют один из наиболее естественных способов приближенного моделирования цели.

Предполагаем, что $f_0(x) = p$, $x \in M \cap D$ совместна.

Общий шаг. Пусть уже построены множества

$$A_j^k \subset M_j, B_j^k \subset \mathbb{R}^n \setminus M_j \quad (\forall j \in I), \quad S_j^k \subset D: \\ f_0(x) = j \quad (\forall x \in S_j^k).$$

Полагаем

$$N_j^k \in \tilde{\varphi}(A_j^k, B_j^k) \quad (j \in I) \quad (\text{другой вариант: } \varphi \text{ вместо } \tilde{\varphi}),$$

$$M_k = \bigcap_{j \in I} N_j^k, \quad Q_k \in \tilde{\varphi}\left(S_p^k, \bigcup_{i=1}^{p-1} S_i^k\right) \quad (\text{другой вариант: } \varphi$$

вместо $\tilde{\varphi}$),

$$x_k \in Q_k \cap M_k \cap D,$$

$$\{A_j^{k+1}, B_j^{k+1}\} = \Phi_j(A_j^k, B_j^k, x_k) \quad (\forall j \in I),$$

$\{S_1^{k+1}, \dots, S_p^{k+1}\} = \Phi(S_1^k, \dots, S_p^k, x_k)$, где оператор Φ отвечает образам $\{x: f_0(x) = j\}$ ($j \in N_p$).

5. Алгоритм Z_4 (моделирование графиков зависимостей).

Общий шаг. Пусть построены множества

$$A_j^k \subset \{[x, y] \in \mathbb{R}^{n+1}: x \in \mathbb{R}^n, y > f_j(x)\},$$

$$B_j^k \subset \{[x, y] \in \mathbb{R}^{n+1}: x \in \mathbb{R}^n, y < f_j(x)\} \quad (j \in \{0\} \cup I).$$

Полагаем:

$N_j^k \in \tilde{\varphi}(A_j^k, B_j^k)$ (другой вариант: φ вместо $\tilde{\varphi}$);

$M_j^k = \{x \in \mathbb{R}^n: \exists y \leq 0, [x; y] \in N_j^k\} (\forall j \in I)$;

$f_0^k(x) = \inf \{y: [x, y] \in N_0^k\}$;

$M_k = \bigcap_{i \in I} M_i^k$;

$x_{k+1} = \tilde{\psi}_1(f_0^k, M_k, D, x)$ (другой вариант: ψ_1 вместо $\tilde{\psi}_1$);

$C^k \subset \{[x_{k+1}, y]: y \in \mathbb{R}\}$;

$\{A_j^{k+1}, B_j^{k+1}\} = \Phi_j(A_j^k, B_j^k, C^k)$.

6. Алгоритм Z_5 (случайное генерирование планов).

Общий шаг. Пусть уже построено множество $A^k = \{[a_t, z_t]\}_{t=1}^k \subset \mathbb{R}^n \times S$, $S = \{z: |z| = 1\}$, $A^k = B_1^k \cup B_2^k$, причем

$$E_2(a, z) = 1 \quad (\forall [a, z] \in B_1^k),$$

$$E_2(a, z) = 0 \quad (\forall [a, z] \in B_2^k).$$

Полагаем

$N^k \in \tilde{\varphi}(B_1^k, B_2^k)$ (другой вариант: φ вместо $\tilde{\varphi}$);

$a_{k+1} = a_k + \varepsilon E_2(a_k, z_k) z_k$;

$z_{k+1} \in \{z: [a_{k+1}, z] \in N^k, |z| = 1\}$;

$\{B_1^{k+1}, B_2^{k+1}\} = \Phi(B_1^k, B_2^k, [a_{k+1}, z_{k+1}])$.

7. Алгоритм Z_6 (замена экспертизы таксономией).

Пусть в задаче (5.1.1) \tilde{f}_0 и D заданы, M неизвестно, $M \cap D \neq \emptyset$, $\sup \{|x|: x \in M\} < +\infty$. Пусть задано $A \subset \mathbb{R}^n$, $|A| < +\infty$, причем $\forall a \in A: |a - M| = \inf \{|a - x|: x \in M\} \leq \varepsilon$, где $\varepsilon > 0$ — достаточно малое число.

Проводим таксономию множества A ; пусть $G = \{g_1, \dots, g_r\}$ — совокупность центров тяжести таксонов. Находим множество G' крайних точек (или вершин, что одно и то же) выпуклой оболочки множества G . Определяем «разделяющее» множество

$$N \in \tilde{\varphi}(G \setminus G', G'),$$

а затем и

$$\tilde{x} = \tilde{\psi}_1(f_0, N, D, x_0).$$

Элемент \tilde{x} является оптимальным вектором для задачи

$$\sup \{f_0(x): x \in N \cap D\},$$

где N моделирует множество допустимых по плохо формализуемому ограничению решений.

1.4. Варианты операторов дискриминации и экспертизы.

1. Варианты операторов дискриминации:

оператор линейной дискриминации LD (реализующие его алгоритмы: линейной коррекции LC , фейеровского типа F , симплекс-метода SM);

оператор аффинной дискриминации AD (те же алгоритмы; кроме того, алгоритм $\mathfrak{N}1$, дающий гиперплоскость, делящую пополам разделяющий плоский слой наибольшей толщины, и алгоритм $\mathfrak{N}2$, отличающийся от $\mathfrak{N}1$ тем, что нормаль слоя фиксирована);

оператор построения разделяющего комитета функционалов CD ;

оператор DR , отвечающий экспертизам E_4^j .

2. Варианты операторов оптимизации:

симплекс-метод SM ;

фейеровского типа F ;

случайный поиск RS .

В следующем параграфе исследуются алгоритмы $Z_0 - Z_6$ при различных вариантах вхождений в них перечисленных операторов.

§ 2. Учет плохо определенных ограничений методами дискриминантного анализа и таксономии

2.1. Алгоритм Z_0 при условии $\mathfrak{N}1$. Через $\mathfrak{N}1$ обозначается выбор нормального полупространства $P \in \tilde{\varphi}(A, B)$, т. е. такого, что граничная гиперплоскость ∂P делит пополам плоский разделяющий слой наибольшей толщины.

Рассмотрим следующую задачу ЛП: найти

$$\bar{x} = \arg \max \{(c, x) : x \in M \cap D\}, \quad (5.2.1)$$

где

$$D = \{x : (a_j, x) - b_j \leq 0 \ (\forall j \in N_m)\},$$

$$M = \{x : (c_1, x) - t_1 \leq 0\}.$$

В этой ситуации предполагается наличие только одного плохо формализуемого ограничения. Рассматривать ее имеет смысл потому, что часто в практических задачах множество D отражает почти все ограничения, а число плохо формализуемых ограничений достаточно мало. Так, в задачах оптимального использования ресурсной базы производства может присутствовать лишь небольшая

группа плохо формализуемых ограничений, отражающих физические свойства природных веществ.

Теорема 5.2.1. Пусть D — выпуклое многогранное телесное множество, \mathbf{x}_0 — единственное оптимальное решение задачи $\text{тах} \{(\mathbf{c}, \mathbf{x}): \mathbf{x} \in D\}$, $\bar{\mathbf{x}}$ — единственное оптимальное решение задачи $\text{тах} \{(\mathbf{c}, \mathbf{x}): \mathbf{x} \in M \cap D\}$, M — замкнутое полупространство, l — ребро многогранника D , причем $\mathbf{x}_0 \in l$, $\mathbf{x}_1 \in l$, $\bar{\mathbf{x}} \in l$. Пусть в алгоритме Z_0 при варианте $\mathfrak{N}1$: $\mathbf{x}_k \in l \cap \partial N^k$ ($\forall k$). Тогда $\{\mathbf{x}_k\} \rightarrow \bar{\mathbf{x}}$.

Доказательство. Гиперплоскость ∂N_1^1 делит отрезок $\overline{\mathbf{x}_0, \mathbf{x}_1}$ пополам и ортогональна ему;

$$\{\mathbf{x}_2\} = \partial N_1^1 \cap \overline{\mathbf{x}_0, \mathbf{x}_1}.$$

Если $\mathbf{x}_2 \in M$, то ∂N_1^1 делит пополам отрезок $\overline{\mathbf{x}_1, \mathbf{x}_2}$, причем $\bar{\mathbf{x}} \in \overline{\mathbf{x}_0, \mathbf{x}_2}$.

Вообще, на каждом шаге гиперплоскость делит пополам тот отрезок, которому принадлежит $\bar{\mathbf{x}}$. Теорема доказана.

Замечание. Условие $\mathfrak{N}1$ в теореме 5.2.1 можно ослабить, требуя, чтобы плоскость ∂N_1^k делила плоский слой не пополам, а в отношении $\lambda_k/(1-\lambda_k)$, $\inf_{(k)} \lambda_k > 0$,

$$\sup_{(k)} \lambda_k < 1.$$

Если $A, B \subset \mathbb{R}^n$, то полагаем

$$\tilde{\rho}(A, B) = \inf \{|\mathbf{x} - \mathbf{y}|: \mathbf{x} \in A, \mathbf{y} \in B\}.$$

Теорема 5.2.2. Пусть в задаче (5.1.1) $\{f_j\}$ — аффинные функции относительно \mathbf{x} , D — ограниченное замкнутое выпуклое множество, $M \cap D \neq \emptyset$, и применяется алгоритм Z_0 с условием $\mathfrak{N}1$. Предположим, что \mathcal{N} — класс всех замкнутых полупространств, $J(\mathbf{x}_k) = \{j: \mathbf{x}_k \in \partial N_j^k\}$, $J' = \{j: j \in J(\mathbf{x}_k) \text{ для бесконечной последовательности элементов } \mathbf{x}_k\}$. Тогда

$$\lim_{(k)} \tilde{\rho}(\text{co } A_j^k, \text{co } B_j^k) = 0, \quad \forall j \in J'.$$

Доказательство. Зафиксируем $j \in J'$ и обозначим: $\rho_k = \tilde{\rho}(\text{co } A_j^k, \text{co } B_j^k)$. Так как $A_j^{k+1} \supset A_j^k$, $B_j^{k+1} \supset B_j^k$, то $\rho_{k+1} \leq \rho_k$ ($\forall k$). Поэтому $\exists \rho \geq 0$: $\rho_k \rightarrow \rho$ ($k \rightarrow +\infty$). Покажем, что неравенство $\rho > 0$ невозможно.

Действительно, предположим противное, т. е. $\rho > 0$. Тогда существует $\theta \in (0; 1)$ такое, что если $f_j(\bar{x}_{k+1}) = 0$, то $\rho_{k+1} \leq \theta \rho_k$. В самом деле, пусть $\rho_k = |\bar{x} - \bar{y}|$, $\bar{x} \in \text{co } A_j^k$, $\bar{y} \in \text{co } B_j^k$. Не ограничивая общности, предположим, что \bar{x}_{k+1} лежит по ту же сторону от гиперплоскости

$$\{x: f_j(x) = 0\},$$

что и \bar{x} . Тогда $\rho_{k+1} \leq \tilde{\rho}(\bar{y}, \overline{\bar{x}, x_{k+1}})$. Так как

$$|\bar{x} - \bar{y}| \geq \rho, \quad |\bar{y} - x_{k+1}| = |\bar{x} - x_{k+1}| \leq R,$$

где R — радиус шара, включающего D , то $\exists \theta \in (0; 1)$:

$$[z = (\bar{x} + \bar{y})/2] \Rightarrow |x_{k+1} - z| / |x_{k+1} - \bar{x}| \leq \theta.$$

Но тогда и $\rho_k(\bar{y}, \overline{\bar{x}, x_{k+1}}) \leq \theta |\bar{x} - \bar{y}|$.

Таким образом, $\rho_{k+1} \leq \theta \rho_k (\forall k)$. Однако это противоречит тому, что $\rho_k \rightarrow \rho > 0$. Теорема доказана.

Теорема 5.2.3. Пусть в задаче (5.1.1) D — многогранное ограниченное замкнутое выпуклое множество, $\{f_j\}$ — аффинные функции, $|I| = 1$ и применяется алгоритм Z_0 с условием $\mathfrak{N}1$, \mathfrak{N} — класс всех полупространств. Предположим, что $\forall k: x_k$ — базисное решение задачи $\max \{f_0(x): x \in D, x \in N_1^k\}$. Если $A_1^0 \neq \emptyset$, $B_1^0 \neq \emptyset$, то либо $\{x_k\} \rightarrow \bar{x} = \arg \max \{f_0(x): x \in M \cap D\}$, либо

$$\lim_{(k)} \tilde{\rho}(\text{co } A_1^k, \text{co } B_1^k) = 0. \quad (5.2.2)$$

Доказательство. Имеются две возможности.

1. $\forall k, \exists k' \geq k: x_{k'} \in \partial N_1^{k'}$. Тогда по теореме 5.3.2 соотношение (5.2.2) выполняется.

2. $\exists \bar{k}, \forall k' \geq \bar{k}: x_{k'} \notin \partial N_1^{k'}$. Если при этом $\exists k' \geq \bar{k}: x_{k'} \in M$, то, очевидно, $x_{k'} = \bar{x}$. Поэтому предположим, что $\forall k' \geq \bar{k}: x_{k'} \notin M$. Тогда $\forall k' \geq \bar{k}$ все точки последовательности $\{x_k\}_{\bar{k}}^{k'}$ недопустимы и находятся по одну сторону от $\partial N_1^{k'+1}$, а $x_{k'}$ — оптимальное решение задачи

$$\max \{f_0(x): x \in D \cap N_1^{k'}\}.$$

Так как $x_{k'} \notin \partial N_1^{k'}$, то ограничение $x \in N_1^{k'}$ является несущественным, т. е.

$$x_{k'} = \arg \max \{f_0(x): x \in D\}.$$

Так как $\mathbf{x}_{k'+1} \notin \partial N_1^{k'+1}$, то точно так же $\mathbf{x}_{k'+1} = \arg \max \{f_0(\mathbf{x}) : \mathbf{x} \in D\}$. Но соотношение $\mathbf{x}_{k'} = \mathbf{x}_{k'+1}$ невозможно, поскольку $N_1^{k'+1}$ не включает недопустимую точку $\mathbf{x}_{k'}$. Теорема доказана.

Теорема 5.2.4. При условиях случая 1 теоремы 5.2.3 справедливы соотношения

$$\begin{aligned} & \overline{\left(\text{co } \bigcup_{k=1}^{\infty} A_1^k \right)} \cap \overline{\left(\text{co } \bigcup_{k=1}^{\infty} B_1^k \right)} \cap \partial M \neq \emptyset; \\ & \overline{\bigcup_{k=1}^{\infty} A_1^k} \cap \partial M \neq \emptyset; \quad \overline{\bigcup_{k=1}^{\infty} B_1^k} \cap \partial M \neq \emptyset. \end{aligned}$$

Здесь черта над множествами означает их топологическое замыкание.

Доказательство. Перейдя в соотношении

$$0 \leq \tilde{\rho} \left(\text{co } \bigcup_{k=1}^{\infty} A_1^k, \text{co } \bigcup_{k=1}^{\infty} B_1^k \right) \leq \tilde{\rho}(\text{co } A_1^k, \text{co } B_1^k) \quad (\forall k)$$

к пределу при $k \rightarrow +\infty$, получим

$$\tilde{\rho} \left(\text{co } \bigcup_{k=1}^{\infty} A_1^k, \text{co } \bigcup_{k=1}^{\infty} B_1^k \right) = 0.$$

Отсюда следует:

$$\begin{aligned} & \overline{\text{co } \bigcup_1^{\infty} A_1^k} \cap \overline{\text{co } \bigcup_1^{\infty} B_1^k} \cap \partial M \neq \emptyset; \\ & \overline{\text{co } \bigcup_1^{\infty} A_1^k} \cap \overline{\text{co } \bigcup_1^{\infty} B_1^k} \subset \partial M. \end{aligned}$$

Но если некоторое замкнутое множество U лежит по одну сторону от гиперплоскости V , причем $\text{co } U \cap V \neq \emptyset$, то $U \cap V \neq \emptyset$. Поэтому

$$\begin{aligned} & \overline{\bigcup_1^{\infty} A_1^k} \cap \partial M \neq \emptyset, \\ & \overline{\bigcup_1^{\infty} B_1^k} \cap \partial M \neq \emptyset. \end{aligned}$$

Теорема доказана.

Приведем пример, показывающий, что даже в случае ограниченного D последовательность $\{\mathbf{x}_k\}$ не обязана сходиться к оптимальному решению.

Пример 1. Пусть $\mathbf{x} = [x_1, x_2] \in \mathbf{R}^2$, $\mathbf{c} = [0, 1] \in \mathbf{R}^2$,

$$D = \{\mathbf{x}: x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \leq 1\};$$

$$M = \{\mathbf{x}: x_1 + 3x_2 \leq 2\}; A_1^0 = \{[1, 0]\}, B_1^0 = \{[0, 1]\}.$$

Тогда $\bar{\mathbf{x}} = [0, 2/3]$; $\mathbf{x}' = \lim_{(k)} \mathbf{x}_k = [1/2, 1/2] \in \partial M \cap \partial D$,

$$\mathbf{x}' \neq \bar{\mathbf{x}}.$$

Этот пример показывает, что в некоторых случаях алгоритм Z_0 в варианте $\mathfrak{N}1$ целесообразно дополнить каким-либо механизмом (случайным или детерминированным) генерирования дополнительных точек $\mathbf{x} \in D$, так как иначе процесс может не приводить к нужному результату.

Определение 5.2.1. Пусть

$$M \cap D = \{\mathbf{x}: (c_j, \mathbf{x}) \leq b_j (\forall j \in J)\}, \bar{M} = \text{Arg} (5.2.1),$$

$$J(\bar{M}) = \{j: \exists \bar{\mathbf{x}} \in \bar{M}: (c_j, \bar{\mathbf{x}}) = b_j\}, H_j = \{\mathbf{x}: (c_j, \mathbf{x}) = b_j\}.$$

Квазиоптимальным решением задачи (5.2.1) назовем всякий элемент

$$\mathbf{x} \in M \cap D \cap \left(\bigcup_{j \in J(\bar{M})} H_j \right).$$

Определение 5.2.2. Полупространство $P \in \bar{\varphi}(A, B)$ называется *нормальным*, если его граница ∂P делит пополам разделяющий A и B плоский слой наибольшей толщины.

Теорема 5.2.5. Пусть выполнены условия теоремы 5.2.3 и применяется алгоритм Z_0 . Если N_1^k — нормальный элемент из $\bar{\varphi}(A_1^k, B_1^k)$, $\bar{M} \neq \emptyset$, то существует квазиоптимальное решение \mathbf{x}' такое, что $\{\mathbf{x}_k\} \rightarrow \mathbf{x}'$, где $\{\mathbf{x}_k\}$ вырабатывается в силу алгоритма Z_0 .

Доказательство следует из теоремы 5.2.3.

2.2. Алгоритм Z_0 при условии $\mathfrak{N}2$. Условие $\mathfrak{N}2$ следующим образом выделяет класс \mathcal{N} подмножеств, определяющий оператор дискриминации $\bar{\varphi}$. Фиксируя вектор $\mathbf{d} \in \mathbf{R}^n$, определим класс $\mathcal{N} = \{N = \{\mathbf{x}: (\mathbf{d}, \mathbf{x}) \leq \alpha\}: \alpha \in \mathbf{R}\}$. Положим $\bar{\varphi}(A, B) = \{N \in \mathcal{N}: \partial N \text{ делит пополам плоский слой с нормалью } \mathbf{d}, \text{ разделяющий } A \text{ и } B, \text{ наибольшей толщины}\}$.

Теорема 5.2.6. Пусть для задачи (5.1.1) $M \cap D \neq \emptyset$; множества $A_j^k \neq \emptyset$, $B_j^k \neq \emptyset$ ($\forall j \in I, \forall k$) делимы слоем

с нормалью \mathbf{d} . Если применяется алгоритм Z_0 с условием $\mathfrak{N}2$, то $\forall j$ существует гиперплоскость H_j такая, что

$$|\partial N_j^k - H_j|_+ \rightarrow 0 \quad (k \rightarrow +\infty).$$

Доказательство. На каждом шаге плоскость ∂N_j^k делит пополам предыдущий разделяющий плоский слой. Если μ_k — толщина слоя на k -м шаге, то $\mu_{k+1} = \mu_k/2$ ($\forall k$). Отсюда $\{\mu_k\} \rightarrow 0$, что и требовалось.

Далее предполагается: M и D — выпуклые многогранные множества, при этом D ограничено.

Приведем пример, показывающий, что даже при этих предположениях для процедуры Z_0 при условии $\mathfrak{N}2$ нельзя утверждать, что $\tilde{\rho}$ (со A_1^k , со B_1^k) $\rightarrow 0$ ($k \rightarrow +\infty$).

Пример 2. $D = \{\mathbf{x} \in \mathbb{R}^2: \mathbf{h} \geq \mathbf{x} \geq 0\}$, $\mathbf{h} = [10, 10]$, $M = \{\mathbf{x}: x_1 + x_2 \geq 2\}$, $A_1^0 = \{[3 + \varepsilon, 0]\}$, $B_1^0 = \{[0, 1]\}$, $\mathbf{d} = [1, 3]$, $\varepsilon > 0$, $\mathbf{c} = [-1, 1]$. Вычисления показывают, что

$$\tilde{\rho}(\text{со } A_1^k, \text{ со } B_1^k) = (10 + 6\varepsilon + \varepsilon^2)^{1/2}.$$

Теорема 5.2.7. Пусть для задачи (5.1.1) $M \cap D \neq \emptyset$, $|I| = 1$, A_1^k и B_1^k не пусты и делимы некоторыми плоскими слоями с нормалью \mathbf{d} . Если применяется алгоритм Z_0 с условием $\mathfrak{N}2$, то либо $\exists k: \mathbf{x}_k = \arg(5.1.1)$, либо $\{\mu_k\} \rightarrow 0$, где μ_k — толщина разделяющего A_1^k и B_1^k слоя.

Доказательство. На k -м шаге имеются три возможности.

1. $\mathbf{x}_k \in \partial N_1^k$. Тогда $\mu_{k+1} = \mu_k/2$.

2. $\mathbf{x}_k \notin \partial N_1^k$, \mathbf{x}_k — допустимая точка. Тогда $\mathbf{x}_k = \arg \sup \{f_0(\mathbf{x}): \mathbf{x} \in D\} = \arg(5.1.1)$.

3. $\mathbf{x}_k \notin \partial N_1^k$, \mathbf{x}_k — недопустимая точка. Тогда $\mu_{k+1} < \mu_k/2$. Следовательно, либо $\exists k: \mathbf{x}_k = \arg(5.1.1)$, либо $\{\mu_k\} \rightarrow 0$. Теорема доказана.

Теорема 5.2.8. Пусть в задаче (5.2.1) ограничение $\mathbf{x} \in M$ существенно, $D \cap M \neq \emptyset$, A_1^k и B_1^k не пусты и делимы плоским слоем с нормалью \mathbf{d} , $\{\bar{\mathbf{x}}\} = \text{Arg}(5.2.1)$, число $\varepsilon = |\mathbf{c}_1|/c_1 - \mathbf{d}/|\mathbf{d}|$ достаточно мало, $(\mathbf{c}, \mathbf{x}) \geq (\mathbf{c}, \bar{\mathbf{x}})$ ($\forall \mathbf{x} \in B_1^k$, $\forall k$). Если применяется процедура Z_0 с условием $\mathfrak{N}2$, то $\mathbf{x}_k \rightarrow \bar{\mathbf{x}}$ ($k \rightarrow +\infty$).

Доказательство. По теореме 5.2.7 либо $\exists k: \mathbf{x}_k = \bar{\mathbf{x}}$, либо $\exists N: |\partial N_1^k - H|_+ \rightarrow 0$. Пусть ε так мало, что огра-

нение $x \in N_1^k$ существенно. Тогда и ограничение $x \in N$, где N — полупространство с границей H , существенно. Пусть $\arg \max \{(c, x) : x \in D \cap N\} = x'$. Тогда $x' \in H$, $x' \in \partial D$. Кроме того, $x' \in \partial M$, так как по определению разделяющего слоя, в любой окрестности множества $H \cap \partial D$ имеются как точки из $D \cap M$, так и точки из $D \setminus M$. Итак, $\{x'\} = \partial M \cap H \cap \partial D = \bar{\{x\}}$, что и требовалось доказать.

2.3. Алгоритм Z_0 с линейной коррекцией. Рассмотрим применительно к задаче (5.2.1) алгоритм Z_0 , в котором $A_1^k = A \neq \emptyset$, $B_1^k = B \neq \emptyset$ ($\forall k$), φ — итерационный оператор, соответствующий методу линейной коррекции для разделения множеств A и B (см. § 2 гл. IV). В качестве оператора оптимизации для алгоритма Z_0 применим

$$\psi_s [D, N, c](x) = \pi [D, N](x) + \varepsilon c,$$

где π — оператор проектирования на наиболее удаленное из полупространств $\{x : (a_j, x) - b_j \leq 0\}$ ($\forall j \in N_m$) и N таких, что их пересечение не пусто (пересечение первых m полупространств есть D). Этим самым для задачи (5.2.1) определен алгоритм

$$\begin{aligned} N_1^k &= \varphi^k(A, B), \\ x_{k+1} &= \psi_s [D, N_1^k, c](x_k). \end{aligned} \quad (5.2.3)$$

Первое из выписанных соотношений соответствует компоненте алгоритма, реализующей дискриминацию множеств A и B с целью моделирования плохо формализуемого ограничения $(c_1, x) - t_1 \leq 0$, а второе — одному из вариантов метода фейеровских приближений, используемого собственно для оптимизации.

Теорема 5.2.9. Пусть множества A и B конечны и аффинно разделимы, задача (5.2.1) разрешима. Тогда для последовательности $\{x_k\}$, возникающей в силу (5.2.3), выполняется свойство

$$\forall \delta > 0, \exists \varepsilon > 0, \exists k_0 \in N : [k \geq k_0] \Rightarrow [|\bar{f} - (c, x_k)| \leq \delta],$$

где \bar{f} — оптимальное значение задачи

$$\max \{(c, x) : x \in D \cap P\},$$

в которой P — некоторое полупространство со свойствами

$$P \supset A, P \cap B = \emptyset.$$

Доказательство компилирует основную теорему из работы [34] и тот факт, что метод линейной коррекции решает в принятых предположениях задачу разделения множеств A и B за конечное число шагов (это вытекает из теоремы 4.2.1).

2.4. Алгоритм Z_1 . Обратимся вновь к примеру 1. В случае алгоритма Z_0 мы получили: $\bar{x} = [0, 2/3]$, $x' = \lim_{(k)} x_k = [1/2, 1/2]$, $x' \neq \bar{x}$. Нетрудно убедиться, что этот же результат получается, если воспользоваться алгоритмом Z_1 . Это объясняется тем, что на грани

$$\{x: x_1 + x_2 = 1, x_1 \geq 0, x_2 \geq 0\}$$

предпочтение, устанавливаемое функцией $x_1 + 3x_2$, совпадает с предпочтением, отвечающим функции $-x_1 + x_2$, которая определяет разделяющий плоский слой наибольшей толщины.

Действительно, если $x_1 + x_2 = 1$, $y_1 + y_2 = 1$, $x_1 + 3x_2 \geq y_1 + 3y_2$, то $-x_1 + x_2 \geq -y_1 + y_2$.

Однако если множества A_1^0 и B_1^0 содержат достаточное «разнообразие» элементов, то алгоритм Z_1 может существенно раньше, чем Z_0 , «выправить» разделяющую гиперплоскость. Пусть в примере 1 $B_1^0 = \{x_1 = [0, 1]\}$, $A_1^0 = \{y_1 = [1/2, 1/2], y_2 = [1, 0], y_3 = [1/2, 0]\}$. Слой наибольшей толщины, разделяющий A_1^0 и B_1^0 , не удовлетворяет предпочтениям. Поэтому поворачиваем соответствующую ему нормаль, пока предпочтения не станут удовлетворяться. Получаем нормаль $[0, 1]$. Тогда имеет место сходимость к оптимальному решению. В случае же алгоритма Z_0 при этих множествах A_1^0 и B_1^0 снова получим $x_k \rightarrow [1/2, 1/2]$ при $k \rightarrow +\infty$.

2.5. Условия сходимости за один шаг в алгоритме Z_2 . Будем предполагать множество $D \cap M$ ограниченным и замкнутым, а решение \bar{x} единственным: $\bar{M} = \{\bar{x}\}$. Отсюда следует, что \bar{x} — вершина многогранного множества $D \cap M$. Не ограничивая общности, предположим, что первые n неравенств системы, задающей множество $D \cap M$, а именно:

$$(c_j, x) - b_j \leq 0 \quad (j = 1, \dots, n),$$

обладают следующим свойством. Векторы $\{c_j: \forall j \in N_n\}$ линейно независимы, а в точке \bar{x} все эти неравенства обращаются в равенства.

Определим множество

$$G(\bar{\mathbf{x}}) = \{g \in \mathbb{R}^n: \bar{\mathbf{x}} = \arg \max \{(g, \mathbf{x}): (c_j, \mathbf{x}) - b_j \leq 0, \forall j \in N_n\}\}.$$

Согласно теореме Фаркаша—Минковского [111], оно исчерпывается элементами вида

$$g = \sum_{j=1}^n \lambda_j c_j, \quad \lambda_j \geq 0 \quad (\forall j \in N_n).$$

Если C —множество всех решений системы

$$(a'_1, \mathbf{y}) < 0, \dots, (a'_{k-1}, \mathbf{y}) < 0, \quad a'_i = a_i - a_{i-1}, \quad (5.2.4)$$

причем $C \subset G(\bar{\mathbf{x}})$, то для всякого $\mathbf{y} \in C$

$$\bar{\mathbf{x}} = \arg \max \{(\mathbf{y}, \mathbf{x}): \mathbf{x} \in M \cap D\}.$$

В этом случае за один шаг мы приходим к решению, т. е. $a_{k+1} = \bar{\mathbf{x}}$.

2.6. Ситуации заикливания в алгоритме Z_2 . Пусть, применяя алгоритм Z_2 , мы на k -м шаге получили упорядоченное множество векторов $\{a_1, \dots, a_k\}$, затем определили $a_{k+1} = \arg \max \{(\bar{\mathbf{y}}, \mathbf{x}): \mathbf{x} \in D \cap M\}$, где $\bar{\mathbf{y}}$ —некоторое решение системы (5.2.4). Если для $\lambda_1 \geq 0, \dots, \dots, \lambda_{k-1} \geq 0$ выполняется соотношение $a'_k = \sum_{i=1}^{k-1} \lambda_i a'_i$, где $a'_i = a_i - a_{i-1}$, то по теореме Фаркаша—Минковского неравенство

$$(a'_k, \mathbf{y}) \leq 0 \quad (5.2.5)$$

является следствием системы (5.2.4), и тогда на этом шаге не произойдет сужения конуса C решений этой системы. Если же неравенство (5.2.5) не является следствием, то происходит сужение конуса C .

Предположим, что \mathbf{v} и $\bar{\mathbf{v}}$ —две некоторые вершины множества $D \cap M$, причем $C \cap G(\mathbf{v}) \neq \emptyset$, $C \cap G(\bar{\mathbf{v}}) \neq \emptyset$, а также

$$\bar{\mathbf{v}} = \arg \max \{(c, \mathbf{x}): \mathbf{x} \in D \cap M\}.$$

Тогда если в качестве решения системы (5.2.4) выбрать вектор $\bar{\mathbf{y}} \in G(\bar{\mathbf{v}})$, то получим

$$\mathbf{v} = \arg \max \{(\bar{\mathbf{y}}, \mathbf{x}): \mathbf{x} \in D \cap M\}.$$

Если неравенства, полученные добавлением вектора \mathbf{a}_{k+1} к множеству A , будут следствиями системы (5.2.4), то сужения конуса C не будет, и вектор $\bar{\mathbf{y}}$ можно вновь взять в качестве решения полученной системы. В этом случае происходит заикливание: $\mathbf{a}_{k+1} = \mathbf{a}_{k+2} = \dots = \mathbf{v}$. Следовательно, сходимости к решению не будет.

Приведем конкретный пример заикливания в \mathbf{R}^2 .

Пример 3. Рассмотрим задачу $\max \{(c, \mathbf{x}) : \mathbf{x} \in D \cap M\}$, где $c = [0, 1]$, $D \cap M$ — множество решений системы

$$\begin{aligned} (c_1, \mathbf{x}) - 2 &\leq 0, & c_1 &= [-4, 3], \\ (c_2, \mathbf{x}) - 34 &\leq 0, & c_2 &= [1, 5], \\ (c_3, \mathbf{x}) - 46 &\leq 0, & c_3 &= [4, 2], \\ (c_4, \mathbf{x}) - 27 &\leq 0, & c_4 &= [3, -6], \\ (c_5, \mathbf{x}) + 3 &\leq 0, & c_5 &= [-1, -1]. \end{aligned}$$

Вершины множества $D \cap M$:

$$\begin{aligned} \mathbf{v}_1 &= [4, 6], & \mathbf{v}_2 &= [9, 5], & \mathbf{v}_3 &= [1, 2], \\ \mathbf{v}_4 &= [11, 1], & \mathbf{v}_5 &= [5, -2]. \end{aligned}$$

Упорядочим вершины по убыванию значений (c, \mathbf{x}) при $\mathbf{x} = \mathbf{v}_i$:

$$\begin{aligned} (c, \mathbf{v}_1) = 6 &> (c, \mathbf{v}_2) = 5 > (c, \mathbf{v}_3) = \\ &= 2 > (c, \mathbf{v}_4) = 1 > (c, \mathbf{v}_5) = -2. \end{aligned}$$

Вершина $\mathbf{v}_1 = [4, 6]$ является оптимальным вектором задачи.

Пусть множество $A = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ таково, что $\mathbf{a}_1 = [6, 4]$, $\mathbf{a}_2 = [7, 2]$, $\mathbf{a}_3 = [4, 1]$. Введем порядок в множестве A в соответствии с порядком значений (c, \mathbf{a}_i) :

$$(c, \mathbf{a}_1) = 4 > (c, \mathbf{a}_2) = 2 > (c, \mathbf{a}_3) = 1.$$

Нужно решить систему неравенств

$$(\mathbf{a}_1 - \mathbf{a}_2, \mathbf{y}) > 0, \quad (\mathbf{a}_2 - \mathbf{a}_3, \mathbf{y}) > 0.$$

Одно из ее решений: $\mathbf{y}_1 = [1, 1]$. Находим

$$\begin{aligned} \max \{(\mathbf{y}_1, \mathbf{x}) : \mathbf{x} \in D \cap M\} &= \max \{(\mathbf{y}_1, \mathbf{v}_1) = 10, (\mathbf{y}_1, \mathbf{v}_2) = 14, \\ &(\mathbf{y}_1, \mathbf{v}_3) = 3, (\mathbf{y}_1, \mathbf{v}_4) = 12, (\mathbf{y}_1, \mathbf{v}_5) = 3\} = (\mathbf{y}_1, \mathbf{v}_2). \end{aligned}$$

Полагаем $a_4 = v_2$. Определим порядок в расширенном множестве $\{a_1, a_2, a_3, a_4\}$:

$$(a_4, c) = 5 > (a_1, c) = 4 > (a_2, c) = 2 > (a_3, c) = 1.$$

Теперь нужно решить систему

$$(a_4 - a_1, y) > 0, \quad (a_1 - a_2, y) > 0, \quad (a_2 - a_3, y) > 0.$$

Новое неравенство $(a_4 - a_1, y) > 0$ является следствием предыдущей системы, так как совпадает с ее вторым неравенством.

Поэтому можно положить $y_2 = y_1$, но тогда $a_5 = v_2$ и т. д., т. е. произошло зацикливание.

2.7. Характеризация сходимости в алгоритме Z_3 .

Теорема 5.2.10. Пусть к задаче (5.2.1) применяется алгоритм Z_3 , $N_1^k \cap Q_1^k \cap D \neq \emptyset$ ($\forall k$), N_1^k и Q_1^k — нормальные элементы, а каждое из включений

$$x_k \in \partial N_1^k, \quad x_k \in \partial Q_1^k$$

выполняется для бесконечного числа индексов k . Тогда $\bar{\rho}(co A_1^k, co B_1^k) \rightarrow 0$, $\bar{\rho}(co S_p^k, co \bigcup_{j < p} S_j^k) \rightarrow 0$ ($k \rightarrow +\infty$).

Обоснование аналогично доказательству теоремы 5.2.2.

Напомним, что в соответствии с алгоритмом Z_3 : A_1^k — множество представителей допустимой области относительно плохо формализуемого ограничения $(c_1, x) - t_1 \leq 0$ на k -й итерации; B_1^k — аналогичное множество из дополнения до \mathbf{R}^n допустимой области; N_1^k — полупространство, граница которого ∂N_1^k делит пополам разделяющий A_1^k и B_1^k плоский слой наибольшей толщины; Q_1^k — аналогичное полупространство для S_p^k и $\bigcup_{j < p} S_j^k$; S_j^k — множество точек из материала обучения $A_1^k \cup B_1^k \cup S_1^k \cup \dots \cup S_p^k$, в которых значение ступенчатой целевой функции $f_0(x)$ равно j ($\forall j \in N_p$); последняя является плохо формализуемой.

2.8. Алгоритм Z_5 . Алгоритм связан с выбором направления e_t в процедуре

$$x_{t+1} = x_t + \varepsilon e_t, \quad (5.2.6)$$

поставленной в соответствие задаче

$$\max \{f(x) : x \in M\}.$$

На шаге t известны конечные множества

$$A_t = \{[\mathbf{x}, \mathbf{e}] \in \mathbb{R}^{2n}: f(\mathbf{x} + \varepsilon \mathbf{e}) > f(\mathbf{x}) + \delta, \mathbf{x} + \varepsilon \mathbf{e} \in M, |\mathbf{e}| = 1\};$$

$$B_t = \{[\mathbf{x}, \mathbf{e}] \in \mathbb{R}^{2n}: \text{либо } f(\mathbf{x} + \varepsilon \mathbf{e}) \leq f(\mathbf{x}) + \delta, \\ \text{либо } \mathbf{x} + \varepsilon \mathbf{e} \notin M, |\mathbf{e}| = 1\}.$$

Для выбора \mathbf{e}_t решается задача дискриминантного анализа:

$$g(\mathbf{a}) > 0 \quad (\forall \mathbf{a} \in A_t), \quad g(\mathbf{b}) < 0 \quad (\forall \mathbf{b} \in B_t), \quad g \in G. \quad (5.2.7)$$

Здесь G — выбранный класс разделяющих функций. Если g_t — некоторое решение системы (5.2.7), то в качестве \mathbf{e}_t в процедуре (5.2.6) выбирается одно из решений системы

$$g_t(\mathbf{x}_t, \mathbf{e}) > 0, \quad |\mathbf{e}| = 1.$$

Исследование этого алгоритма связано с выяснением вопросов:

1) возможность разделения множеств A_t и B_t ;

2) характеристика точки \mathbf{x}_t , в которой происходит останов.

Будем рассматривать упрощенный вариант: $\mathbf{x}_{t+1} = \mathbf{x}_t + \mathbf{e}_t$, условие $|\mathbf{e}_t| = 1$ снимаем. Введем обозначения:

$$A = \{[\mathbf{x}, \mathbf{e}]: f(\mathbf{x} + \mathbf{e}) > f(\mathbf{x}) + \delta; \mathbf{x} + \mathbf{e} \in M\}; \quad B = \mathbb{R}^{2n} \setminus A.$$

Пусть

$$f(\mathbf{x}) = (\mathbf{c}, \mathbf{x}), \quad M = \{\mathbf{x}: (\mathbf{c}_j, \mathbf{x}) \leq b_j \quad (\forall j \in \mathbb{N}_m)\},$$

т. е. исходная задача является задачей линейного программирования. Тогда $A = \{[\mathbf{x}, \mathbf{e}]: (\mathbf{c}, \mathbf{e}) > \delta, (\mathbf{c}_j, \mathbf{x}) + (\mathbf{c}_j, \mathbf{e}) \leq b_j \quad (\forall j \in \mathbb{N}_m)\}$. Следовательно, класс G должен позволять отделять выпуклое многогранное множество от его дополнения. Этому условию удовлетворяет, например, класс комитетных конструкций.

Переходим к вопросу 2). Будем считать M ограниченным многогранником. Процесс заканчивается в такой точке $\bar{\mathbf{x}} \in M$, для которой система

$$(\mathbf{c}, \mathbf{e}) > \delta, \quad (\mathbf{c}_j, \bar{\mathbf{x}}) + (\mathbf{c}_j, \mathbf{e}) \leq b_j \quad (\forall j \in \mathbb{N}_m)$$

несовместна относительно \mathbf{e} . Такой момент обязательно наступит, так как $\forall t: f(\mathbf{x}_{t+1}) > f(\mathbf{x}_t) + \delta$.

Таким образом, неравенство $(c, e) \leq \delta$ есть следствие системы

$$(c_j, e) \leq b_j - (c_j, \bar{x}) \quad (\forall j \in N_m).$$

Согласно теореме Фаркаша — Минковского в этом случае существуют числа

$$y_0 \geq 0, y_1 \geq 0, \dots, y_m \geq 0,$$

при которых

$$(c, e) - \delta \equiv \sum_{j=1}^m y_j [(c_j, e) - b_j + (c_j, \bar{x})] - y_0,$$

откуда

$$\begin{aligned} c &= \sum_{j=1}^m y_j c_j; \\ -\delta &= - \sum_{j=1}^m y_j b_j + \sum_{j=1}^m y_j (c_j, \bar{x}) - y_0. \end{aligned} \quad (5.2.8)$$

Из последнего соотношения

$$\sum_{j=1}^m y_j (b_j - (c_j, \bar{x})) \leq \delta. \quad (5.2.9)$$

Здесь $b_j - (c_j, \bar{x}) \geq 0$ ($\forall j$), так как $\bar{x} \in M$. Подставляя (5.2.8) в (5.2.9), получим

$$\sum_{j=1}^m y_j b_j - (c, \bar{x}) \leq \delta.$$

Это соотношение характеризует близость значений функционалов прямой и двойственной задач в точках \bar{x} и $\underline{y} = [y_1, \dots, y_m]$ соответственно, и тем самым — близость \bar{x} к оптимальному множеству рассматриваемой задачи ЛП.

2.9. Алгоритм, использующий экспертизу E_1^i . Пусть имеется одно плохо формализуемое ограничение $\hat{f}(x) \geq 0$, \hat{f} неизвестно, однако заданы конечные множества $A \subset R^n$ и $B \subset R^n$:

$$\hat{f}(x) > 0 \quad (\forall x \in A), \quad \hat{f}(x) < 0 \quad (\forall x \in B),$$

несущие информацию о $\hat{f}(x)$.

Пусть для всех x из $A \cup B$ известно число

$$\tilde{f}(x) \in [\hat{f}(x) - \varepsilon \hat{f}(x); \hat{f}(x) + \varepsilon \hat{f}(x)],$$

где ε — достаточно малое положительное число, меньшее 1, $\text{sign } \hat{f}(x) = \text{sign } \bar{f}(x)$.

Массив наблюдений $\bar{f}(A \cup B)$ можно использовать, строя поверхность регрессии $\text{graf } r(x)$, аппроксимирующую $\text{graf } \hat{f}(x)$. Для приближенного восстановления функции $\hat{f}(x)$ можно также использовать решения следующей системы неравенств:

$$\begin{aligned} (1 + \varepsilon) f(x) &\geq \bar{f}(x) \quad (\forall x \in A), \\ (1 + \varepsilon) f(x) &\leq \bar{f}(x) \quad (\forall x \in B), \\ (1 - \varepsilon) f(x) &\leq \bar{f}(x) \quad (\forall x \in A), \\ (1 - \varepsilon) f(x) &\geq \bar{f}(x) \quad (\forall x \in B), \\ f &\in F. \end{aligned} \quad (5.2.10)$$

Заметим, что если \bar{f} — решение системы (5.2.10), то \bar{f} — дискриминантная функция для множеств A и B :

$\bar{f}(x) \geq 0 \quad (\forall x \in A), \quad \bar{f}(x) \leq 0 \quad (\forall x \in B), \quad \bar{f} \in F,$
так как

$$\text{sign } \bar{f}(x) = \text{sign } \hat{f}(x) = \begin{cases} 1, & \text{если } x \in A, \\ -1, & \text{если } x \in B. \end{cases}$$

Введем обозначения:

$$D_+ = \{x \in D: \hat{f}(x) > 0\}, \quad D_- = \{x \in D: \hat{f}(x) < 0\}.$$

Рассмотрим систему относительно f :

$$\begin{aligned} (1 + \varepsilon) f(x) &\geq d \quad (\forall [x, d]: x \in D_+, d \in \hat{f}(x)[1 - \varepsilon; 1 + \varepsilon]), \\ (1 + \varepsilon) f(x) &\leq d \quad (\forall [x, d]: x \in D_-, d \in \hat{f}(x)[1 - \varepsilon; 1 + \varepsilon]), \\ (1 - \varepsilon) f(x) &\leq d \quad (\forall [x, d]: x \in D_+, d \in \hat{f}(x)[1 - \varepsilon; 1 + \varepsilon]), \\ (1 - \varepsilon) f(x) &\geq d \quad (\forall [x, d]: x \in D_-, d \in \hat{f}(x)[1 - \varepsilon; 1 + \varepsilon]), \\ f &\in F. \end{aligned} \quad (5.2.11)$$

Теорема 5.2.11. Если множество $\{x: \hat{f}(x) = 0\}$ не имеет внутренних точек, функция \hat{f} непрерывна, множество D телесно и замкнуто, F — класс всех непрерывных определенных на D функций, то решение системы (5.2.11) единственно и совпадает с \hat{f} .

Доказательство. Очевидно, \hat{f} — решение системы (5.2.11). Докажем единственность. Пусть $\hat{f} + f_\Delta$ — тоже

решение системы. Тогда

$$\begin{aligned} f_{\Delta}(\mathbf{x}) &\geq 0 \quad (\forall \mathbf{x} \in D_+ \cup D_-), \\ f_{\Delta}(\mathbf{x}) &\leq 0 \quad (\forall \mathbf{x} \in D_+ \cup D_-). \end{aligned}$$

D телесно и замкнуто, $D_+ \cup D_- = D \setminus \{\mathbf{x}: \hat{f}(\mathbf{x}) = 0\}$, f_{Δ} — непрерывная функция, поэтому $f_{\Delta}(\mathbf{x}) = 0$ ($\forall \mathbf{x} \in D$). Следовательно, теорема доказана.

Если заранее известно, что $\text{co}A \cap \text{co}B = \emptyset$ и \hat{f} — аффинная функция, то в качестве F можно взять класс аффинных функций. В этом случае

$$\hat{f}(\mathbf{x}) = (\hat{u}, \mathbf{x}) - \hat{v}, \quad f(\mathbf{x}) = (u, \mathbf{x}) - v, \quad u \in \mathbb{R}^n, v \in \mathbb{R}.$$

Система (5.2.10) примет вид:

$$\begin{aligned} (1 + \varepsilon) [(u, \mathbf{x}) - v] &\geq \bar{f}(\mathbf{x}) \quad (\forall \mathbf{x} \in A), \\ (1 + \varepsilon) [(u, \mathbf{x}) - v] &\leq \bar{f}(\mathbf{x}) \quad (\forall \mathbf{x} \in B), \\ (1 - \varepsilon) [(u, \mathbf{x}) - v] &\leq \bar{f}(\mathbf{x}) \quad (\forall \mathbf{x} \in A), \\ (1 - \varepsilon) [(u, \mathbf{x}) - v] &\geq \bar{f}(\mathbf{x}) \quad (\forall \mathbf{x} \in B). \end{aligned} \quad (5.2.12)$$

Так как множество решений системы (5.2.12) может быть вырожденным, то для повышения устойчивости процесса вычислений заменим ε на $\delta > \varepsilon$, \geq (\leq) — на $>$ ($<$). Получим систему, которую можно свести к следующей однородной системе:

$$\begin{aligned} (1 + \delta) [(u, \mathbf{x}) - v] &> \bar{f}(\mathbf{x}) \omega \quad (\forall \mathbf{x} \in A), \\ (1 + \delta) [(u, \mathbf{x}) - v] &< \bar{f}(\mathbf{x}) \omega \quad (\forall \mathbf{x} \in B), \\ (1 - \delta) [(u, \mathbf{x}) - v] &< \bar{f}(\mathbf{x}) \omega \quad (\forall \mathbf{x} \in A), \\ (1 - \delta) [(u, \mathbf{x}) - v] &> \bar{f}(\mathbf{x}) \omega \quad (\forall \mathbf{x} \in B), \\ \omega &> 0, \omega \in \mathbb{R}. \end{aligned}$$

Для решения этой системы, где множества A и B могут быть бесконечными (например, $A = D_+$, $B = D_-$), можно применить метод линейной коррекции. Если A и B ограничены, то последовательность приближений в методе коррекции стабилизируется через конечное число шагов.

Рассмотрим теперь один вычислительный процесс для задачи

$$\max \{(c, \bar{\mathbf{x}}): \mathbf{x} \in D, \hat{f}(\mathbf{x}) \geq 0\},$$

где $\hat{f}(x) = (\hat{u}, x) - \hat{v}$, \hat{u} и \hat{v} неизвестны, D — выпуклый многогранник.

Предположим, что известны конечные множества $A_0 \subset D_+$, $B_0 \subset D_-$. Найдем решение $[u_0, v_0]$ системы (5.2.12), где $A = A_0$, $B = B_0$. Пусть

$$x_0 = \arg \max \{(c, x) : x \in D, (u_0, x) - v_0 \geq 0\}.$$

Полагаем

$$A_1 = \begin{cases} A_0 \cup \{x_0\}, & \text{если } \hat{f}(x_0) \geq 0; \text{ иначе} \\ A_0; \end{cases}$$

$$B_1 = \begin{cases} B_0 \cup \{x_0\}, & \text{если } \hat{f}(x_0) < 0; \text{ иначе} \\ B_0. \end{cases}$$

Затем находим решение системы (5.2.12) при $A = A_1$, $B = B_1$, а далее повторяем все остальные шаги и т. д.

2.10. Процесс, использующий метод комитетов. Рассмотрим следующий частный случай задачи (5.1.1):

$$\sup \{f_0(x) : f_j(x) \geq 0 \ (\forall j \in I), x \in D \subset \mathbb{R}^n\}. \quad (5.2.13)$$

Так как в практических задачах функции f_j отражают часто очень сложные аналитические зависимости, то возникает вопрос о методах их построения и о методах решения соответствующих задач (5.2.13), позволяющих эффективно учитывать условие $x \in D$. В ряде задач функции f_j ввиду их сложности могут быть использованы лишь как имитаторы.

В большинстве приведенных выше результатов предполагалось, что разделение пространства на допустимую и недопустимую области осуществляется гиперплоскостью, и в рассмотренных алгоритмах при некоторых условиях строилась последовательность приближений к этой гиперплоскости по мере накопления информации о допустимых и недопустимых векторах в виде множеств A_j^k и B_j^k . Предположим теперь, что допустимое и недопустимое множества аффинно неразделимы. В этом случае можно воспользоваться комитетной дискриминантной функцией, т. е. функцией вида

$$h_i^*(x) = \sum_{i=1}^q \text{sign}^i h_i(x),$$

где h_i — аффинные функции. Возможность использования функций такого вида вытекает из того факта, что для любых двух непересекающихся конечных множеств существует разделяющий их аффинный комитет (§ 3 гл. IV).

Комитетные функции относятся к классу кусочно-аффинных разделяющих функций. Касаясь вопроса о возможности разделения бесконечных множеств, покажем, что если некоторые множества A и B разделяются выпуклой кусочно-линейной функцией, то существует и разделяющая их комитетная функция.

Теорема 5.2.12. *Если $A' \subset A$, A — множество всех решений некоторой конечной системы строгих линейных неравенств, а $B \subset \mathbb{R}^n \setminus \bar{A}$, где \bar{A} — топологическое замыкание множества A , то существует аффинный комитет, разделяющий A' и B' , с числом членов $2m - 1$, где m — число граней максимальной размерности множества A .*

Доказательство. Можно предполагать, что число m граней максимальной размерности равно числу неравенств, т. е.

$$A = \{x: h_j(x) = l_j(x) + \alpha_j > 0 \ (\forall j \in N_m)\},$$

где $l_j(x)$ — линейные функции. Рассмотрим совокупность функций

$$\{h_j: j \in N_m\} \cup \{g_j: j \in N_{m-1}\},$$

где

$$g_j(x) \equiv -1 \ (\forall j \in N_{m-1}).$$

Покажем, что это и есть разделяющий комитет.

Действительно, если $\bar{x} \in A$, то \bar{x} удовлетворяет более чем половине из неравенств $l_j(x) + \alpha_j > 0 \ (\forall j \in N_m)$, $g_j(x) > 0 \ (\forall j \in N_{m-1})$. Если же $\bar{x} \in B$, то \bar{x} удовлетворяет ограничениям $g_j(x) < 0 \ (\forall j \in N_{m-1})$ и хотя бы одному из ограничений $h_j(x) \leq 0$. Теорема доказана.

Очевидно, эту теорему вместе с доказательством можно модифицировать, рассматривая функции $h_j(x)$ из произвольного класса (не обязательно класса аффинных функций).

Пусть $D = \{x: g_j(x) \geq 0 \ (\forall j \in N_k)\}$, $I = N_m$, $D \neq \emptyset$. Система ограничений

$$g_j(x) \geq 0 \ (\forall j \in N_k), \ f_j(x) \geq 0 \ (\forall j \in N_m) \quad (5.2.14)$$

может оказаться несовместной.

В связи с этим возникают следующие постановки задач.

1. Найти все μ -подсистемы для (5.2.14), включающие систему

$$g_j(\mathbf{x}) \geq 0 \quad (\forall j \in N_k). \quad (5.2.15)$$

2. Найти все μ -подсистемы для (5.2.14), включающие (5.2.15), при условии, что функции $\{f_j\}$ допускают их малые вариации, не выходящие за пределы некоторой области (эта постановка связана с вопросами устойчивости комитетных конструкций).

3. Максимизировать $f_0(\mathbf{x})$ на μ -подсистемах системы (5.2.14), включающих (5.2.15), при условии, что возможны малые вариации в задании функций f_j .

4. Максимизировать $f_0(\mathbf{x})$ на звездных множествах системы (5.2.14). Звездным множеством системы (5.2.14) при фиксированном $p \leq t$ и фиксированной подсистеме (5.2.15) называется

$$\{\mathbf{x}: \exists J \subset N_m: g_j(\mathbf{x}) \geq 0 \quad (\forall j \in N_k), \\ f_j(\mathbf{x}) \geq 0 \quad (\forall j \in J), |J| = p\}.$$

Так как решения f_j системы

$$f_j(A_j) \subset \mathbf{R}_+, f_j(B_j) \subset \mathbf{R}_-, f_j \in F, \quad (5.2.16)$$

где F — некоторый класс функций, определяются, вообще говоря, неоднозначно, то естественно рассматривать также следующие задачи.

5. Найти такие решения f_j системы (5.2.16), чтобы система (5.2.14) была совместной.

6. Найти такие решения f_j систем (5.2.16), чтобы минимальный комитет системы (5.2.14) состоял из как можно меньшего числа членов.

7. Найти такие решения f_j систем (5.2.16), которые максимизируют функцию

$$\varphi(f_1, \dots, f_m) = \\ = \max \{f_0(\mathbf{x}): f_j(\mathbf{x}) \geq 0 \quad (\forall j \in N_m), g_j(\mathbf{x}) \geq 0 \quad (\forall j \in N_k)\}.$$

Рассмотрим задачу 7 в линейном случае, т. е.

$$\varphi(\mathbf{x}_1, y_1, \dots, \mathbf{x}_m, y_m) = \\ = \max \{(c, \mathbf{x}): (\mathbf{x}_j, \mathbf{x}) \geq y_j \quad (\forall j \in N_m), (c_j, \mathbf{x}) \geq b_j \quad (\forall j \in N_k)\},$$

где $\forall j \in N_m: [x_j, y_j]$ — решение систем (5.2.16) при условии, что F — класс всех аффинных функций; тогда

$$(x_j, c) > y_j \quad (\forall c \in A_j), \quad (x_j, c) < y_j \quad (\forall c \in B_j). \quad (5.2.17)$$

Требуется найти $\arg \max \{\varphi(x_1, y_1, \dots, x_m, y_m): (5.2.17)\}$. При решении этой задачи можно использовать известные формулы для вычисления производных от функции φ по направлению в пространстве переменных [91].

Рассмотрим теперь один из вариантов использования метода комитетов в процедурах математического программирования. Пусть требуется найти

$$\max \{(c, x): x \in D, h(x) > 0\}, \quad (5.2.18)$$

где

$$h(x) = \sum_{i=1}^q \text{sign}[(z^i, x) - v^i], \quad (5.2.19)$$

$$\text{sign } t = \begin{cases} +1, & t \geq 0, \\ -1, & t < 0, \end{cases} \quad q = 2k + 1.$$

Предполагается, что ограничение $h(x) > 0$ отражает некоторое плохо формализуемое требование. Функцию h можно называть разделяющей функцией (комитетом) эксперта.

Введем обозначения:

$$M(J) = \{x: (z^i, x) - v^i \geq 0 \quad (\forall i \in J)\},$$

$$Y = \{J \subset N_q: |J| = k + 1\},$$

$$M = \bigcup_{J \in Y} M(J).$$

Задача (5.2.18) состоит в нахождении

$$\max \{(c, x): x \in D \cap M\}.$$

Пусть имеются конечные множества $A_0, B_0 \subset \mathbb{R}^n$:

$$h(A_0) \subset \mathbb{R}_+, \quad h(B_0) \subset \mathbb{R}_-.$$

Строим функцию h_0 , разделяющую множества A_0 и B_0 :

$$h_0(x) = \sum_{i=1}^q \text{sign}[(z^{i,0}, x) - v^{i,0}],$$

где $[z^{i,0}, v^{i,0}]$ — решение i -й μ -подсистемы системы:

$$(z, x) - v > 0 \quad (\forall x \in A_0), \quad (z, x) - v < 0 \quad (\forall x \in B_0). \quad (5.2.20)$$

Будем предполагать, что система (5.2.20) имеет не менее q μ -подсистем и что $[z^i, v^i]$ — решение i -й μ -подсистемы. Пусть при этом векторы коэффициентов левых частей этой μ -подсистемы составляют множества $A_{i,0}, B_{i,0}$; пусть, далее, плоскость

$$P_{i,0} = \{x: (z^{i,0}, x) - v^{i,0} = 0\}$$

делит пополам разделяющей $A_{i,0}$ и $B_{i,0}$ плоский слой наибольшей толщины.

Определяем

$$x_0 = \arg \max \{(c, x): x \in D, h_0(x) > 0\}.$$

Положим $\forall i \in N_q$:

$$A_{i,1} = \begin{cases} A_{i,0} \cup \{x_0\}, & \text{если } x_0 \text{ допустимо по всем} \\ & \text{ограничениям и множества } B_{i,0} \text{ и} \\ & A_{i,0} \cup \{x_0\} \text{ аффинно разделимы,} \\ A_{i,0} & \text{в противном случае;} \end{cases}$$

$$B_{i,1} = \begin{cases} B_{i,0} \cup \{x_0\}, & \text{если } x_0 \text{ недопустимо и} \\ & \text{множества } A_{i,0} \text{ и } B_{i,0} \cup \{x_0\} \\ & \text{аффинно разделимы,} \\ B_{i,0} & \text{в противном случае.} \end{cases}$$

Затем для всякого $i \in N_q$ находим гиперплоскость

$$P_{i,1} = \{x: (z^{i,1}, x) - v^{i,1} = 0\},$$

делящую пополам плоский слой наибольшей толщины, разделяющий множества $A_{i,1}$ и $B_{i,1}$. Полагаем

$$h_1(x) = \sum_{i=1}^q \text{sign}[(z^{i,1}, x) - v^{i,1}],$$

находим $x_1 = \arg \max \{(c, x): x \in D, h_1(x) > 0\}$, а далее переходим к следующей итерации. В результате получаем последовательность $\{x_k\}$.

Теорема 5.2.13. Пусть D — выпуклое замкнутое ограниченное множество и $\forall k: x_k \in \bigcup_{i=1}^q (A_{i,k+1} \cup B_{i,k+1})$. Тогда либо

- а) $\exists k, \exists J \in \bar{Y}: x_k = \arg \max \{(c, x): x \in D \cap M(J)\}$, либо
 б) $\exists i \in N_q: \bar{\rho}$ (со $A_{i,k}$, со $B_{i,k}$) $\rightarrow 0$ ($k \rightarrow +\infty$).

Доказательство. Возможен лишь один из следующих трех случаев:

- 1) $\forall \omega \in N, \exists k_\omega \geq \omega, \forall i \in N_q: x_{k_\omega} \notin P_{i,k_\omega}; \exists \omega: x_{k_\omega}$
 допустимо;
 2) $\forall \omega \in N, \exists k_\omega \geq \omega, \forall i \in N_q: x_{k_\omega} \notin P_{i,k_\omega}; \forall \omega: x_{k_\omega}$
 недопустимо;
 3) $\exists \omega \in N, \forall k_\omega \geq \omega, \exists i \in N_q: x_{k_\omega} \in P_{i,k_\omega}$.

Пусть имеет место случай 1). Тогда

$$\exists J \in Y: x_{k_\omega} \in D \cap M(J).$$

Так как $x_{k_\omega} \notin P_{i,k_\omega} (\forall i)$, то

$$x_{k_\omega} = \arg \max \{(c, x): x \in D \cap M(J)\},$$

т. е. выполняется а).

Если имеет место случай 3), то $\exists i = \tau \in N_q$: для некоторого бесконечного подмножества

$$\{x_{k_l}\}: x_{k_l} \in P_{\tau, k_l} (\forall l).$$

Докажем, что

$$\bar{\rho} (\text{со } A_{\tau, k}, \text{ со } B_{\tau, k}) \rightarrow 0 (k \rightarrow +\infty).$$

Так как $A_{\tau, k+1} \supset A_{\tau, k}, B_{\tau, k+1} \supset B_{\tau, k}$, то

$$\bar{\rho} (\text{со } A_{\tau, k+1}, \text{ со } B_{\tau, k+1}) = \rho_{k+1} \leq \rho_k (\forall k).$$

Поэтому $\{\rho_k\} \rightarrow \rho \geq 0$.

Покажем, что соотношение $\rho > 0$ невозможно. Предположим, от противного, что $\{\rho_k\} \rightarrow \rho > 0$. Тогда $\exists \theta, 0 < \theta < 1: \rho_{k_l} \leq \rho_{l-1} \theta$. Действительно, пусть $\rho_{k_{l-1}} = |\bar{x} - \bar{y}|$, $\bar{x} \in \text{со } A_{\tau, k_{l-1}}, \bar{y} \in \text{со } B_{\tau, k_{l-1}}$. Не ограничивая общности, можно предположить, что x_{k_l} находится по ту же сторону от гиперплоскости

$$P_\tau = \{x: (z^\tau, x) - v^\tau = 0\},$$

что и \bar{x} . Тогда $\rho_{k_l} \leq \bar{\rho}(\bar{y}, \bar{x}, x_{k_l})$. Так как

$$|\bar{x} - \bar{y}| \geq \rho, |\bar{y} - x_{k_l}| = |\bar{x} - x_{k_l}| \leq R,$$

где R — радиус шара, включающего ограниченное множество D , то $\exists \theta$, $0 < \theta < 1$:

$$[u = (\bar{x} + \bar{y})/2] \Rightarrow [|x_{k_l} - u| / |x_{k_l} - \bar{x}| \leq \theta].$$

Но тогда и $\tilde{\rho}(\bar{y}, \overline{x_{k_l}}, \bar{x}) \leq \theta |\bar{x} - \bar{y}|$. Таким образом, $\rho_{k_l} \leq \theta \rho_{k_l-1}$ ($\forall l$). Но это противоречит тому, что $\{\rho_k\} \rightarrow \rho > 0$. Этим доказано утверждение б).

Пусть имеет место случай 2). Тогда существует бесконечная последовательность $\{x_{k_l}\}$: $x_{k_l} \notin P_{i, k_l} (\forall i)$, x_{k_l} — недопустимые точки, при этом $\exists \tau$: $x_{k_l} \in B_{\tau, k_l+1}$. Пусть $\rho_{k_l-1} = |\bar{x} - \bar{y}|$, $\bar{x} \in \text{co } B_{\tau, k_l-1}$, $\bar{y} \in \text{co } A_{\tau, k_l-1}$. Точка x_{k_l} находится по ту же сторону от плоскости P_{τ} , что и точка \bar{x} . Тогда $\rho_{k_l} \leq \tilde{\rho}(\bar{y}, \overline{x_{k_l}}, \bar{x})$. Поэтому, как и выше,

$$\exists \theta, 0 < \theta < 1: \rho_{k_l} \leq \theta \rho_{k_l-1} (\forall l),$$

откуда следует 2). Теорема доказана.

Теорема 5.2.14. Пусть D — телесный выпуклый многогранник, \bar{x} — единственное решение задачи (5.2.13) и существует единственный номер $i = i'$ такой, что $(z^{i'}, \bar{x}) - v^{i'} = 0$. Пусть, далее, для всех разделяющих A_0 и B_0 комитетных функций \tilde{h} :

$$\tilde{h}(x) = \sum_{i=1}^q \text{sign}[(\tilde{z}^i, x) - \tilde{v}^i],$$

где $[\tilde{z}^i, \tilde{v}^i]$ — решение i -й μ -подсистемы для (5.2.20), выполняется соотношение

$$|\tilde{z}^i - z^i| \leq \varepsilon, |\tilde{v}^i - v^i| \leq \varepsilon \quad (\forall i \in N_q), \quad (5.2.21)$$

где $[z^i, v^i]$ — некоторое фиксированное решение, $\varepsilon > 0$ достаточно мало. Тогда $\{x_k\} \rightarrow \bar{x}$.

Доказательство. Из (5.2.21) следует что для всякого k , $(z^{i', k}, x_k) - v^{i', k} = 0$; $(z^{i, k}, x_k) - v^{i, k} \neq 0$ ($i \neq k$). Следовательно, множество существенных ограничений k -й задачи оптимизации определяет одно и то же ребро многогранника D . Алгоритм построения x_k таков, что на каждом шаге происходит деление пополам отрезка на этом ребре, содержащего \bar{x} . Следовательно, $\{x_k\} \rightarrow \bar{x}$. Теорема доказана.

2.11. Моделирование плохо формализуемой цели и ее последовательное уточнение. Пусть $\Omega_j = \{\omega_{j1}, \dots, \omega_{jk}\}$ и заданы конечные множества H_i :

$$f_j(H_i) = \{\omega_{ji}\} \quad (j \in \{0\} \cup I, i \in N_k),$$

где $\{\omega_{ji}\}$ — элементы произвольной природы, чаще — действительные числа или интервалы.

Множества H_i служат материалом обучения.

Решим задачу дискриминантного анализа: найти

$$\varphi_{ji} \in \{\mathbf{R}^n \rightarrow \mathbf{R}\} \quad (\forall i \in N_{k-1})$$

из системы соотношений

$$\varphi_{ji}(H_i) \subset \mathbf{R}_+, \quad \varphi_{ji}\left(\bigcup_{r>i} H_r\right) \subset \mathbf{R}_-, \quad \varphi_{ji} \in \Phi;$$

Φ — фиксированный класс отображений.

Модель цели и показателя в этом случае примет вид

$$f_j(\mathbf{x}) = \begin{cases} \omega_{j1}, & \text{если } \varphi_{j1}(\mathbf{x}) > 0; & \text{иначе} \\ \omega_{j2}, & \text{если } \varphi_{j2}(\mathbf{x}) > 0; & \text{иначе} \\ \dots & \dots & \dots \\ \omega_{j, k-1}, & \text{если } \varphi_{j, k-1}(\mathbf{x}) > 0; & \text{иначе} \\ \omega_{jk}. & & \end{cases}$$

Задача поиска оптимального с точки зрения цели плана $\tilde{\mathbf{x}}$ может быть решена путем последовательного случайного генерирования планов \mathbf{x} и получения оценок $f_j(\mathbf{x})$ с помощью построения моделей этих зависимостей.

Можно моделировать цель, используя обратную связь. Оптимальный вектор задачи, в которой в качестве целевой функции берется очередное приближение к ней, оценивается с точки зрения значения в нем действительной функции цели, и на основе этой информации строится очередная, более точная, модель цели. В результате этой процедуры можно получать расширяющиеся от шага к шагу множества H_j :

$$H_j \subset G_j = \{\mathbf{x}: f_0(\mathbf{x}) \in \Delta_j\},$$

где Δ_j — некоторые множества из \mathbf{R} (числа, интервалы и т. д.).

Более общий подход к моделированию цели связан с синтезом информации о ней. Сложная модель цели опре-

деляется представлением о существовании сложного механизма оценивания вида критериальных функций и определения их коэффициентов на основе алгоритмов идентификации.

§ 3. Автоматизация управления параметрами итерационного процесса для задач математического программирования

Управление итерационными процессами счета, основанное на точных утверждениях о свойствах последовательности приближений, часто оказывается неэффективным. Поэтому во многих случаях организация процедур счета опирается на эвристические соображения и интуицию, возникающие при решении большого числа задач определенного класса. Однако эффективность счета может быть существенно повышена, если снабдить программы средствами самообучения управлению параметрами численных процедур, такими, например, как распознавание образов.

3.1. Управление параметрами по ситуационным признакам на основе методов распознавания образов. Пусть $x_{t+1} = \varphi[p](x_t)$ — итерационный процесс (для той или иной задачи оптимизации), заданный с помощью однозначного итерационного оператора $\varphi[p](x)$, зависящего от векторного параметра $p \in R^N$. В качестве параметров, задающих вектор p , могут выступать константы точности, константы метода и т. д. В данном параграфе рассматриваются приемы формализации управления последовательностью $\{p_t\}$ с целью повышения эффективности вычислительного процесса

$$x_{t+1} = \varphi[p_t](x_t) \quad (5.3.1)$$

в условиях, когда признаки (характеристики) процесса, по которым вырабатываются управления $u_t: p_t \rightarrow p_{t+1}$, могут быть косвенными или плохо формализуемыми.

Добавим, что управление u_t может состоять в переключении с одного вида отображения φ на другой; оно может быть как детерминированным, так и стохастическим (например, может настраивать на смешанную стратегию выбора отображения из некоторой фиксированной их совокупности или просто осуществлять выбор отобра-

жения на каждом шаге; вырабатывать значения для субъективных констант счета).

В качестве иллюстрационной рассмотрим задачу линейного программирования над \mathbf{R}^n : $\max(\mathbf{c}, \mathbf{x})$ при ограничениях (2.1.9), т. е.

$$\max \{(\mathbf{c}, \mathbf{x}): L_j(\mathbf{x}) = (\mathbf{h}_j, \mathbf{x}) - b_j \leq 0, \forall j \in N_m\}. \quad (5.3.2)$$

Запишем для нее процесс (градиентного типа):

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \beta(\mathbf{x}_{t-1} - \mathbf{x}_t) - \frac{\lambda}{\delta} \sum_{j=1}^m L_j^+(\mathbf{x}_t) \mathbf{h}_j + \alpha \mathbf{c}; \quad (5.3.3)$$

здесь

$$\delta = \sum_{j=1}^m |\mathbf{h}_j|^2, \quad \lambda \in (0; 2), \quad \beta > 0, \quad \alpha > 0.$$

При достаточно малых положительных α и β процесс (5.3.3) сходится к приближенному решению задачи (5.3.2) [65]. Если в (5.3.3) положить $\alpha = \alpha_t$ и потребовать $\{\alpha_t \geq 0\} \rightarrow 0, \sum_{(t)} \alpha_t = +\infty$, то будет иметь место сходимость процесса к оптимальному вектору задачи (5.3.2). Однако сходимость может оказаться достаточно медленной. Эффективность процесса может быть повышена за счет целесообразного (но неизвестного нам) алгоритма пересчета на каждой итерации (частичного или полного) параметров α, β и λ ; этим самым процесс (5.3.3) заменяется на

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \beta_t(\mathbf{x}_{t-1} - \mathbf{x}_t) - \frac{\lambda_t}{\delta} \sum_{j=1}^m L_j^+(\mathbf{x}_t) \mathbf{h}_j + \alpha_t \mathbf{c}. \quad (5.3.4)$$

Но трудность как раз и состоит в формализации управления u_t , реализующего соответствие

$$\mathbf{p}_t \xrightarrow{u_t} \mathbf{p}_{t+1}, \quad (5.3.5)$$

где $\mathbf{p}_t = [\beta_t, \lambda_t, \alpha_t]$ ($\forall t \in \mathbf{N}$). Следовательно, речь здесь идет о трудности, связанной с плохой формализуемостью управления u_t , хотя примеров удачных управлений (т. е.

способов пересчета вектора p_t на итерации с номером t) может быть достаточно много. Возникает типичная ситуация, в которой целесообразно применение аппарата формализации средствами распознавания образов.

Если реализован отрезок x_0, \dots, x_t траектории, порожаемой процессом (5.3.4), то ситуация, сложившаяся к моменту t вычислительного процесса, может быть формализована (хотя далеко и не однозначным образом) последовательностью y_1, \dots, y_{N_0} признаков y_i — скалярных, векторных или предикатных. В роли скалярных признаков могут выступать, например, значения $(x_{i-1} - x_t, c)$, $i = t - t_0, \dots, t$ (t_0 — заранее выбранное число, задающее отрезок траектории, подлежащей анализу на итерации с номером t); в роли векторных: x_{t-t_0}, \dots, x_t ; в роли предикатных: π_i ($i = 1, 2, 3, 4$), каждый из которых принимает значение $+1$, если выполняется соответствующий признак:

- 1) $\sum_{i=t-t_0+1}^t |x_i - x_{i-1}| \leq \varepsilon_1$;
- 2) $\sum_{j=1}^m L_j^+(x_t) \leq \varepsilon_2$;
- 3) $(c, x_t - x_{t-1}) \geq \varepsilon_3$;
- 4) $|A^T v^t - c| \leq \varepsilon_4$;

в противном случае значение равно -1 ; здесь $v^t = [v_1^t, \dots, v_m^t]$ и $v_j^t = \lambda_j L_j^+(x_t) / \delta \alpha_j$, A — матрица коэффициентов системы ограничений в задаче (5.3.2), ε_i ($i = 1, 2, 3, 4$) — константы точности, которые сами могут быть включены в список скалярных признаков (и, следовательно, могут пересчитываться от шага к шагу). Заметим, что признак π_1 характеризует степень сгущения точек итерационной последовательности, π_2 — близость точки x_t к допустимой области задачи (5.3.2), π_3 — степень возрастания максимизируемой функции (c, x) , π_4 — близость вектора v^t к допустимой области задачи $\min \{(b, v) : A^T v = c, v \geq 0\}$ — двойственной к (5.3.2).

Таким образом, ситуация, сложившаяся на момент t вычислительного процесса, описывается (моделируется) вектором $y^t = [y_1^t, \dots, y_{N_0}^t] \in R^{N_0}$. Для некоторых из них могут быть рекомендованы управления u_t : $u_t(p_t) = p_{t+1}$, подсказанные либо здравым смыслом, либо интуицией и

опытом. Примерами управлений u_i^t могут быть: $u_i^t(\mathbf{p}_t) = \mathbf{p}_t$, если $\pi_4^t = +1$; $u_i^t(\mathbf{p}_t) = [\beta_t, \lambda_t, \alpha_t/2]$, если $\pi_1^t = +1$ и $\pi_2^t = -1$, и т. д.

Естественно, выработка эффективного универсального управления u , реализующего алгоритм перехода $\mathbf{p}_t \rightarrow \mathbf{p}_{t+1}$, существенным образом будет зависеть от объемности материала обучения $M \subset \mathbf{R}^N$, для каждого элемента \mathbf{y} которого считается заданным управление u . Точнее: предполагается, что конечное множество M из пространства ситуаций \mathbf{R}^N разбито на конечное число множеств M_i ($\forall i \in \mathbf{N}_l$) так, что если $\mathbf{y} \in M_i$, то этим определяется однозначным образом управление u_i . Итак, мы теперь предполагаем, что если реализован отрезок $\{x_k\}_0^{t+1}$ итерационной последовательности, порождаемой в силу соотношения (5.3.1) (этим самым реализован и отрезок $\{p_k\}_0^t$), то в случае, если $\mathbf{y}^t \in M_i$, последующий элемент процесса вычисляется по формуле $x_{t+2} = \Phi[p_{t+1}](x_{t+1})$, где $p_{t+1} = u_i(\mathbf{p}_t)$.

Выбор же управления u на случай произвольного $\mathbf{y} \in \mathbf{R}^N$ может быть осуществлен на основе применения методов дискриминантного анализа. А именно, для всякого $j \in \mathbf{N}_{l-1}$ решаем задачу дискриминантного анализа: найти f из условий

$$\begin{aligned} f(\mathbf{y}) &> 0 \quad (\forall \mathbf{y} \in M_j), \\ f(\mathbf{y}) &< 0 \quad (\forall \mathbf{y} \in \bigcup_{i>j} M_i), \\ f &\in F, \end{aligned}$$

где F — выбранный класс разделяющих функционалов. Пусть f_j — решение этой задачи ($j=1, \dots, l-1$). Тогда вводим следующее правило выбора u по ситуации $\mathbf{y} \in \mathbf{R}^N$:

$$u = \begin{cases} u_1, & \text{если } f_1(\mathbf{y}) > 0; \text{ иначе} \\ u_2, & \text{если } f_2(\mathbf{y}) > 0; \text{ иначе} \\ \dots & \dots \\ u_{l-1}, & \text{если } f_{l-1}(\mathbf{y}) > 0; \text{ иначе} \\ u_l. & \end{cases}$$

Заметим, что логическим основанием рассмотрения процесса (5.3.4), играющего здесь иллюстративную роль, служит

Утверждение 5.3.1. Пусть оптимальное множество задачи (5.3.2) не пусто и ограничено. Тогда при

произвольном генерировании параметров $\beta_t, \lambda_t, \alpha_t$ таких, что $\beta_t \leq \beta, \lambda_t \in [\varepsilon; 2 - \varepsilon], \{\alpha_t \geq 0\} \rightarrow 0$ и $\sum_{(t)} \alpha_t = +\infty$ (здесь $\varepsilon < 1, \beta$ — достаточно мало), последовательность $\{x_t\}$, задаваемая в силу (5.3.4), ограничена и

$$\{x_t\}' \subset \text{Arg}(5.3.2),$$

т. е. все предельные точки ограниченной последовательности принадлежат оптимальному множеству задачи (5.3.2).

В соответствии с описанными методами управления параметрами по ситуации проводились обширные численные эксперименты (на задачах линейного программирования), которые подтвердили возможность существенного повышения эффективности вычислительных процессов.

3.2. Управление рандомизированными фейеровскими процессами. В § 5 гл. II был рассмотрен процесс (2.5.1), порождаемый совокупностью фейеровских отображений $\{\varphi_1, \dots, \varphi_m\}$, применяемых для порождения процесса в той или иной их последовательности. Заменим процесс (2.5.1) на однозначный:

$$\{x_{t+1} = \varphi[p_t](x_t)\}_{v_i}; \quad (5.3.6)$$

здесь $\varphi[p](x)$ задается согласно (2.5.2), $p_t = [p_t^1, \dots, p_t^m] \geq 0$,

$\sum_{i=1}^m p_t^i = 1$. В § 5 гл. II были рассмотрены некоторые детерминированные методы управления последовательностью $\{p_t\}$. Однако в основу автоматизации управления параметром p могут быть положены и приемы предыдущего пункта без каких-либо модификаций. Что касается вопросов сходимости возникающих при этом процессов (5.3.6), то они полностью решаются теоремами из § 5 гл. II.

ПРИЛОЖЕНИЯ В ЗАДАЧАХ ПЛАНИРОВАНИЯ И УПРАВЛЕНИЯ ТЕХНИКО-ЭКОНОМИЧЕСКИМИ И ПРИРОДНЫМИ СИСТЕМАМИ

Вопросы прикладных возможностей итерационных методов в их нестационарном варианте были обсуждены в гл. I. Интерпретации нестационарного процесса, данные в §§ 3, 4 этой главы, по существу, указывают и на области некоторых применений, в частности и в самой математике (вычислительной математике).

Ниже будет идти речь о некоторых прикладных задачах, по некоторым из которых накоплен реальный опыт счета. В решении этих задач большое место занимали, наравне с оптимизацией, методы классификации. Последние упорядочивают исходную информацию и результаты счета, подводя определенный итог как этапам решения задачи, так и решению задачи в целом.

§ 1. Технико-экономические приложения

В сфере планирования и управления можно выделить следующие направления приложений, в которых существенно применение алгоритмов распознавания образов: структуризация информации и ее сжатие, типология технико-экономических систем, декомпозиция систем, диагностика систем, предсказание характеристик.

Многие из возникающих здесь задач могут быть естественным образом погружены в более широкие постановки задач оптимизации, рассматриваемых в единстве их информационного, модельного и вычислительного аспектов.

1.1. Агрегирование в многомерных задачах принятия решений. Прикладные задачи ЛП (такие, как задачи планирования работы предприятий, задача проектирования технических систем и другие) имеют как правило

большую размерность, и их решение на ЭВМ с помощью общих методов ЛП практически неосуществимо. В связи с этим разрабатываются специальные методы декомпозиции (например, метод Данцига—Вулфа), различные точные методы, учитывающие специальную структуру матрицы коэффициентов задачи. Однако и эти методы не решают в полной мере проблемы большой размерности, при счете на ЭВМ происходит неконтролируемое накопление ошибок, что не позволяет получить достаточно приемлемое приближение.

Ввиду этого разрабатываются и применяются различные эвристические подходы, позволяющие выделить значимые ограничения и переменные задачи, заменять группы ограничений их выпуклыми комбинациями и т. д.

Однако эти подходы нуждаются в точном математическом обосновании, иначе их применение может снизить объективность и обоснованность принимаемых практических решений. В качестве аппарата для обоснования методов агрегирования ограничений и переменных могут служить математические методы распознавания образов. Действительно, задача разбиения множества ограничений на группы такие, что ограничения одной группы достаточно близки друг к другу, а ограничения разных групп достаточно далеки друг от друга, представляют собой типичную задачу распознавания образов.

Приведем механизм *агрегирования ограничений и переменных* в задаче линейного программирования

$$\max \{ (c, x) : Ax \leq b, x \geq 0 \} = \bar{f}, \quad (6.1.1)$$

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}, \quad b = [b_1, \dots, b_m]^T,$$

моделирующей некоторую технико-экономическую систему. Обозначим через p_j j -й столбец матрицы A , через h_i — ее i -ю строку.

Возможны две схемы агрегирования ограничений и переменных в задаче (6.1.1): *замкнутая* и *открытая*.

Замкнутая схема отвечает предположению, что будет использована только информация о задаче (6.1.1) в виде элементов A , b , c и что нет возможности обращаться к другим источникам сведений, имеющих, как правило,

качественный характер. В этом случае проводится таксономия строк h_i , предварительно нормализованных, и таксономия столбцов p_j , взятых с весами, выравнивающими влияния различия в координатах вектора c .

Открытая схема используется в предположении, что мы не отрываемся от практического смысла коэффициентов матрицы A и векторов b, c, x . В этом случае типология строк и столбцов и их объединение в таксоны создаются, исходя из содержательного смысла.

Таксономия в замкнутой схеме дает классы технологических способов $P_k = \{p_i: i \in J_k\}$ ($k = 1, \dots, l$), $\bigcup_{k=1}^l P_k = \{p_i: \forall i \in N_n\}$, $P_i \cap P_j = \emptyset$ при $i \neq j$. Взяв из каждого класса P_k по представителю p_{i_k} — в качестве эталона, получим агрегированную задачу:

$$\max \left\{ \sum_{k=1}^l c_{i_k} y_{i_k}; \sum_{k=1}^l y_{i_k} p_{i_k} \leq b, y_{i_k} \geq 0, \forall k \in N_l \right\}. \quad (6.1.2)$$

Пусть $\tilde{f}(p_{i_1}, \dots, p_{i_l})$ — оптимальное значение этой задачи. Качество агрегирования можно подчинить требованию:

$$|\tilde{f} - \tilde{f}(p_{i_1}, \dots, p_{i_l})| \leq \varepsilon \quad (6.1.3)$$

при заданном ε и любом варианте выбора $p_{i_k} \in P_k$ ($\forall k \in N_l$). Исходя из заданного результата таксономии, можно добиться обеспечения

$$\min \{ |\tilde{f} - \tilde{f}(p_{i_1}, \dots, p_{i_l})| : p_{i_k} \in P_k, \forall k \in N_l \}. \quad (6.1.4)$$

Аналогичная конструкция может быть применена и к задаче агрегирования ограничений. Если D_k ($\forall k \in N_q$) — таксоны массива строк $[h_j, b_j]$ ($j = 1, \dots, m$), то агрегированная задача запишется следующим образом:

$$\max \{ (c, x) : (h_{i_k}, x) \leq b_{i_k}, \forall k \in N_q, x \geq 0 \}.$$

Переходя к открытой схеме, мы будем смотреть на выражение $\sum_{i=1}^n p_i x_i$ как на грубую модель исчисления результата взаимодействия технологических способов

с номерами $1, \dots, n$, применяемых с интенсивностями x_1, \dots, x_n . Действительный же результат этого взаимодействия может быть иным, обозначим его через $G(\{x_i: i \in N_n\})$.

По аналогии с (6.1.3) может быть рассмотрена задача: обеспечить по заданному $\varepsilon > 0$

$$\left| G(\{\bar{x}_i: i \in N_n\}) - \sum_{k=1}^l p_{i_k} \tilde{x}_{i_k} \right| \leq \varepsilon,$$

где $[\bar{x}_1, \dots, \bar{x}_n]$:—оптимальный план задачи (6.1.1), а $[\tilde{x}_{i_1}, \dots, \tilde{x}_{i_l}]$ —оптимальный план агрегированной задачи (6.1.2).

В рамках открытой схемы можно решать и задачу, аналогичную задаче (6.1.4), а именно:

$$\min \left\{ \left| G(\{\bar{x}_i: i \in N_n\}) - \sum_{k=1}^l p_{i_k} \tilde{x}_{i_k} : p_{i_k} \in P_k, \forall k \in N_l \right. \right\}.$$

В открытой схеме также может быть рассмотрено агрегирование строк с уже введенными типами требований качества.

На самом деле может быть осуществлено агрегирование и строк, и столбцов одновременно. Один из способов такого агрегирования состоит в следующем. Выписываем симметрическую задачу, соответствующую задаче (6.1.1): найти $[x, u]$ из условий

$$\begin{aligned} Ax &\leq b, \\ A^T u &\geq c, \\ x &\geq 0, \quad u \geq 0, \\ (c, x) - (b, u) &\geq 0. \end{aligned} \tag{6.1.5}$$

Агрегирование по столбцам системы (6.1.5) осуществляет одновременное «сжатие» и строк, и столбцов матрицы исходных данных задачи (6.1.1) и отнесено не к преобразованию допустимой области, а к преобразованию декартова произведения оптимальных множеств задачи (6.1.1) и двойственной к ней, ибо именно это произведение задается системой (6.1.5).

Описанный подход к агрегированию апробирован на задаче *объемно-календарного планирования* одного круп-

ного машиностроительного предприятия [60]. Параметры задачи: число ограничений $m \geq 2000$ (без учета условий типа $\underline{x}_i \leq x_i \leq \bar{x}_i$), число переменных $n \geq 10000$. Программа преобразования массива данных позволяет формировать по исходной задаче ряд ее гомоморфных образов с понижением размерности. Применение этой программы не ограничено алгоритмами таксономии. Для решения задачи в целом используется пакет программ производственного планирования, реализующий подготовку данных, их проверку и коррекцию, алгоритмы оптимизации (конечные и итерационные), анализ полученных результатов, генерацию отчетных документов. Пакет обеспечен средствами решения задачи в диалоговом режиме.

Добавим, что решение указанной задачи не сводится к одноразовому (или несколькоразовому) просчету, хотя и объемному, а является непрерывным процессом, направленным на отслеживание динамики производства по системе его различных характеристик (в частности, по эволюционирующей системе исходных данных) и адаптацию рабочей модели к условиям реального производства.

Сформулированный здесь подход к обеспечению решения задач оптимального текущего и оперативного планирования производства, характеризующегося подвижностью, и должен быть, на взгляд авторов, положен в основу создания подсистем АСУП, реализующих такое планирование.

1.2. Идентификация экономико-математических моделей. При планировании управления плохо структурированными технико-экономическими системами с помощью математических моделей целесообразно использование итерационных процессов последовательного уточнения модели. В этих процессах результат, полученный при формальном анализе модели, используется, согласно принципу обратной связи, для уточнения модели. Такова схема управления плохо формализуемыми процедурами планирования.

Рассмотрим вопрос о построении экономико-математической модели, некоторые из ограничений которой могут представлять собой требования допустимости планов по плохо формализуемым критериям. Эти ограниче-

ния идентифицируются, как уже было показано в гл. V, на основе методов распознавания образов, причем для обучения распознаванию используются конечные множества допустимых и недопустимых по плохо формализуемым критериям планов. Если указанные множества пополняются с течением времени, то возникает итерационный процесс решения исходной задачи (см. § 2 гл. IV).

Пусть \mathfrak{X} — некоторая технико-экономическая система, $\mathbb{R}^n \ni \mathbf{x}$ — вектор ее состояния, $\mathbb{R}^m \ni \mathbf{y}$ — вектор состояния среды, $\mathbb{R}^p \ni \mathbf{u}$ — вектор управления, $\mathbb{R}^n \supset X$ — допустимое множество состояний системы, Y и U — аналогичные множества соответственно для состояния среды и возможных управлений. Через $f[\mathbf{y}_t, \mathbf{u}_t](\mathbf{x})$ обозначим показатель эффективности функционирования системы на момент t в состоянии системы \mathbf{x} .

Если информация о системе задана в виде конечных подмножеств $X'_t \subset X_t$, $Y'_t \subset Y_t$, $U'_t \subset U_t$, $X''_t \subset \mathbb{R}^n \setminus X_t$, $Y''_t \subset \mathbb{R}^m \setminus Y_t$, $U''_t \subset \mathbb{R}^p \setminus U_t$, а также множества

$$\{f[\mathbf{y}_t, \mathbf{u}_t](\mathbf{x}): \mathbf{x} \in X'_t, \mathbf{y}_t \in Y'_t, \mathbf{u}_t \in U'_t\},$$

то они по совокупности выступают в качестве материала для приближенного представления множеств X_t , Y_t и U_t , а также функции f . Эти приближения \tilde{X}_t , \tilde{Y}_t , \tilde{U}_t и \tilde{f} могут быть построены, как уже было показано в гл. V, на основе методов дискриминантного анализа. Тогда модель

$$\sup \{\tilde{f}[\mathbf{y}_t, \mathbf{u}_t](\mathbf{x}): \mathbf{x} \in \tilde{X}_t, \mathbf{y}_t \in \tilde{Y}_t, \mathbf{u}_t \in \tilde{U}_t,$$

описывает объект по системе априорных сведений о ней (на момент t). Так как модель нестационарна, то она допускает организацию для нее нестационарного процесса в силу того или иного итерационного оператора.

Описанный подход применялся к решению задач планирования ресурсной базы металлургического производства, планирования использования трудовых ресурсов, оптимального выбора структуры парка машин для ведения открытых горных работ.

1.3. Управление технологическими процессами. Пусть состояние производства (технологического процесса) характеризуется вектором \mathbf{x} параметров и признаков (прямых или косвенных): $\mathbf{x} = [x_1, \dots, x_n]$. При введении

управляющего воздействия $u \in U$ можно вычислить значение функции $f(x, u)$, характеризующей эффективность функционирования производства (процесса). Всякому состоянию x соответствует наиболее эффективное управление $u = u(x)$, реализующее

$$\max \{f(x, u): u \in U\}.$$

Будем предполагать, что элементы u (их природа при описываемом подходе не имеет значения) составляют конечное множество $U = \{u_1, \dots, u_m\}$. Совокупность M всех возможных состояний процесса разбиваем на m подмножеств M_1, \dots, M_m так, что всякому $x \in M_j$ соответствует наилучшее управление $u(x) = u_j$. Задачу нахождения такого управления для любого x теперь следует понимать как задачу установления разбиения

$M = \bigcup_{j=1}^m M_j$, если известны конечные подмножества $M'_j \subset M_j$ ($\forall j \in N_m$), для элементов которых известны наилучшие управления.

В качестве примера конкретной ситуации, отвечающей возможности применения данного метода, рассмотрим задачу управления технологическими процессами извлечения полезных компонентов из руды на горно-обогатительном комбинате. Рассматриваемые процессы включают в себя дробление, измельчение, флотацию и другие операции. Введем в рассмотрение векторы параметров сырья (таких, например, как дробимость руды, ее выщелачиваемость, удельный вес, минеральный состав и др.), параметров технологического режима (величина щели между дробящими органами, загрузка дробилки и др.), показателей качества процесса (таких, как потребление мощности, выход дробленого продукта, шум, производимый дробилкой, и др.). При этом зависимости показателей качества от параметров сырья и технологического режима часто бывают плохо определенными.

Помимо этой задачи, с помощью описанного метода с привлечением методов математического программирования могут быть решены следующие задачи:

1) диагностика технологического процесса с целью управления им;

2) моделирование зависимостей показателей качества от параметров сырья и технологического режима;

3) оптимизация качества при условии допустимости всех параметров.

Кроме того, в соответствии с рассмотренной схемой решались задачи управления процессом переработки пластмасс, управления и прогнозирования качества продукции металлургического производства, определения маршрутов обработки деталей.

1.4. Выбор оптимальных параметров и определение области эффективности технических систем. Пусть $x \in R^n$ — вектор параметров технической системы, $\omega \in R^m$ — вектор состояния среды, в которой работает система. Имеется материал обучения: множества $A, B \subset R^n$, причем, если $[x, \omega] \in A$, то техническая система с вектором параметров x эффективно работает в среде, характеризующейся вектором ω . Если же $[x, \omega] \in B$, то система работает неэффективно. Решаем задачу дискриминантного анализа:

$$x^*(x, \omega) > 0 \quad (\forall [x, \omega] \in A),$$

$$x^*(x, \omega) < 0 \quad (\forall [x, \omega] \in B), \quad x^* \in X^*,$$

где X^* — выбранный класс разделяющих функций. Пусть \bar{x}^* — решение этой системы. Рассмотрим неравенство

$$\bar{x}^*(x, \omega) > 0.$$

Если задан вектор состояния среды ω , то выбираем вектор x так, чтобы выполнялось это неравенство. Если задано x , то можно определить область Ω эффективности использования технической системы:

$$\Omega = \{\omega: \bar{x}^*(x, \omega) > 0\}.$$

Рассмотрим далее конкретные примеры.

В качестве примера технической системы рассмотрим самолет. Он состоит из планера, двигателя, аппаратуры слепой посадки и других подсистем. Имеется функциональное назначение этой системы, оно определяет цель ее функционирования. В соответствии с этим может быть введен показатель качества:

$$f = f_1/f_2,$$

где f_1 — уровень качества (например, грузооборот), f_2 — стоимость системы. Необходимо, чтобы выполнялись следующие соотношения:

$$f_1 \geq y,$$

где y — некоторый фиксированный уровень;

$$f_1 = f_1(y_1, \dots, y_n),$$

$$f_2 = f_2(y_1, \dots, y_n),$$

где y_i — технико-экономические показатели ($i \in \mathbf{N}_n$). При этих условиях нужно найти наибольшее значение показателя f . Сами показатели в свою очередь могут быть результатом расчетов по сложным алгоритмам и иметь ограниченную достоверность.

В этом смысле ограничения и цель данной оптимизационной задачи являются не полностью определенными: мы имеем только возможность проверить их выполнение для некоторого конечного множества вариантов (для обучающего множества). Описанной схеме естественным образом можно придать динамику.

1.5. Прогнозирование технико-экономических показателей. Задачи технико-экономического планирования на предприятии носят существенно творческий характер, требуют применения широких знаний, большого опыта и развитой интуиции специалиста. Поэтому переход к математической формализации этапов принятия экономических решений, необходимость которого диктуется сложностью технологии и организации современного производства, наталкивается на целый ряд трудностей, связанных с проблемой моделирования плохо формализуемых ограничений, связей и целей.

Действительно, рассмотрим, например, задачу исчисления технико-экономических показателей. Опытный экономист (эксперт) обладает сложным неформализованным правилом, моделирующим процесс формирования показателей. В целом такой процесс переработки информации основан на богатом опыте работы с таблицами показателей, на знании конкретных трудно выделяемых факторов, истории развития соответствующих экономических структур, сравнении с близкими ситуациями, встречавшимися ранее.

При формировании новых вариантов организации производства эксперт пользуется своим плохо формализуемым знанием практической реализуемости тех или иных хозяйственных ситуаций. При моделировании зависимостей, связывающих технико-экономические показатели с большим числом факторов, он проявляет способность ранжировать факторы, выделять первостепенные, затем факторы второй очереди и так далее. Эксперт интуитивно угадывает предельные, пороговые значения параметров, при которых рассматриваемый процесс или продукт вытесняется другим, более эффективным для новых значений параметров. Плохо формализуем также анализ чувствительности и эластичности показателей, анализ с помощью допустимых вариаций параметров.

Технико-экономические показатели, различным образом формализуемые (вещественные числа, интервальные оценки, качественные признаки, балльные оценки и т. п.), отражают те или иные объективные характеристики технологических процессов и организации производства. Эти показатели, очевидно, зависят друг от друга и от ряда косвенных параметров производства. Отсюда вытекает возможность прогнозирования показателей, характеризующих новые варианты производства, на основе составления ряда вариантов новых технологий и их оценки с помощью предыдущего опыта.

Имитация процесса переработки информации экспертом — типичная задача, для решения которой привлечение алгоритмов распознавания образов является естественным способом преодоления плохой формализуемости этого процесса.

1.6. Оценка качества продукции. Отметим ряд трудностей, связанных с задачей оценки качества продукции. Прежде всего, сложна проблема соизмерения разнородных видов продукции, товаров, видов деятельности и т. д. Вторая трудность состоит в необходимости квантификации качественных признаков. Задача становится еще более сложной, если речь идет о нескольких качественных признаках и нужно сформировать некоторый синтетический показатель. Следует также сказать о том, что оценка часто связана с неопределенными целями и их эволюцией.

Один из практических подходов, позволяющих в какой-то мере справиться с указанными трудностями, состоит в передаче ЭВМ опыта эксперта, в обучении ЭВМ с помощью распознавания образов.

1.7. Приложения в экономической географии. Укажем на работы этого направления, сводящиеся к агрегированию строк и столбцов матрицы экономико-географических данных. Строки матрицы — пункты или экономико-географические объекты, столбцы — те или иные их характеристики. Выделим следующие задачи этого направления приложений:

изучение пространственного распределения человеческой деятельности;

решение задачи размещения: выбор эталонных пунктов размещения (что связано с разбиением географического пространства на однородные районы), закономерности размещения населенных пунктов в географическом пространстве;

районирование;

оценка ресурсов экономико-географического района: трудовых, климатических, биологических; прогнозирование месторождений полезных ископаемых; комплексная оценка и диагностика района;

типология экономико-географических объектов: предприятий, ресурсов, районов, населенных пунктов;

прогноз влияния осуществления технологических проектов на элементы биогеоценозов.

Учет неструктуризованных данных о ресурсах в моделях математического программирования может быть осуществлен на основе методов дискриминантного анализа. Приведем пример. Пусть участки земли описываются системой признаков различного характера. Для части участков известны данные о наличии тех или иных полезных ископаемых. Пусть $M \subset \mathbb{R}^n$ — множество всех участков, описываемых n -мерными векторами \mathbf{x} в пространстве \mathbb{R}^n признаков; A_1, \dots, A_k — подмножества из M такие, что векторы из A_j соответствуют участкам, где оценка наличия полезного ископаемого равна некоторой величине p_j . Задача прогнозирования оценки участка, описываемого вектором $\mathbf{x} \in M$, на наличие полезного ископаемого сводится к разбиению $M = \bigcup_{i=1}^k M_i$,

$M_i \cap M_j = \emptyset$ ($i \neq j$), $M_i \supset A_i$, ($\forall i \in N_k$), получаемому методами дискриминантного анализа.

1.8. Обзор областей технико-экономических приложений. Перечислим задачи, характеризующиеся наличием плохо определенных факторов, а потому решение их может осуществляться методами нестационарного программирования:

оценка ситуаций в методе сценариев;

нахождение области неопределенности развития больших систем;

предсказание отдельных последствий принимаемых решений;

прогнозирование результатов управляющих воздействий, в частности при управлении большими технико-экономическими и природными системами;

моделирование функций предпочтения при прогнозировании;

выбор вариантов технологического оформления производственных процессов;

выбор эталонных пунктов размещения предприятий при прогнозировании развития отрасли;

диагностика ситуаций при прогнозировании развития крупных фирм;

прогноз интегральных свойств системы по характеристикам ее подсистем;

классификация технических, экономических, географических и биологических объектов с целью выделения эталонов, пользуясь которыми можно прогнозировать поведение близких объектов;

прогноз территориального районирования;

моделирование сложных производственных зависимостей;

прогноз нормативов на новые изделия и прогноз области их применения;

оценка ресурсов районов, комплексная оценка районов, прогнозирование рудоносности;

типология технико-экономических систем;

учет социальных факторов в задачах экономического планирования (социально-экономическое планирование)

и др.

§ 2. Приложения в биологии и медицине

Перечислим некоторые виды приложений математического программирования и распознавания образов в биологии и медицине:

создание систематики;

оценка сходства организмов;

диагностика и проектирование природно-технических систем;

определение границ существования биологических систем в пространстве факторов;

построение математических моделей типа межпопуляционного баланса, а также оптимизационных моделей с учетом плохо формализуемых требований (пример: моделирование закономерностей роста и видового состава той или иной популяции растений в лесу в зависимости от факторов: тепло, влажность, почва, состав древесных пород и др.);

интерпретация данных (например, распознавание по аэрофотосъемкам видового состава лесных массивов, запасов древесины);

управление биологическими системами;

медицинская диагностика;

моделирование возникновения и протекания болезней;

управление процессом лечения;

оценка информативности симптомов;

выделение областей критического риска в пространстве факторов.

2.1. Оценка ресурсной базы в биогеоценозах. Ресурсы биогеоценозов, понимаемые в широком смысле (солнечная энергия, кислород, вода, неорганические вещества, а затем и биомассы различных иерархически организованных уровней жизни), не всегда удовлетворяют потребности животных и растений. Ограниченность некоторых ресурсов оказывает влияние на интенсивность биохимического круговорота.

Запишем в статическом варианте следующую модель-схему:

$$\max \{f(x): g(x) \leq v, x \in Q, x \geq 0\}, \quad (6.2.1)$$

в которой x — вектор параметров биогеоценоза (включая численности популяций), $g(x) \leq v$ — векторное ограни-

чение на потребление ресурсов \mathbf{v} , Q —множество допустимых состояний биогеоценоза, $f(\mathbf{x})$ —критерий его качества, например суммарный выход биомассы.

Пусть все элементы модели (6.2.1) идентифицированы и $\bar{\mathbf{x}} = \bar{\mathbf{x}}(\mathbf{v})$ —ее оптимальный вектор.

Рассматривая узкий диапазон изменения вектора \mathbf{v} , приблизим задачу (6.2.1) задачей линейного программирования L (на основе принципа линеаризации). Если $\mathbf{y} = [y_1, y_2, \dots]$ —вектор двойственных оценок ресурсов, то, согласно одной из интерпретаций, координата этого вектора есть мера дефицитности соответствующего ресурса.

Знание таких оценок может служить инструментом управления биогеоценозами. В случае, когда модель (6.2.1) заменяется на нестационарную, возникает возможность отслеживать динамику вектора двойственных оценок, а вместе с тем и тенденции в потреблении ресурсов.

2.2. Динамика возрастно-генетического состава биопопуляции. Рассматриваемая ниже модель является отражением изучения реальной популяции лягушек, модель была положена в основу машинной имитации развития последней. В популяции выделим особи 1-го, 2-го, ..., m -го генотипов. Продолжительность жизни в этой популяции ограничена сверху числом лет $n \geq 1$. Состояние популяции в момент $t=0, 1, 2, \dots$ описывается матрицей

$$X_t = [x_{ij}^t] = \begin{bmatrix} x_{10}^t & x_{11}^t & \dots & x_{1n}^t \\ x_{20}^t & x_{21}^t & \dots & x_{2n}^t \\ \dots & \dots & \dots & \dots \\ x_{m0}^t & x_{m1}^t & \dots & x_{mn}^t \end{bmatrix},$$

где x_{ij}^t —численность особей i -го генотипа в возрасте j лет. Процесс развития популяции записывается с помощью следующих рекуррентных соотношений:

$$\begin{aligned} x_{i0}^{t+1} &= f_i(y_{t+1}), \\ x_{ij}^{t+1} &= \alpha_{i,j-1}^t x_{i,j-1}^t \quad (j \neq 0), \\ y_{t+1} &= \sum_{i=1}^m \sum_{j=1}^n \beta_{ij}^{t+1} x_{ij}^{t+1}. \end{aligned} \quad (6.2.2)$$

Здесь $0 \leq \alpha_{ij}^t$ и $0 \leq \beta_{ij}^t$ —коэффициенты соответственно выживания и выхода икры для особей i -го генотипа

в возрасте j в период t , y_t — суммарный выход икры, $f_t(y_t)$ — зависимость числа сеголеток от количества икры y_t .

Помимо вычислительных экспериментов, проведенных по просьбе биологов, был проведен формальный анализ модели (6.2.2) (при некоторых предположениях относительно функций f_t), в результате чего были установлены условия, при которых популяция вымирает, либо со временем ее генетико-возрастная структура стабилизируется вместе с общей численностью, либо эта численность неограниченно возрастает.

2.3. Медицинская диагностика и управление процессом лечения. Задача медицинской диагностики может быть сведена к задачам дискриминантного анализа и таксономии, так как она представляет собой разбиение пространства векторов состояния на классы, соответствующие различным видам заболеваний, по конечному материалу обучения.

Из практических задач медицинской диагностики, решавшихся с участием одного из авторов, упомянем следующие: диагностика некоторых онкологических заболеваний (в частности, тиреоидных состояний), инсультов.

Рассмотрим теперь задачу управления процессом лечения. Повсюду ниже t означает момент времени, которому соотнесено состояние модели.

Пусть x_t — n -мерный вектор параметров больного, выявленных при его обследовании, y — m -мерный вектор параметров лечебного процесса, $f_j(x_t, y)$ — значение j -го параметра состояния больного, являющегося реакцией на входные векторы x_t и y ($j=0, 1, \dots, p$), a_j и b_j — допустимые верхняя и нижняя границы значения j -го параметра.

В качестве критерия оптимизации может быть взята одна из реакций, например f_0 со смыслом ее максимизации. Возникает задача

$$\max_{(y)} \{f_0(x_t, y) : b_j \leq f_j(x_t, y) \leq a_j \ (\forall j \in N_m)\}.$$

Анализ и организация численных процедур для этой модели могут быть осуществлены в соответствии с методами нестационарного программирования.

§ 3. Программное обеспечение решения задач производственного планирования и распознавания образов

Ниже будет дана общая информация о двух пакетах прикладных программ, используемых для решения конкретных задач, охарактеризованных в предыдущем параграфе.

Пакеты прикладных программ (ППП) — это новая специальная форма организации комплекса программ, обеспечивающая возможность эффективного решения большого круга задач из заданного класса; повышение гибкости использования всей совокупности программ (с различным функциональным назначением), составляющей модульную базу пакета; открытость, т. е. возможность развития пакета за счет включения в него новых программ; снижение сроков исполнения этапов или частей сложной прикладной задачи; резкое сокращение неавтоматизированных операций в процессе счета; возможность интерактивного режима решения задач.

Внедрение идеологии ППП меняет саму технологию программирования прикладных задач, предъявляя новые требования к исполнению его модулей, к его организации, уровню стандартизации и документированию [62, 98, 110, 111].

3.1. Пакет программ оптимального производственного планирования. Названный пакет является проблемно-ориентированным и предназначен для решения задач линейного программирования большой размерности в режиме потока, упорядоченного во времени и удовлетворяющего условию «близости» задач, его составляющих. Другое назначение пакета — оперативное опробование, внедрение в практику новых методов и модификации известных, т. е. пакет является и методо-ориентированным. Сфера его возможных применений в области оптимального планирования является весьма широкой. Однако специфические черты пакета, связанные в первую очередь с возможностью программных преобразований модели по ходу дела, внесенных в процесс счета корректив определились в результате накопления опыта решения конкретной задачи, а именно: задачи динамического объемно-календарного планирования работы крупного машинострои-

тельного предприятия с мелкосерийным и индивидуальным характером производства.

Пакет «ОКП» создан для машины БЭСМ-6, используемая система программирования — «Автокод-СОМИ», операционные системы «ДИСПАК» и «ДИАПАК».

При разработке пакета учитывалась необходимость обеспечения оперативного изменения исходных технико-экономических данных, диалогового режима, позволяющего вмешиваться в процесс решения задачи и учитывать внемоделльные ограничения. Это дает возможность на процесс решения смотреть как на процесс отслеживания происходящих в реальной системе изменений и гибко учитывать их при человеко-машинных процедурах получения оптимальных планов. Учет внемоделльных условий позволяет формировать достаточно хорошее приближение оптимального плана без применения точных методов на основе эвристических приемов и соображений.

Математическое содержание пакета составляют:

- итерационные процедуры, использующие методы штрафных и модифицированных штрафных функций;
- итерационные методы типа Эрроу—Гурвица;
- итерационные методы фейеровского типа;
- модифицированный симплекс-метод.

Для решения задач линейного программирования большой размерности был специально разработан механизм стыковки итерационных алгоритмов и симплекс-метода. Комплексное применение этих алгоритмов позволяет существенно ускорить решение за счет рационального использования информации о приближенном решении для формирования первоначального плана в симплекс-методе.

На рис. 8 приведена структура функционального наполнения пакета «ОКП», отражающая также и технологию прохождения на ЭВМ решаемой задачи.

Коснемся системной части пакета, выделив в ней:

- монитор (управляющая программа),
- планировщик,
- таблицы стандартных величин ТРП (таблицы распределения памяти);
- шкалы прерываний;
- «Архив» модулей.

На решение задачи составляется «Задание» с помощью специального языка, состоящего из операторов управле-

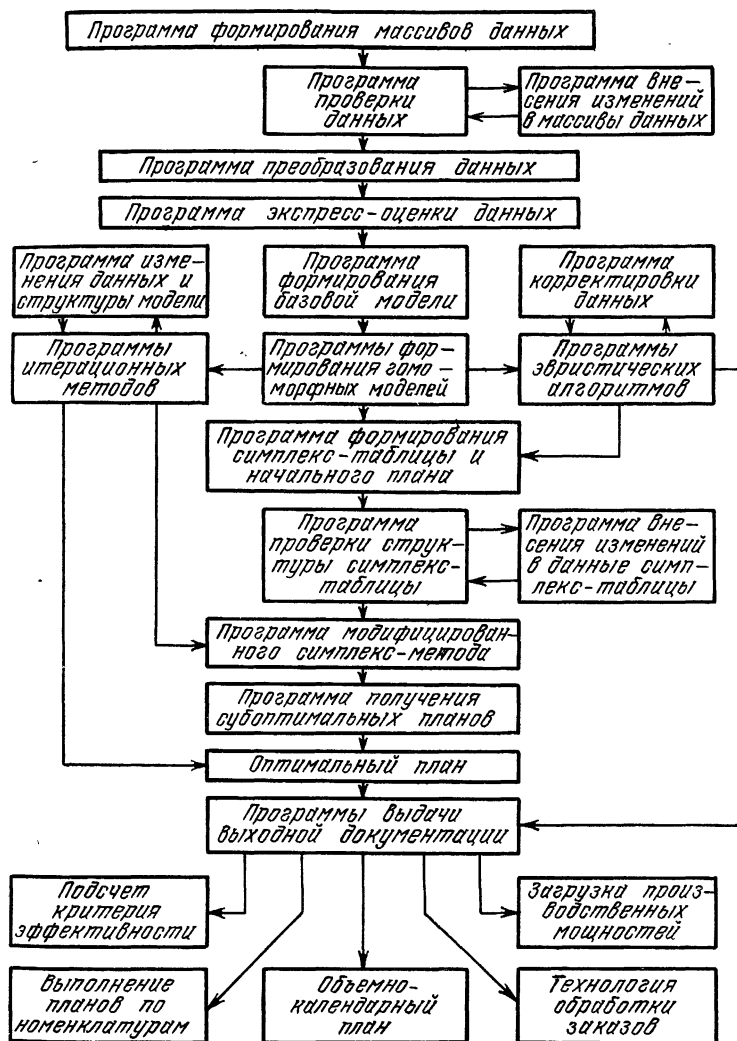


Рис. 8.

ния, которые читаются монитором последовательно. Каждый оператор «Задания» транслируется, затем планировщик формирует управляющие карты для работы с «Архивом» и транслятором с «Автокода-СОМИ»: в память вызываются необходимые модули из «Архива», транслируются и выполняются. Выполняемая программа единолично распоряжается всей доступной ей оперативной и внешней памятью ЭВМ.

Результаты работы такой программы — решение базисной задачи и заполнение соответствующей шкалы прерывания, с помощью которой изменяется по желанию последовательность выполнения управляющих операторов «Задания» в зависимости от полученных результатов вычислений (например, неограниченность целевой функции на допустимом множестве, ошибки в вычислениях и т. д.). Когда все это исполнено, монитор читает следующий оператор «Задания».

Таблица стандартных величин ТРП служит для организации интерфейсов между модулями. В ней указывается, где находится данная стандартная величина. Каждый модуль извлекает из ТРП нужные ему аргументы и по окончании работы заносит в нее информацию о месте нахождения полученных результатов.

Модули, содержащиеся в архиве, позволяют генерировать программы алгоритмов решения задач предусмотренного класса. Кроме того, архив модулей содержит сервисные и вспомогательные программы, сопровождающие решение задачи.

Пакет «ОКП» наследует многие элементы структуры и организации пакета «ОПТИМА-1» [63].

В заключение отметим, что пакет находится в рабочем состоянии и может принять вид конечной продукции после утверждения всех государственных стандартов на оформление пакетов ППП и на языки программирования.

3.2. Пакет прикладных программ распознавания образов. Пакет, именуемый «КВАЗАР», представляет собой систему программного обеспечения (на основе методов распознавания образов и многомерной статистики) достаточно широкого класса практических задач диагностики, классификации и прогнозирования. Он апробирован на многих технико-экономических, социологических, медицинских, биологических и других задачах [54, 55].

Система модулей пакета реализует алгоритмы решения следующих задач распознавания образов: задачи «обучения с учителем» (или дискриминантного анализа); задачи «обучения без учителя» (или таксономии — кластерного анализа); задачи выбора информативного пространства признаков (представления многомерных данных в пространстве меньшей размерности).

Кроме того, с помощью пакета могут решаться задачи многофакторного регрессионного анализа и ряд вспомогательных задач обработки информации:

- исключение по заданной или автоматически формирующейся маске столбцов и строк исходной матрицы данных;
- преобразование пространства R^n в пространство векторов с координатами 0 или 1;

- формирование и выдача на графопостроитель гистограмм распределения значений количественных признаков;
- обмен информацией между оперативной памятью и внешними запоминающими устройствами.

В пакете реализованы следующие алгоритмы дискриминантного анализа: алгоритм комитетного распознавания; рекуррентный алгоритм аффинного разделения выпуклых многогранных множеств; алгоритм, соответствующий методу потенциалов. Из алгоритмов таксономии в пакете функционируют: алгоритм типа «корреляционных плеяд»; алгоритм, использующий выделение максимальных совместных подсистем некоторой системы линейных неравенств и некоторые другие. Наконец, в пакет входят программы нахождения информативного признакового пространства: программа, основанная на некоторой простой эвристической процедуре отбора двоичных признаков; программа, использующая случайный поиск с адаптацией; программа построения главных компонент.

Взаимодействие программ отражает схема на рис. 9.

Остановимся вкратце на основных вопросах организации пакета. Он представляет собой систему программ, выполненных на алгоритмическом языке ФОРТРАН (версия ЦЕРН) для ЭВМ БЭСМ-6 и рассчитан на эксплуатацию в рамках мониторинговой системы «ДУБНА» и операционных систем «ДИСПАК» и «ДИАПАК».

В пакете можно выделить несколько уровней системы модулей. К первому уровню относятся модули, обращение к которым производится непосредственно из головной



Рис. 9.

программы пакета, ко второму — модули, к которым обращаются модули первого уровня, и т. д. При этом один и тот же модуль может принадлежать нескольким уровням. Кроме того, все модули пакета можно условно разделить на основные и вспомогательные. Управление работой пакета производится с помощью специальных входных инструкций.

Головная программа включает в себя таблицу входных инструкций, блок обработки заданий, блок планирования решения, модель класса решаемых задач (блок управления решением), необходимые справочные данные и некоторые вспомогательные элементы.

На рис. 10 показана схема модели класса решаемых задач. Как видно из этой схемы, модель представляет собой последовательность однотипных элементов $\{P_i \rightarrow S_i\}_1^n$. Здесь P_1, \dots, P_n — предикаты; $P_i \in \{0, 1\} (\forall i \in N_n)$. Если i -й предикат принимает значение «истина» ($P_i = 1$), то управление передается блоку S_i . Каждый блок содержит обращение к тому или иному модулю (или несколько обращений к различным модулям) в режиме динамической загрузки.

В исходном состоянии все предикаты принимают значение «ложь», т. е. $P_i = 0 (\forall i \in N_n)$. Это означает, что при прохождении последовательности от оператора «вход» к оператору «выход» ни одному из блоков S_1, \dots, S_n управление передано не будет, т. е. не будет обращений ни к одному из модулей пакета. При включении же пакета в решение конкретной задачи для некоторых $i = i_1, \dots, i_k$ будет выполнено: $P_i = 1$; при этом блоки S_{i_1}, \dots, S_{i_k} получат управление и обратятся к соответствующим модулям.

Работа головной программы отражена в функциональной схеме управляющей части пакета, изображенной на рис. 11.

Документирование пакета в связи с отсутствием единых стандартов использует ряд инструкций и положений, в частности ГОСТ 19427-74, ГОСТ 19428-74, ОСТ 25-231-74 (см. [58]).

§ 4. Задача объемно-календарного планирования машиностроительного производства

В данном параграфе описывается созданная в ИММ система математического и программного обеспечения задачи динамического объемно-календарного планирования работы крупного машиностроительного предприятия с индивидуальным и мелкосерийным характером произ-

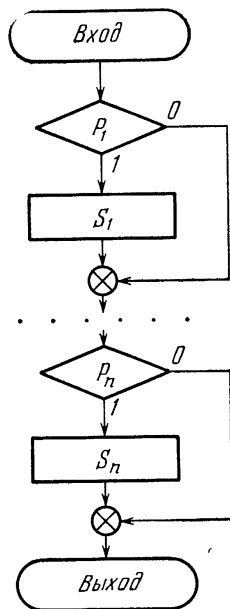


Рис. 10.



Рис. 11.

водства. Этот материал непосредственно связан с п. 3.1 «Пакет оптимального производственного планирования», где описываются элементы программного обеспечения.

Необходимо отметить принципиальную новизну описываемой системы как в теоретическом, так и в прикладном аспектах. Система внедряется на одном из конкретных крупных предприятий и является уникальной для соответствующей отрасли. Существенным достоинством системы можно считать то, что она включает развитую совокупность средств решения оптимизационных задач планирования, которые в рассматриваемой предметной области отличаются трудоемкостью.

4.1. Техничко-экономическая постановка задачи. Рассматривается задача оптимального объемно-календарного планирования деятельности крупного машиностроительно-

го предприятия с индивидуальным и мелкосерийным типом производства. Планово-учетной единицей в таком производстве является заказ. Заказы, характеризующие выпуск однородных видов оборудования (заказы доменного оборудования, заказы бурового оборудования и т. д.), образуют номенклатуры. Все заказы, которые могут быть реализованы в планируемом периоде (за планируемый период выбирается год), составляют портфель заказов. Для каждого заказа или для представителя однотипных заказов составляется технологический маршрут, который определяет последовательность обработки заказа по стадиям производства (заготовительной, обрабатывающей и сборочной стадий) с указанием месяца начала и месяца завершения изготовления заказа для каждой из стадий производства. Общее количество месяцев изготовления изделия с учетом технологического маршрута определяет суммарный производственный цикл.

На каждой стадии производственного процесса имеются виды оборудования, имеющие наибольшее использование их действительного фонда рабочего времени (лимитирующие виды оборудования), и виды оборудования, действительные фонды рабочего времени которого используются не полностью (нелимитирующие виды оборудования). Загрузка производственных мощностей лимитирующих видов оборудования определяет возможность выполнения производственной программы и системы технико-экономических показателей деятельности предприятия. С другой стороны, требуемый для выполнения плана действительный фонд рабочего времени лимитирующего оборудования зависит от времени запуска заказов в производство с учетом суммарного цикла производства.

Выпуск заказов и их реализация находят отражение в системе технико-экономических показателей деятельности предприятия, например, величина товарной реализации в денежном выражении или выполнение плана по номенклатуре в заданных единицах измерения (тыс. руб., тоннах, количестве комплектов). Можно установить такие сроки запуска и выпуска заказов, которые давали бы возможность выполнить систему технико-экономических показателей. Но это приводит к недопустимым перегрузкам производственных мощностей лимитирующих видов обо-

рудования. Если же составлять план производства с учетом допустимого использования действительного фонда, то это может привести к невыполнению системы технико-экономических показателей не только в отдельные месяцы, но и в целом по году.

Таким образом, возникает задача о составлении годового плана производства с распределением по месяцам сроков запуска заказов с учетом суммарного цикла производства так, чтобы выполнялись условия использования действительного фонда рабочего времени лимитирующих видов оборудования и технико-экономические показатели.

4.2. Особенности планирования индивидуального и мелкосерийного производства. Индивидуальный характер производства большинства видов оборудования, большая номенклатура изделий, непрерывный рост объемов производства, повышенные требования к выпускаемому оборудованию, сложность производственных связей между цехами, участками и другими подразделениями завода, жесткость директивных сроков и договорных обязательств по изготовлению и поставке продукции, большая длительность цикла изготовления сложного и уникального оборудования и другие особенности производства ставят работников плановых служб предприятия в условия, при которых они не в состоянии вручную произвести соответствующие расчеты, подготовить варианты графиков выпуска оборудования, четко увязанных по циклу изготовления с производственными мощностями завода. В этих условиях одной из основных форм повышения эффективности производства является дальнейшее совершенствование организации, управления и планирования производства на базе автоматизированной системы управления, основной частью которой должна стать система оптимального объемно-календарного планирования производства (система ОКП). Система ОКП предназначена для генерирования и оценки оптимальных планов с последующей выработкой на их основе вариантов производственных планов, которые будут иметь большее обоснование, чем при традиционном планировании. Рассматривая систему ОКП в развитии, на первом этапе ее внедрения широко используются положительные стороны традиционного метода планирования и осознаются отрицательные моменты такой организации процесса планирования с тем, чтобы выработать техно-

логию формирования производственных планов завода работниками плановых служб с помощью экономико-математических моделей, методов и ЭВМ.

Система ОКП, оперируя портфелем заказов по всем номенклатурам выпускаемых изделий, учитывая динамику его изменений и возмущения, влияющие на ход производственного процесса, распределяет запуски заказов в производство и их выпуски таким образом, чтобы наиболее эффективно загрузить наличные мощности, выполнить план в установленном объеме и номенклатуре и, обеспечив выполнение основных технико-экономических показателей плана с соблюдением технологических ограничений по циклам изготовления продукции, добиться оптимального объема выпуска продукции в денежном выражении в определенные отрезки времени (год, квартал, месяц).

Описанная постановка задачи укладывается в рамки моделей линейного программирования, и потому была создана линейная экономико-математическая модель объемно-календарного планирования. Наполнение модели конкретным содержанием привело к задаче большой размерности как по числу существенных переменных (100 000) и общему объему чисел, входящих в систему ограничений, так и по числу существенных ограничений (10 000). Решение подобной задачи затруднительно не только с вычислительной точки зрения, но в связи со сложностью, возникающей при принятии решений. Поиски путей уменьшения размерности задачи привели не только к необходимости использования агрегирования данных, но и к созданию системы взаимоувязанных экономико-математических моделей, поэтапное применение которых позволило детализировать агрегированные данные, уточнять и конкретизировать планы и принимаемые решения при переходе от одного этапа к другому.

Наиболее характерными особенностями оптимального объемно-календарного планирования производства указанного типа, которые должны учитываться при создании системы ОКП, являются:

- 1) сравнительно большой объем исходной технико-экономической информации; невозможность априорного определения наиболее существенных разделов этой информации и, в силу этого, невозможность априорной оценки складывающейся на предприятии ситуации в целом;

2) сложность отражения в одной экономико-математической модели всех существенных факторов, влияющих на ход производства, что объясняется не только многообразием факторов, но и, как правило, невозможностью для конкретно складывающейся на предприятии ситуации выделить из огромного числа влияющих на ход производства факторов те из них, которые наиболее существенно лимитируют производственную программу, поэтому оказывается возможным наличие в модели несущественно влияющих факторов;

3) наличие неформализуемых условий производства и противоречивый характер некоторых соотношений в экономико-математических моделях, когда эти соотношения должны выполняться лишь как приближенные, причем степень приближения может быть оценена лишь после получения общей картины хода производства (например, после получения оптимального плана);

4) наличие постоянных возмущений в ходе производства, что вынуждает предъявлять к плану ряд требований, связанных с устойчивостью предлагаемых рекомендаций, с необходимостью обеспечения заделов на случай сбоев производства;

5) необходимость учета эвристических предложений по составлению производственной программы, что позволяет, с одной стороны, уменьшить затраты на получение оптимального плана производства и, с другой стороны, учесть некоторые неформализованные условия производства, которые в обычных условиях могут быть оценены лишь после получения плана;

б) необходимость прогнозирования критических ситуаций в ходе производства.

4.3. Основные требования к системе ОКП. При создании системы ОКП целесообразно расчленить ее на ряд основных подсистем.

1. Информационная подсистема. Данная подсистема состоит из подсистемы ведения, исправления и контроля исходной технико-экономической информации, подсистемы формирования и выдачи выходных форм, пригодных для использования без предварительной расшифровки соответствующими службами предприятия.

2. Подсистема, реализующая методы оптимизации задач большой размерности (оптимизационная подсистема).

3. Подсистема корректировки эвристических планов, позволяющая получать комплексную оценку производственных планов, либо предлагаемых службами предприятия, либо получаемых эвристическими методами, и намечающая пути улучшения этих планов (имитационная подсистема).

Требования, налагаемые на все подсистемы, вытекают из условия взаимоувязки их входной и выходной информации, возможности принятия качественных решений при передаче информации из одной подсистемы в другую и должны обеспечивать получение оптимальных планов производства в реальном масштабе времени.

Система ОКП должна предусматривать возможность поэтапного получения оптимальных планов производства в зависимости от детализации исходной технико-экономической информации и выбранных единичных периодов планирования (месяц, квартал).

Должна быть предусмотрена возможность получения планов производства в режиме диалога человека и ЭВМ. Это требование может быть обосновано тем, что получение планов производства — процесс итерационный. В поисках компромисса требования к плану неоднократно пересматриваются (уточняются, корректируются), причем важность отдельных требований в сильной степени зависит от получаемого решения. Система должна обеспечивать, по крайней мере, автоматизированный режим диалога. Для этого необходимо органически сочетать возможность автоматической оптимизации планов производства с возможностью для пользователей частично или полностью составлять план по своему усмотрению. В последнем случае система производит только комплексную оценку ручного (эвристического) плана.

Система должна обеспечивать возможность достаточно быстро и просто изменять условия задачи и адаптировать алгоритм ее решения к полученным промежуточным результатам. Для этого в системе должно быть предусмотрено выполнение следующих операций:

- а) агрегирование условий задачи;
- б) исключение несущественных ограничений;
- в) изменение границ вариации основных параметров;
- г) вариация портфеля заказов;
- д) фиксирование значений некоторых переменных, и целый ряд других операций.

Система должна обеспечивать достоверность и наглядность процесса составления плана производства. Для этого должна быть предусмотрена возможность выдачи любого раздела исходной технико-экономической информации, выдачи на любом этапе оптимизации плана промежуточной информации и комментариев к ним, а выдача конечных результатов должна производиться в наглядной форме, не требующей расшифровки или обработки, и понятной работникам планово-производственных служб предприятия.

Должны быть предусмотрены различные варианты (по степени детализации) распечатки конечных и промежуточных результатов.

Система ОКП должна функционировать в реальном масштабе времени, т. е. должна гарантировать своевременность получения результатов расчетов по оптимизации планов. Скорость получения результатов зависит не только от мощности и быстродействия используемых ЭВМ, но и от качества сбора и ввода данных, надежности технических средств, а также от организации вычислительного процесса.

4.4. Информационная подсистема. Информационная подсистема предназначена для

а) ведения, исправления и контроля исходной, промежуточной и выходной технико-экономической информации;

б) формирования и выдачи выходных форм.

Технико-экономическая информация, которая используется в процессе оптимального планирования для удобства ее обработки, разделена на ряд логически структурированных массивов. Структура массивов для упрощения создания программ контроля по возможности должна быть одинаковой. Структуризация массивов позволяет при обнаружении ошибок указывать их характер и местонахождение.

В системе ОКП используется следующая технико-экономическая информация:

- 1) массив «портфель заказов»;
- 2) массив «действительные фонды рабочего времени и оборудования»;
- 3) массив «плановые технико-экономические показатели»;

4) массив «наименования типов оборудования и единицы измерения действительного фонда рабочего времени этого оборудования»;

5) массив «наименования и единицы измерения шифров продукции»;

6) массив «условные заказы»;

7) массив «серийные заказы»;

8) массив «комплектные заказы»;

9) массив «заделы»;

10) массив «директивные заказы»;

11) массив «приоритеты заказов»;

12) массив «директивные сроки выпуска заказов»;

13) массив «приоритеты на сроки выпуска заказов».

Информация о перечисленных массивах перфорируется с входных таблиц, либо поступает из других подсистем и записывается на внешние машинные накопители (ВМН).

Контроль массивов может осуществляться либо по мере их формирования, либо после того, как они уже сформированы.

При контроле осуществляются проверки:

1) логической структуры;

2) целочисленности задания отдельных реквизитов;

3) границы задания реквизитов;

4) последовательности следования реквизитов;

5) наличия информации по заказанным массивам, и осуществляется подсчет некоторых характеристик массивов, например, длины массива и т. д.

При контроле исходный массив может быть разделен на два подмассива. Первый подмассив содержит правильную информацию, второй подмассив содержит неправильную информацию. Такое разбиение массивов позволяет с меньшими затратами произвести корректировку, с последующим их слиянием.

Исходя из этого, для каждого массива указывается его местонахождение на внешних машинных накопителях (например, номер начальной зоны первого подмассива, номер начальной зоны второго подмассива). Проверка массивов может осуществляться как по всем массивам одновременно, так и по списку массивов или по отдельным массивам. Если при контроле исходных массивов ошибок не обнаружено, то выдается машинная табу-

ляграмма с указанием того, что ошибок в массивах нет, и с указанием местонахождения на ВМН этих массивов.

Если же ошибки обнаружены, то выдается машинная табуляграмма с указанием массива, в котором обнаружена ошибка, местонахождения массива, характера обнаруженной ошибки и ее местонахождения.

После контроля исходной технико-экономической информации необходимо произвести ее исправление. Исправления можно производить по вторым подмассивам, в которых содержатся ошибки. Исправления должны охватывать:

1) структуру массива;

2) реквизиты массива;

3) изменения по заданным структурам массива (например, если необходимо вычеркнуть какой-либо одинаковый для всех карт заказов тип оборудования, то достаточно только указать, что необходимо произвести это изменение в технологическом маршруте карты заказа).

После проведенных исправлений выдается машинная табуляграмма с указанием массива, в котором производились изменения, его местонахождения на ВМН, характера и местонахождения обнаруженной ошибки и характера проведенного исправления.

Исправленный второй подмассив объединяется с правильным первым подмассивом, и данный массив считается подготовленным к его использованию при решении задачи ОКП.

В процессе решения задачи ОКП возникает необходимость в использовании массива, который объединяет другие массивы и удобен при обработке информации. Например, имеющийся в задаче ОКП массив «фонды и ограничения» есть объединение двух массивов «действительные фонды рабочего времени оборудования» и «плановые технико-экономические показатели» со своей логической структурой.

Основное место в информационной подсистеме отводится ведению исходных технико-экономических массивов. Оно включает:

1) изменение информации в массивах с течением времени;

2) создание поколений массивов.

Процесс ведения массивов заключается в отслеживании возникающих изменений в информации и подготовке массивов с учетом произошедших изменений к моменту начала составления объемно-календарного плана.

Помимо ведения, контроля и исправления массивов требуется производить агрегирование и дезагрегирование технико-экономической информации, которые соответствуют определенным этапам решения задачи планирования. Каждому этапу составления плана соответствует свой набор массивов. Поэтому информационная подсистема должна обеспечивать сборку всех требуемых массивов по автоматическому или автоматизированному заданию на сборку.

Итак, программное обеспечение информационной подсистемы состоит из комплекса проблемно-ориентированных многоцелевых программ, реализующих различные режимы обработки исходной технико-экономической информации. Подсистема должна выполнять следующие основные операции:

- 1) ввод-вывод;
- 2) контроль исходных массивов;
- 3) исправление ошибок в массивах;
- 4) ведение массивов;
- 5) агрегирование и дезагрегирование массивов;
- 6) агрегирование и дезагрегирование исходной технико-экономической информации;
- 7) подготовка необходимых массивов в соответствии с этапом решения задачи ОКП.

Все вышеописанное относится и к обработке промежуточной и выходной технико-экономической информации.

Вторая подсистема — подсистема формирования и выдачи выходных форм — дает возможность представлять результаты решения задачи в удобном для анализа и принятия решений виде начальных, промежуточных и конечных результатов различных этапов составления производственных планов.

Большой объем представляемых для анализа и принятия решений данных, поэтапное планирование накладывают соответствующие требования к данной подсистеме.

Подсистема должна обеспечить:

- 1) документирование всех этапов составления плана;
- 2) согласование выходных форм одного этапа с входными формами других этапов (без дополнительной или ручной переработки выходных форм);
- 3) унификацию выходных форм для различных уровней агрегирования и дезагрегирования моделей;
- 4) наглядность представляемых данных;
- 5) выдачу как всего объема данных, так и отдельных его частей по требованию;
- 6) компактность;
- 7) тиражирование;
- 8) возможность выдачи выходных форм не только по данным, записанным на ВМН, но и по вводимым с перфокарт.

Данная подсистема должна обеспечивать работу имитационной подсистемы и оптимизационной подсистемы на всех этапах составления оптимального плана по моделям с первой по шестую.

Подсистема формирования и выдачи выходных форм представляет для анализа и принятия решений следующие выходные документы.

1) Объемно-календарный план деятельности предприятия:

а) годовая производственная программа с перечнем заказов. Итоговые цифры по выполнению годовых плановых технико-экономических показателей;

б) распределение годовой производственной программы по кварталам с перечнем заказов, которые выпускаются в кварталах. Итоговые цифры выполнения квартальных технико-экономических показателей и итогов выполнения годовых технико-экономических показателей;

в) распределение квартальной производственной программы по месяцам с указанием перечня заказов, которые могут быть выпущены в месяцах квартала. Итоговые цифры выполнения квартальных технико-экономических показателей;

г) распределение годовой производственной программы по месяцам с указанием перечня заказов, которые выпускаются в месяцах года. Итоговые цифры выполнения квартальных и годовых технико-экономических показателей;

2) Использование фонда рабочего времени:

а) годовая загрузка производственных мощностей с перечнем заказов, которые обрабатываются на выделенных видах оборудования;

б) квартальная загрузка производственных мощностей с перечнем заказов по каждому виду оборудования. Итоговые цифры годовой загрузки производственных мощностей;

в) месячная загрузка производственных мощностей с перечнем заказов по каждому виду оборудования. Итоговые цифры квартальной и годовой загрузки производственных мощностей.

3) Движение во времени каждого из рассматриваемых заказов по циклу производства. Указывается квартал или месяц запуска заказа в производство, кварталы и месяцы прохождения заказа по стадиям производства, виды оборудования, на которых он обрабатывается, квартал или месяц выпуска заказа, квартальный или месячный объем выпуска заказа.

4) Выполнение планового показателя по товарной реализации (тыс. руб.):

а) годовое значение;

б) квартальное с подсчетом годового значения;

в) месячное с подсчетом квартальных и годового значений.

5) Выполнение плановых показателей по номенклатурам (в тыс. руб., тоннах, количестве комплектов) — аналогично п. 4).

Кроме того, данная подсистема формирует информацию, используемую при диалоге. Часть этой информации записывается на внешние магнитные накопители и используется в автоматическом режиме другими подсистемами.

Следует отметить, что большие объемы представляемой технико-экономической информации (до 500 метров АЦПУ по основным формам) требует уделения внимания содержанию и форме представления таблиц по итогам решения. Выдача полной отчетной документации требует более сложной организации по формированию и выдаче выходных форм.

Так как одни и те же данные могут одновременно анализироваться различными плановыми службами, не-

обходимо предусмотреть тиражирование выходных форм при одноразовой работе программы формирования выходных форм.

Программное обеспечение подсистемы формирования и выдачи выходных форм должно включать следующие основные программные модули:

- 1) выдачи таблиц объемно-календарных планов;
- 2) выдачи таблицы использования фонда рабочего времени;
- 3) выдачи таблицы движения заказов по циклу производства.

Все выходные формы должны иметь варианты представления информации в кратком, сжатом и полном виде по требованию. Должна иметься возможность сборки программ на выдачу желаемого списка выходных форм. Отсутствие удобной организации в подсистеме формирования и выдачи выходных форм приводит к снижению качества обработки полученных результатов.

4.5. Оптимизационная подсистема. Исходя из принципа поэтапности формирования оптимальной производственной программы, построим ряд экономико-математических моделей с различной степенью детализации исходной технико-экономической информации и с различными единичными временными периодами.

Модель статической задачи годового планирования (модель I). Найти

$$\max \sum_{q \in Q} c_q x_q \quad (6.4.1)$$

при условиях

$$\sum_{q \in Q} a_{iq} x_q \leq F_i \quad (\forall i \in N_n), \quad (6.4.2)$$

$$\sum_{q \in Q_j} r_q x_q \geq R_j \quad (\forall j \in N_m), \quad (6.4.3)$$

$$a_q \leq x_q \leq A_q \quad (\forall q \in N_g), \quad (6.4.4)$$

где

$I = N_n$ — множество различных типов оборудования;

$J = N_m$ — множество номенклатур;

$Q = N_g$ — портфель заказов;

a_{iq} — норма затрат i -го типа оборудования на изготовление одного комплекта q -го заказа;

r_q — вклад q -го заказа в выполнение планового показателя j -й номенклатуры (в соответствующей единице измерения — тыс. руб., тонны, количество комплектов);

c_q — цена одного комплекта q -го заказа;

F_i — годовой действительный фонд рабочего времени i -го вида оборудования;

R_j — годовой плановый показатель для j -й номенклатуры (в соответствующей единице измерения);

a_q — нижняя граница объема выпуска изделий q -го заказа;

A_q — верхняя граница объема выпуска изделий q -го заказа;

x_q — объем изделий q -го заказа, который может быть реализован.

Функция цели (6.4.1) — суммарная стоимость произведенной продукции в денежном выражении за год; ограничения (6.4.2) обеспечивают загрузку оборудования в заданных пределах, ограничения (6.4.3) требуют выполнения плановых годовых заданий по каждой номенклатуре в натуральных или денежных единицах, ограничения (6.4.4) устанавливают границы вариации объемов производства изделий по каждому заказу в течение года.

Оптимальный план задачи (6.4.1) — (6.4.4) обозначим через $\{\tilde{x}_q\}$. Он позволяет оценить в первую очередь объемы выпуска продукции в целом и по каждой номенклатуре

$$\tilde{C} = \sum_{q \in Q} c_q \tilde{x}_q, \quad \tilde{R}_j = \sum_{q \in Q_j} r_q \tilde{x}_q \quad (\forall j \in N_m),$$

а также степень загрузки основных видов оборудования

$$F_i - \sum_{q \in Q} a_{iq} \tilde{x}_q \quad (\forall i \in N_n).$$

Симплекс-матрица этой задачи автоматически формируется из исходной технико-экономической информации путем ее агрегирования. Размеры этой задачи сравнительно невелики, и поэтому она может быть решена на ЭВМ средней мощности.

В ходе решения устраняются существенные невязки в исходной технико-экономической информации, достигается соответствие между годовым планом производства и существующими мощностями, обеспечивается конкретными заказами выполнение плана по отдельным номенкла-

турам, формируется такая производственная программа, которая позволяет наилучшим образом обеспечить выполнение экономических показателей плана. Для решения этих вопросов должна быть предусмотрена возможность проведения серии решений задачи (6.4.1)—(6.4.4), когда после каждого решения оперативно производится его анализ и исходная информация корректируется в автоматизированном режиме.

Оптимальный план статической задачи позволяет выделить те виды оборудования, которые будут существенно лимитировать годовую производственную программу, а также те виды оборудования, которые в конкретной ситуации не влияют на оптимальность плана и могут быть исключены из дальнейшего рассмотрения (последующей конкретизации планов).

Полученные величины (\tilde{C} — годовой объем выпуска продукции и \tilde{R}_j — выпуски продукции по номенклатурам) являются верхними оценками соответствующих показателей для следующих этапов составления производственных планов. Кроме этого, решение двойственной задачи позволяет установить некоторую приоритетность заказов, вытекающую из оценок технологических способов.

Модель квартального планирования (модель II). Найти

$$\max \sum_{q \in Q} \sum_{t=1}^4 d_q^t z_q^t \quad (6.4.5)$$

при условиях

$$\sum_{q \in Q} \sum_{t=1}^4 b_{iq}^v z_q^t \leq F_i^v \quad (\forall i \in N_n, \forall v \in N_4), \quad (6.4.6)$$

$$\sum_{q \in Q_j} r_q^t z_q^t \geq R_j^t \quad (\forall j \in N_m, \forall t \in N_4), \quad (6.4.7)$$

$$\sum_{t=1}^4 z_q^t \leq A_q \quad (\forall q \in N_g), \quad (6.4.8)$$

$$\gamma_q^t \leq z_q^t \leq \delta_q^t \quad (\forall q \in N_g, \forall t \in N_4), \quad (6.4.9)$$

где

v — текущий квартал производства;

t — квартал, в котором возможен выпуск продукции;

b_{iq}^{vt} — норма затрат i -го типа оборудования на изготовление одного комплекта q -го заказа в квартале v при условии, что изделие будет выпущено в квартале t ;

r_q^t — вклад q -го заказа в выполнение планового показателя j -й номенклатуры в квартале t (в соответствующей единице измерения — тыс. руб., тонны, количество комплектов);

d_q^t — цена одного комплекта q -го заказа;

F_i^v — действительный фонд рабочего времени i -го типа оборудования в квартале v ;

R_j^t — плановый показатель для j -й номенклатуры в t квартале (в соответствующей единице измерения);

γ_q^t — нижняя граница объема выпуска изделий q -го заказа в квартале t ;

δ_q^t — верхняя граница объема выпуска изделий q -го заказа в квартале t ;

z_q^t — объем изделий q -го заказа, который может быть реализован в квартале t .

Функция цели (6.4.5) — суммарная стоимость произведенной продукции в денежном выражении за год; ограничения (6.4.6) обеспечивают квартальную загрузку каждого вида оборудования в заданных пределах, ограничения (6.4.7) обеспечивают выполнение квартальных плановых показателей по объемам производства продукции каждой номенклатуры. Условия (6.4.8) и (6.4.9) накладывают ограничения на объемы изготовления изделий по каждому заказу.

Оптимальный план задачи (6.4.5) — (6.4.9) $\{\tilde{z}_q^t\}$ позволяет оценить объемы выпуска продукции в каждом квартале в денежном выражении и по отдельным номенклатурам

$$\bar{D}^t = \sum_{q \in Q} d_q^t \tilde{z}_q^t, \quad \bar{R}_j^t = \sum_{q \in Q_j} r_q^t \tilde{z}_q^t \quad (\forall j \in N_m, \forall t \in N_4)$$

и степень загрузки основных видов оборудования по кварталам

$$F_i^v = \sum_{q \in Q} \sum_{t=1}^4 b_{iq}^{vt} \tilde{z}_q^t \quad (\forall i \in N_n, \forall v \in N_4).$$

Оптимальное решение задачи (6.4.5) — (6.4.9) не только конкретизирует годовую производственную программу,

но и разбивает ее оптимальным образом на квартальные задания. При этом выполняются следующие соотношения:

$$\bar{C} \geq \sum_{t=1}^4 D^t, \quad \bar{R}_j \geq \sum_{t=1}^4 R_j^t \quad (\forall j \in N_m).$$

Происходит дальнейшее уточнение плановых заданий. Полученные производственные программы для каждого квартала устанавливают соответствие между производственными мощностями и выполнением директивных плановых показателей как в денежном, так и в натуральном выражении.

Оптимальный план позволяет выделить для ряда видов оборудования те периоды, в которые эти виды не лимитируют производственную программу (например, как правило, виды оборудования заготовительной стадии производства не лимитируют производственную программу в четвертом квартале), а для видов оборудования, существенно лимитирующих производственную программу, выделить те единичные периоды, на которые приходится наибольшие нагрузки (перегрузки).

Для каждого квартала t формируется список заказов Q^t , которые либо запущены в производство в этом квартале, либо находятся в стадии производства, либо их производство завершается в этом квартале. Для каждого из таких заказов $q \in Q^t$ автоматически формируется технологический маршрут его прохождения по циклу производства в данном квартале с месячной разбивкой. Обозначим количество изделий q -го заказа, проходящих обработку в t -м квартале, но производство которых не завершается через u_q^t , а через v_q^t — те, чье производство завершается. В этом случае для оптимального распределения квартального плана производства на месячные задания составим следующую модель.

Модель внутриквартального планирования (модель III). Для фиксированного квартала t найти

$$\max \sum_{q \in Q^t} \sum_{s \in S^t} \omega_q u_q^s \quad (6.4.10)$$

при условиях

$$\sum_{s \in S^t} \sum_{q \in Q^t} [a_{iq}^{\tau s} u_q^s + \alpha_{iq}^{\tau s} v_q^s] \leq \Phi_i^{\tau} \quad (\forall i \in N_n, \forall \tau \in S^t), \quad (6.4.11)$$

$$\sum_{s \in S^t} v_q^s = z_q^t \quad (\forall q \in Q^t), \quad (6.4.12)$$

$$\alpha_q^s \leq u_q^s \leq \beta_q^s \quad (\forall q \in Q^t, \forall s \in S^t), \quad (6.4.13)$$

$$v_q^s \geq 0 \quad (\forall q \in Q^t, \forall s \in S^t), \quad (6.4.14)$$

где

S^t — множество месяцев выпуска заказов, находящихся в производстве в течение квартала t ;

τ — текущий месяц производства в t квартале;

s — месяц квартала t , в котором возможен выпуск продукции;

$\alpha_{iq}^{\tau s}$ — норма затрат i -го типа оборудования на изготовление одного комплекта q -го заказа в месяце τ квартала t при условии, что изделие будет реализовано в s месяце этого же квартала;

Φ_i^{τ} — действительный фонд рабочего времени i -го типа оборудования в месяце τ квартала t ;

α_q^s — нижняя граница объема выпуска q -го заказа, который может быть реализован в месяце $s \in S^t$;

β_q^s — верхняя граница объема выпуска q -го заказа, который может быть реализован в месяце $s \in S^t$;

ω_q — оценка q -го заказа, устанавливаемая, исходя из приоритета данного заказа.

Функция цели (6.4.10) обеспечивает прохождение наиболее важных заказов по циклу производства; условия (6.4.11) обеспечивают загрузку оборудования в заданных пределах, условия (6.4.12) гарантируют выполнение квартальных плановых показателей как в денежном, так и в натуральном выражениях. Условия (6.4.13) и (6.4.14) ограничивают вариацию количества изделий заказов, проходящих по циклу производства.

Задача (6.4.10)–(6.4.14) имеет сравнительно небольшую размерность и может быть достаточно эффективно решена для каждого квартала. После решения этих задач можно путем композиции составить годовой производственный план с месячной разбивкой, который является достаточно хорошим приближением к оптимуму задачи объемно-календарного планирования.

Модель годового планирования с разбивкой по месяцам (модель IV). Найти

$$\max \sum_{q \in Q} \sum_{s=1}^T c_q^s y_q^s \quad (6.4.15)$$

при условиях

$$\sum_{q \in Q} \sum_{s=1}^T a_{iq}^{\tau s} y_q^s \leq \Phi_i^{\tau} \quad (\forall i \in N_n, \forall s \in N_{12}), \quad (6.4.16)$$

$$\sum_{q \in Q_j} \sum_{s \in S^t} r_q^s y_q^s \geq R_j^t \quad (\forall j \in N_m, \forall t \in N_4), \quad (6.4.17)$$

$$\sum_{s=1}^T y_q^s \leq A_q \quad (\forall q \in N_g), \quad (6.4.18)$$

$$\alpha_q^s \leq y_q^s \leq \beta_q^s \quad (\forall q \in N_g, \forall s \in N_{12}), \quad (6.4.19)$$

где

T — планируемый период (год);

t — квартал года;

s — месяц, в котором возможен выпуск продукции;

τ — текущий месяц производства;

c_q^s — цена одного комплекта q -го заказа при условии, что изделие будет реализовано в месяце s ;

$a_{iq}^{\tau s}$ — норма затрат i -го типа оборудования на изготовление одного комплекта q -го заказа в месяце τ при условии, что изделие будет реализовано в месяце s ;

r_q^s — вклад q -го заказа в выполнение планового показателя j -й номенклатуры в квартале t при условии, что изделие будет реализовано в месяце s ;

Φ_i^{τ} — действительный фонд рабочего времени i -го типа оборудования в месяце τ ;

R_j^t — плановый показатель для j -й номенклатуры в квартале t (в соответствующей единице измерения);

A_q — допустимый объем реализации q -го заказа;

α_q^s — нижняя граница объема выпуска изделий q -го заказа в месяце s ;

β_q^s — верхняя граница объема выпуска изделий q -го заказа в месяце s ;

y_q^s — объем изделий q -го заказа, который может быть реализован в месяце s .

Функция цели (6.4.15) — суммарная стоимость производственной продукции за год; ограничения (6.4.16) обе-

спечивают месячную загрузку каждого вида оборудования в заданных пределах, ограничения (6.4.17) обеспечивают поквартальное выполнение плановых показателей производства продукции по каждой номенклатуре. Условия (6.4.18), (6.4.19) накладывают ограничения на количество изготавливаемых изделий каждого заказа в целом по году и по каждому месяцу.

Оптимальный план задачи (6.4.15)—(6.4.19) обозначим $\{\tilde{y}_q^s\}$. Он позволяет оценить объемы выпуска продукции в течение года в целом, а по номенклатурам — поквартально:

$$\tilde{C} = \sum_{s \in S^t} \sum_{q \in Q} c_q^s \tilde{y}_q^s, \quad \tilde{R}_j^t = \sum_{q \in Q_j} \sum_{s \in S^t} r_{qj}^s \tilde{y}_q^s \quad (\forall j \in N_m, \forall t \in N_4),$$

а также объем загрузки основных видов оборудования по месяцам периода планирования:

$$\Phi_i^t = \sum_{q \in Q} \sum_{s=1}^T a_{iq}^{ts} \tilde{y}_q^s \quad (\forall i \in N_n, \forall t \in N_{12}).$$

Экономико-математическая модель IV является основной при оптимизации объемно-календарных планов производства, и хотя она формализует не все основные условия производства, ее размерность сравнительно велика. Так, число существенных ограничений (6.4.15)—(6.4.19) достигает 3000, число основных переменных y_q^s — свыше 10 000, а симплекс-матрица ее содержит около полумиллиона ненулевых элементов. Для решения такой задачи классическими методами оптимизации требуется ЭВМ достаточно высокого класса. Поэтому необходимо информацию, полученную на предыдущих этапах оптимизации по укрупненным показателям, использовать для существенного уменьшения размерности решаемой задачи. Это достигается, в первую очередь, за счет следующих операций, выполняемых в автоматизированном режиме:

1) вычеркивание ограничений по использованию фонда рабочего времени видов оборудования, не лимитирующих в данной производственной ситуации план производства;

2) агрегирование ряда ограничений по использованию фонда рабочего времени частично лимитирующих производство видов оборудования;

3) уменьшение интервалов варьирования возможных сроков изготовления изделий ряда заказов;

4) фиксирование значений некоторых основных переменных, что соответствует либо однозначному выбору периода изготовления соответствующих заказов, либо невозможности изготовления заказов в данные периоды.

Эти и другие приемы позволяют существенно уменьшить размерность решаемой задачи, что гарантирует получение оптимального плана за приемлемое время. Однако после получения оптимального плана необходимо проверить выполнение всех исключенных ограничений и, в случае обнаружения несовместности, переформировать задачу, включив несовместные ограничения, и решить задачу заново с использованием заданного начального плана.

Другим эффективным путем сокращения размерности решаемых задач является использование идеи пономенклатурной оптимизации. В этом случае экономико-математическая модель формулируется как модель двухуровневой оптимизации.

Модель пономенклатурной оптимизации (модель V). Все множество заказов Q разбивается на множество заказов каждой номенклатуры Q_j ($\forall j \in N_m$). Для каждой номенклатуры сформулируем оптимизационную задачу: найти

$$\max \sum_{s=1}^T \sum_{q \in Q_j} c_q^s y_q^s \quad (6.4.20)$$

при условиях

$$\sum_{s=1}^T \sum_{q \in Q_j} a_{iq}^{\tau s} y_q^s \leq \varphi_{ij}^{\tau} \quad (\forall i \in N_n, \forall j \in N_m, \forall \tau \in N_{12}), \quad (6.4.21)$$

$$\sum_{s \in S^t} \sum_{q \in Q_j} r_q^s y_q^s \geq R_j^t \quad (\forall j \in N_m, \forall t \in N_4), \quad (6.4.22)$$

$$\sum_{s=1}^T y_q^s \leq A_q \quad (\forall q \in Q_j), \quad (6.4.23)$$

$$\alpha_q^s \leq y_q^s \leq \beta_q^s \quad (\forall q \in Q_j, \forall s \in N_{12}). \quad (6.4.24)$$

Здесь φ_{ij}^{τ} — фонд рабочего времени i -го вида оборудования, выделяемого в τ -м месяце на обработку изделий j -й номенклатуры. Таким образом, формируется m задач вто-

рого уровня планирования. Тогда задача первого уровня (задача центра) формулируется как задача распределения ресурсов (фондов рабочего времени) между производством продукции различных номенклатур: найти

$$\max \sum_{\tau=1}^T \sum_{i=1}^n \sum_{j=1}^m \eta_{ij}^{\tau} \Phi_{ij}^{\tau} \quad (6.4.25)$$

при условиях

$$\sum_{j=1}^m \Phi_{ij}^{\tau} \leq \Phi_i^{\tau} \quad (\forall \tau \in N_{12}, \forall i \in N_n), \quad (6.4.26)$$

$$\gamma_{ij}^{\tau} \leq \Phi_{ij}^{\tau} \leq \delta_{ij}^{\tau} \quad (\forall \tau \in N_{12}, \forall i \in N_n, \forall j \in N_m), \quad (6.4.27)$$

где η_{ij}^{τ} — оптимальные двойственные оценки выделенных фондов рабочего времени в задаче (6.4.20) — (6.2.24).

В частности, ввиду сравнительно небольшого количества серийных, групповых и условных заказов (множество заказов Q^*), можно сформулировать оптимизационную задачу по предварительному оптимальному распределению их производства.

Модель оптимизации производства серийной продукции (модель VI). Найти

$$\min \sum_{i=1}^n f_i \eta_i \quad (6.4.28)$$

при условиях

$$\sum_{s=1}^T \sum_{q \in Q^*} a_{iq}^{\tau s} y_q^s \leq f_i^{\tau} \quad (\forall \tau \in N_{12}, \forall i \in N_n), \quad (6.4.29)$$

$$\sum_{s \in S^t} \sum_{q \in Q_j^*} r_q^s y_q^s \geq R_j^t \quad (\forall t \in N_4, \forall j \in N_m), \quad (6.4.30)$$

$$\sum_{s=1}^T y_q^s \leq A_q \quad (\forall q \in Q^*), \quad (6.4.31)$$

$$\alpha_q^s \leq y_q^s \leq \beta_q^s \quad (\forall q \in Q^*, \forall s \in N_{12}), \quad (6.4.32)$$

$$0 \leq f_i^{\tau} \leq f_i \quad (\forall \tau \in N_{12}, \forall i \in N_n). \quad (6.4.33)$$

Здесь f_i^{τ} — переменная, характеризующая выделенные фонды рабочего времени i -го вида оборудования в τ -й месяц для изготовления серийной продукции; f_i — максимальная из f_i^{τ} по τ ; η_i — стоимость единицы фонда рабочего времени i -го вида оборудования (в качестве η_i

могут использоваться двойственные оценки статической задачи). Функция цели (6.4.28) обеспечивает равномерность загрузки лимитирующих видов оборудования в течение периода планирования. В случае получения удовлетворительных результатов, оптимальный план задачи (6.4.28)—(6.4.33) может быть использован для уменьшения размерности основной задачи (6.4.15)—(6.4.19).

Большие размеры задачи не только затрудняют процесс оптимизации, но и, в силу использования больших объемов исходной технико-экономической информации, не позволяют априорно оценить ситуацию в целом, что ведет к наличию множества ошибок в задании ключевых позиций этой информации. Поэтому до решения задачи (6.4.15)—(6.4.19) необходимо произвести экспресс-анализ всей исходной информации, устранить существенные ошибки и обеспечить проведение всестороннего анализа эвристических объемно-календарных планов предлагаемыми плановыми службами предприятия.

Так как размерность задачи (6.4.1)—(6.4.4) существенно меньше размерности задачи (6.4.15)—(6.4.19), то становится возможным сравнительно быстро решить ее необходимое число раз, последовательно устраняя существенные ошибки в исходной технико-экономической информации.

Решение этой задачи позволяет оценить возможность выполнения плана предприятием в целом (в денежном выражении), выполнение плана по отдельным номенклатурам (в натуральном выражении), выделить существенно лимитирующие план виды оборудования, а также виды оборудования, загрузка которых на данном этапе планирования несущественно влияет на оптимальность плана. Кроме этого, оптимальный план статической задачи устанавливает некоторую приоритетность заказов, вытекающую из оценок технологических способов.

На начальной стадии разработки и внедрения системы ОКП, учитывая непрерывность производственного процесса, для получения оптимальных объемно-календарных планов необходимо не только зафиксировать состояние производства на начало планирования, но и определить наиболее рациональные заделы производства, которые необходимо создать в предшествующие периоды для того, чтобы успешно выполнить план производства в плани-

руемом периоде. Для этого нужно определить выделяемые фонды рабочего времени оборудования в период изготовления заделов и указать перечень заказов, по которым целесообразно изготовление заделов. Для чего формируется соответствующая экономико-математическая модель.

Для автоматизации режима диалога человека и ЭВМ подсистема выполняет следующие операции:

- 1) автоматически формирует из исходной информации симплекс-таблицу любой из задач оптимизации;
- 2) формирует приближенное начальное решение задач эвристическими методами;
- 3) обеспечивает возможность продолжения поиска оптимального плана, начиная с некоторого приближенного решения, не обязательно опорного в любой из перечисленных задач;
- 4) переформирует задачу в автоматизированном режиме, используя информацию о ее приближенных решениях.

4.6. Имитационная подсистема. Основными задачами, решаемыми в имитационной подсистеме, являются следующие.

1. Оценка эвристических планов, предлагаемых различными плановыми службами предприятия или министерства, причем планы могут быть любого уровня: годовые, квартальные, месячные статического и динамического вида.

2. Оценка достаточно широкого спектра предложений по формированию производственной программы, в том числе оценка влияния директивных указаний по изменению плана производства на качество выполнения плановых показателей.

3. Оценка и реализация эвристических приемов составления планов производства.

4. Получение начальных планов производства и существенное сокращение размерности решаемых задач в оптимизационной подсистеме для получения планов производства в реальном масштабе времени.

5. Оценка ряда неформализованных условий производства, влияние которых может быть сравнительно просто оценено имеющейся в наличии информацией, в их числе:

- а) комплектность выпуска ряда заказов;

б) допустимость перегруза рабочего времени некоторых видов оборудования (также перегрузы могут быть ликвидированы с помощью способов, не учитываемых в моделях: кооперация, взаимозаменяемое оборудование и др.);

в) обеспеченность производственной программы лимитирующими ресурсами;

г) выполнение ограничений по использованию фондов рабочего времени видов оборудования, не учитываемых в моделях, и целый ряд других условий.

6. Осуществление автоматизированного режима проведения диалога человека и ЭВМ при составлении объемно-календарных планов производства, что позволяет активно и целенаправленно использовать опыт и знания работников плановых служб предприятия.

Отсутствие блока оптимизации в этой подсистеме позволяет реализовать ее на ЭВМ среднего класса. Подсистема может включать ряд эвристических алгоритмов улучшения предлагаемых планов производства, которые также должны реализовываться в режиме диалога.

В соответствии с принципом поэтапности составления планов производства имитационная подсистема разбивается на ряд основных блоков (уровней).

Блок оценки статических планов производства (блок I). Предлагаемый для оценки план производства задается либо в виде количества изделий, выпускаемых в течение года по всем заказам $\{x_q\}$ ($\forall q \in Q$), либо по некоторой части заказов. В этом случае, в первую очередь, производится оценка выполнения следующих ограничений:

$$\sum_{q \in Q} c_q x_q \geq C, \quad (6.4.34)$$

$$\sum_{q \in Q} a_{iq} x_q \leq F_i \quad (\forall i \in N_n), \quad (6.4.35)$$

$$\sum_{q \in Q_j} r_q x_q \geq R_j \quad (\forall j \in N_m), \quad (6.4.36)$$

$$a_q \leq x_q \leq A_q \quad (\forall q \in N_g). \quad (6.4.37)$$

В случае отсутствия каких-либо рекомендаций по составлению плана производства формируется начальное приближение

$$x_q = A_q \quad (\forall q \in N_g).$$

В результате подстановки $\{x_q\}$ в (6.4.34) — (6.4.36) выдается следующая технико-экономическая информация.

I. Выполнение годовых плановых технико-экономических показателей в принятых единицах измерения:

1) выполнение годового плана по выпуску товарной продукции в денежном выражении, в том числе вклад каждой номенклатуры в выполнение этого показателя;

2) выполнение плана выпуска продукции по каждой из номенклатур в натуральных единицах, в том числе:

а) выделяется множество номенклатур, объем производства по которым больше или равен R_j ;

б) выделяется множество номенклатур, объем производства по которым меньше R_j (т. е. имеет место невыполнение плана выпуска продукции), и указывается объем невыполнения;

в) указывается, в каких единицах измерения и на какую величину плановые показатели по каждой номенклатуре выполняются, перевыполняются или невыполняются.

II. Использование годового фонда рабочего времени каждого учитываемого в системе вида оборудования:

1) выделяется множество видов оборудования, для которых требуемый на выполнение производственной программы фонд рабочего времени превышает или равен плановому F_i (имеет место перегруз оборудования), определяется объем превышения $\sum_{q \in Q} a_{iq}x_q - F_i \geq 0$ ($\forall i \in N_n$) и производится оценка этого превышения;

2) выделяется множество видов оборудования, для которых требуемый на выполнение производственной программы фонд рабочего времени меньше планового F_i (имеет место недогруз оборудования), определяется объем неиспользования годового фонда рабочего времени

$$F_i - \sum_{q \in Q} a_{iq}x_q > 0 \quad (\forall i \in N_n);$$

3) из множества недогруженных видов оборудования при оцениваемой производственной программе выделяются те, которые несущественно ее лимитируют, т. е. которые можно исключить из дальнейшего рассмотрения, а пример, те из них, для которых выполняется

соотношение

$$\frac{\sum_{q \in Q} a_{iq} x_q}{F_i} \cdot 100\% \leq 50\%;$$

4) для всех видов оборудований выдаются списки заказов, проходящих на них обработку с указанием требуемого фонда рабочего времени.

III. Выполнение ограничений на объемы выпуска изделий по каждому заказу:

1) выдается список заказов, для которых $a_q \leq x_q \leq A_q$;

2) выдается список заказов, для которых $x_q > A_q$, и производится оценка возможности перевыполнения;

3) выдается список заказов, для которых $x_q < A_q$, и оценивается невыполнение выпуска изделий по каждому из таких заказов.

IV. Выполнение ряда неформализованных условий производства, например комплектность выпуска заказов.

После получения оценок плана $\{x_q\}$ и их анализа, реализуя режим диалога, подсистема должна обеспечивать возможность проведения следующих операций в автоматизированном режиме для улучшения плана:

1) аннулировать из рассмотрения часть заказов;

2) ввести в рассмотрение новое множество заказов;

3) изменить объемы выпуска заказов, входящих в исходный план $\{x_q\}$;

4) изменить значения основных технико-экономических показателей C , R_j , F_i , определяющих план производства;

5) изменить границы вариации объемов производства отдельных заказов, что особенно важно для условных заказов.

В случае наличия приоритета заказов план должен формироваться по этапам, путем ввода в рассмотрение сначала самых приоритетных заказов, затем все менее и менее важных. Приоритет заказов может быть составлен как исходя из директивных указаний, так и на основе использования оптимальных решений двойственной статической задачи.

В случае получения удовлетворительного статического годового плана (путем согласования решений оптимизационной подсистемы и имитационной подсистемы) пере-

ходим к оценке следующих этапов составления планов производства.

Блок оценки годовых планов с поквартальной разбивкой (блок II). План $\{x_q\}$, полученный на предыдущем этапе, является основой составления начального плана на данном уровне. Начальный план должен быть задан в виде количества изделий, выпускаемых в каждом квартале, по всем заказам, входящим в имеющийся портфель заказов, т. е. в виде набора $\{z_q^t\}$ ($\forall q \in N_g, \forall t \in N_4$).

В случае отсутствия рекомендаций по полному формированию плана производства (т. е. при наличии лишь части информации о производственной программе) блок доформирует начальный план, учитывая циклы изготовления и целочисленность выпуска изделий, по следующему правилу:

$$z_q^t = x_q / T_q,$$

где T_q — количество единичных периодов (кварталов), в которые возможен выпуск изделий заказа q , а индекс t принимает значения, соответствующие этим кварталам. В случае неоднозначности принятия решений целесообразно использовать имеющиеся приоритеты, более приоритетные заказы сдвигают на более ранние сроки изготовления. В случае отсутствия или равенства приоритетов добиваются равномерности выпуска номенклатурной продукции.

В первую очередь оценивается следующая система неравенств:

$$\sum_{q \in Q} \sum_{t=1}^4 b_{iq}^t z_q^t \leq F_i^y \quad (\forall y \in N_4, \forall i \in N_n), \quad (6.4.38)$$

$$\sum_{q \in Q} \sum_{t=1}^4 d_q^t z_q^t \geq D, \quad (6.4.39)$$

$$\sum_{q \in Q_j} r_q^t z_q^t \geq R_j^t \quad (\forall t \in N_4, \forall j \in N_m), \quad (6.4.40)$$

$$\sum_{t=1}^4 z_q^t \leq A_q \quad (\forall q \in N_g), \quad (6.4.41)$$

$$\gamma_q^t \leq z_q^t \leq \delta_q^t \quad (\forall t \in N_4, \forall q \in N_g). \quad (6.4.42)$$

В результате анализа этой системы выдается следующая технико-экономическая информация о плане производства.

I. Выполнение годовых и квартальных технико-экономических показателей плана в принятых единицах измерения:

1) оценивается выполнение годового и квартальных планов по выпуску товарной продукции в денежном выражении, в том числе оценивается вклад каждой номенклатуры в выполнение этих показателей;

2) оценивается выполнение квартальных и годового планов выпуска продукции по каждой номенклатуре в натуральных единицах, выдается следующая информация:

а) перечень номенклатур, план производства по которым выполнен как в каждом из кварталов, так и за год;

б) перечень номенклатур, годовой план выпуска продукции по которым не выполнен, с указанием конкретных кварталов, в которые произошло невыполнение, так и объемов невыполнения;

в) перечень номенклатур с неравномерным выпуском продукции, т. е. годовой план по которым выполнен, но имеются кварталы, в которых имеет место невыполнение квартальных планов, с указанием степени неравномерности выпуска продукции;

3) по каждой номенклатуре выдается перечень заказов, выпускаемых в каждом квартале.

II. Использование фонда рабочего времени каждого учитываемого в системе вида оборудования:

1) выделяются множества видов оборудования, по которым имеет место перегрузка фонда рабочего времени для выполнения производственной программы, хотя бы в одном квартале, указываются конкретные кварталы, в которых наблюдается перегрузка, определяется объем перегрузки

$$\sum_{q \in Q} \sum_{t=1}^4 b_{iq}^{yt} z_q^t - F_i^y > 0$$

и производится оценка этих превышений;

2) выделяется множество видов оборудования, для которых имеет место недоиспользование фонда рабочего

времени в отдельные кварталы

$$F\gamma - \sum_{q \in Q} \sum_{t=1}^4 b_{iq}^{\gamma t} z_q^t > 0,$$

и производится оценка величин недоиспользованного фонда с целью выделения и последующего исключения этих неравенств;

3) выделяются виды оборудования, для которых имеет место неравномерность использования фонда рабочего времени, которая не может быть объяснена особенностями производства, т. е. тех видов оборудования, у которых недогрузка сочетается с перегрузкой;

4) по каждому виду оборудования выдается перечень заказов, проходящих обработку на нем в каждом квартале, с указанием объемов использованных фондов.

III. Выполнение ограничений на объемы выпуска изделий по каждому заказу:

1) выдается список заказов, для которых превышаются годовые объемы производства, т. е. для которых

$$\sum_{t=1}^4 z_q^t > A_q;$$

2) оценивается равномерность выпуска заказов по кварталам с выдачей следующей информации:

а) списка заказов, которые выпускаются равномерно, т. е. для которых выполняются соотношения

$$\gamma_q^t \leq z_q^t \leq \delta_q^t \quad (\forall q \in N_g, \forall t \in N_4);$$

б) списка заказов, для которых нарушены нижние границы выпуска в отдельные кварталы $z_q^t < \gamma_q^t$, с указанием как индексов соответствующих кварталов, так и объемов невыполнения;

в) списка заказов, для которых $z_q^t > \delta_q^t$, с указанием соответствующей информации;

3) выдаются графики движения по циклу производства каждого из рассматриваемых заказов;

IV. Выполнение ряда неформализованных условий производства:

1) выполнение директивных сроков выпуска заказов;

2) комплектность выпуска заказов;

3) учет приоритета выпуска заказов.

4) равномерность загрузки лимитирующих видов оборудования по кварталам, и ряд других условий.

После получения оценок плана $\{z_q^t\}$ и их анализа, реализуя режим диалога, подсистема для частичного улучшения плана производит следующие операции:

1) аннулирует из производственной программы часть заказов и вводит в рассмотрение новые заказы;

2) изменяет объемы выпуска заказов в целом по году и по отдельным кварталам;

3) перераспределяет выпуск заказов по кварталам;

4) изменяет значения основных технико-экономических показателей для D , R_j^t , F_i^y , определяющих план производства.

Затем подсистема проверяет соответствие технико-экономических показателей планов $\{x_q\}$ и $\{z_q^t\}$, а также выполнение ограничений по использованию фондов рабочего времени нелимитирующих видов оборудования. При обнаружении несоответствия планы должны корректироваться.

В случае получения удовлетворительных квартальных планов производства (путем согласования решений оптимизационной подсистемы и имитационной подсистемы) переходим к следующему этапу составления плана производства.

Блок оценки годовых планов производства с месячной разбивкой (Блок III). План $\{z_q^t\}$, полученный на предыдущем этапе, является основой для составления начального плана данного уровня, который задается в виде набора количества изделий, выпускаемых по каждому заказу каждый месяц, т. е. в виде

$$\{y_q^s\} \quad (\forall q \in N_g, \forall s \in N_{12}).$$

Эвристическая информация о плане производства может задаваться в следующих объемах:

1) полностью определен весь план производства, т. е. задан весь набор $\{y_q^s\}$ ($\forall q \in N_g, \forall s \in N_{12}$);

2) план производства задан частично, т. е. зафиксированы некоторые величины y_q^s , $q \in Q^0 \subset Q$, $t \in T_0$; например, задан план выпуска серийной продукции;

3) существенно уменьшена возможность вариации сроков выпуска ряда заказов. Эта информация позволяет учесть директивность изготовления ряда заказов или

обеспечивает более ранние сроки изготовления приоритетных заказов.

В случае отсутствия рекомендаций по полному формированию плана производства (т. е. при наличии лишь части информации о производственной программе) блок доформирует начальный план, учитывает циклы изготовления и целочисленность выпуска изделий по следующему правилу:

$$y_q^s = z_q^t / \sigma \quad (\forall s \in S^t),$$

где σ — количество месяцев в t -м квартале, в которые возможен выпуск изделий q -го заказа. В случае неоднозначности принятия решений либо используют информацию о приоритетах, либо добиваются равномерности выпуска номенклатурной продукции по единичным периодам.

После получения начального плана $\{y_q^s\}$, в первую очередь, оценивается следующая система неравенств:

$$\sum_{q \in Q} \sum_{s=1}^T c_q^s y_q^s \geq C, \quad (6.4.43)$$

$$\sum_{q \in Q} \sum_{s=1}^T a_{iq}^{ts} y_q^s \leq \Phi_i^t \quad (\forall t \in N_{12}, \forall i \in N_n), \quad (6.4.44)$$

$$\sum_{q \in Q_j} \sum_{s \in S^t} r_q^s y_q^s \geq R_j^t \quad (\forall t \in N_4, \forall j \in N_m), \quad (6.4.45)$$

$$\sum_{s=1}^T y_q^s \leq A_q \quad (\forall q \in N_g), \quad (6.4.46)$$

$$\alpha_q^s \leq y_q^s \leq \beta_q^s \quad (\forall s \in N_{12}, \forall q \in N_g). \quad (6.4.47)$$

В результате анализа этой системы выдается технико-экономическая информация о плане производства.

1. Выполнение годовых, квартальных и месячных основных технико-экономических показателей плана:

1) оценивается выполнение годового, квартальных и месячных планов по выпуску товарной продукции в денежном выражении, в том числе оценивается вклад каждой номенклатуры в выполнение этих показателей;

2) оценивается выполнение годового, квартальных и месячных планов выпуска продукции по каждой номенклатуре в соответствующей единице измерения и выдается следующая информация:

а) перечень номенклатур, планы выпуска по которым во все периоды выполняются или перевыполняются;

б) перечень номенклатур, планы выпуска по которым во все периоды недовыполняются с оценкой объемов недо-выполнения;

в) перечень номенклатур с неравномерным выпуском продукции и с характеристикой неравномерности;

3) по каждой номенклатуре выдается перечень заказов, выпускаемых в каждом месяце с указанием их основных реквизитов.

II. Использование фонда рабочего времени всех учитываемых в системе видов оборудования:

1) выделяется множество видов оборудования, по которым использование фонда рабочего времени выходит за установленные границы, хотя бы в одном месяце, указываются эти месяцы и объемы перегрузок

$$\sum_{q \in Q} \sum_{s=1}^T a_{iq}^{rs} y_q^s - \Phi_i^s > 0;$$

производится оценка этих превышений;

2) выделяются виды оборудования, для которых имеет место недоиспользование фонда рабочего времени, а в отдельные месяцы, определяются объемы недоиспользования

$$\Phi_i^s - \sum_{q \in Q} \sum_{s=1}^T a_{iq}^{rs} y_q^s > 0,$$

о чем выдается соответствующая информация;

3) выделяются виды оборудования с неравномерной загрузкой, которую невозможно обосновать особенностями производства, а можно только — качеством плана, и производится оценка этой неравномерности;

4) по каждому виду оборудования выдается перечень заказов, проходящих обработку на них в каждом месяце, с указанием объемов используемых фондов.

III. Выполнение ограничений на объемы выпуска изделий по каждому заказу:

1) выдается список заказов, для которых превышаются годовые объемы производства, т. е. для которых

$$\sum_{s=1}^T y_q^s > A_q;$$

2) оценивается равномерность выпуска изделий заказов по месяцам с выдачей соответствующей информации;

3) для каждого заказа формируется и выдается график его движения по циклу производства.

IV. Выполнение неформализованных условий производства, проверяются:

1) выполнение директивных сроков выпуска ряда заказов (например, обеспечение выпуска условных заказов в конце периода планирования);

2) комплектность выпуска заказов с определением месяца выпуска комплектного заказа;

3) равномерность выпуска продукции;

4) равномерность использования фондов рабочего времени лимитирующих видов оборудования;

5) равномерность выполнения плановых заданий по единичным периодам, и другие условия.

После получения оценок плана $\{y_q^s\}$ подсистема для частичного улучшения плана должна произвести последовательный ряд операций в режиме диалога, в их числе:

1) частично перераспределить объемы выпуска заказов по месяцам;

2) уточнить для ряда заказов возможные периоды их выпуска;

3) изменить значения ряда технико-экономических показателей, лимитирующих план;

4) изменить некоторые реквизиты у ряда заказов.

Затем подсистема проверяет соответствие технико-экономических показателей планов $\{z_q^i\}$ и $\{y_q^s\}$, а также выполнение ограничений по использованию фондов рабочего времени нелимитирующих видов оборудования, которые были изъяты из рассмотрения предшествующими блоками. При обнаружении несоответствия планы корректируются. После этого осуществляется согласование решений оптимизационной и имитационной подсистем, в частности, оценивается близость предлагаемых эвристических вариантов плана к оптимальному решению.

4.7. Технология прохождения задачи на ЭВМ. Процедура решения задачи ОКП не сводится к одноразовому расчету симплекс-методом, а является процессом получения оптимальных решений путем многократного обмена информацией между человеком и ЭВМ. В ходе этого обмена возникает целый поток информации, выдава-

емой ЭВМ, который оперативно подвергается анализу работниками плановых служб предприятия. На основе анализа выдаются конкретные рекомендации по корректировке исходной технико-экономической информации и экономико-математических моделей ОКП.

Приведем краткое описание процедуры получения оптимальных объемно-календарных планов производства.

Основными поставщиками информации о плане производства являются соответствующие планирующие службы предприятия. Информация может задаваться в следующих объемах:

- 1) полностью определен весь план производства, т. е. задано все множество $\{y_q^s\}$ ($\forall s \in N_{12}, \forall q \in Q$);
- 2) план производства задан частично, т. е. зафиксированы некоторые y_q^s , $q \in Q^* \subset Q$, $s \in T^* \subset T$;
- 3) существенно уменьшена возможность вариации сроков выпуска ряда заказов.

Благодаря удобной форме представления исходной информации (в виде симплекс-таблицы) дополнительная эвристическая информация о плане производства может быть легко учтена.

В первом случае, когда определен весь план производства, подсистема, в первую очередь, проверяет заполнение условий (6.4.2) — (6.4.4) и вычисляет значение целевой функции (6.4.1). Как правило, условие (6.4.4) всегда выполняется, условия (6.4.2), (6.4.3) не выполняются.

Во втором и третьем случаях, когда план производства определен частично, подсистема автоматически уменьшает размерность задачи, учитывая информацию о плане путем фиксирования значения ряда переменных и вычеркиванием части ограничений вида (6.4.2), (6.4.3). Как правило, объем эвристической информации таков, что размеры оставшейся задачи позволяют решать ее сравнительно легко. Полученный таким образом план производства будем называть частично эвристическим.

Имитационная подсистема производит анализ предлагаемых эвристических рекомендаций составления плана производства по двум направлениям. Во-первых, проверяется выполнение условий, формализованных в модели (6.4.1) — (6.4.4), и, во-вторых, проверяется ряд условий, существенно влияющих на возможность выполнения плана,

но не учитываемых в рамках этой модели. Из множества неформализованных условий рассматриваются только те, влияние которых может быть сравнительно просто оценено имеющейся в наличии информацией:

- 1) комплектность выпуска ряда заказов;
- 2) допустимость перегрузок рабочего фонда некоторых видов оборудования (такие перегрузки могут быть ликвидированы с помощью способов, не учитываемых в модели: кооперация, взаимозаменяемое оборудование и др.);
- 3) обеспеченность производственной программы лимитирующими ресурсами;
- 4) возможность вариации R_j^t для некоторых номенклатур.

Часть анализа производится в автоматическом режиме, часть — в режиме диалога «человек — ЭВМ». Оптимальный план, получаемый в оптимизационной подсистеме, также подвергается анализу в имитационной подсистеме, которая сравнительно часто обращается к информационной подсистеме, так как выдача выходных документов является частью диалога.

Кроме того, имитационная подсистема включает в себя ряд эвристических алгоритмов улучшения планов, которые реализуются в режиме диалога. Основные этапы получения оптимальных планов следующие:

- 1) информационная подсистема формирует и проверяет всю исходную технико-экономическую информацию;
- 2) имитационная подсистема производит экспресс-анализ информации путем неоднократного решения статической задачи с выделением существенно лимитирующих план видов оборудования и анализирует эвристические рекомендации составления плана;
- 3) информационная подсистема формирует симплекс-таблицу задачи ОКП, используя рекомендации имитационной подсистемы;
- 4) имитационная подсистема строит начальное эвристическое решение; используя решение статической задачи, переформирует задачу, оставляя лишь лимитирующие план виды оборудования;
- 5) начиная с начального решения; оптимизационная подсистема производит оптимизацию плана;

6) имитационная подсистема анализирует оптимальный план, выдает рекомендации и передает управление в информационную подсистему;

7) информационная подсистема производит распечатку оптимального плана.

Опыт решения задачи ОКП показал, что описанная процедура позволяет получать оптимальные объемно-календарные планы в реальном масштабе времени, а выходная информация достаточно полно характеризует состояние и ход производства на предприятии.

Исходя из учета информационного согласования подсистем, из того, что в каждой подсистеме используется определенная совокупность программ, работающих в различных режимах, и из необходимости ведения диалога, к математическому обеспечению задачи ОКП предъявляются следующие требования:

- 1) обеспечение настройки программ внутри подсистем;
- 2) согласование программ между подсистемами;

При этом в конкретной программе должны иметься:

- 1) входная информация;
- 2) промежуточная информация;
- 3) выходная информация;
- 4) контрольная информация;
- 5) корректировочная информация;
- 6) информация о диалоге между человеком и ЭВМ.

При этом во всех программах информация должна быть представлена в виде табуляграмм.

В процессе решения задачи ОКП программы в подсистеме могут работать по нескольку раз, и поэтому необходимо предусматривать:

- 1) краткую печать информации;
- 2) полную печать информации;
- 3) запись фрагментов информации на внешние накопители;

- 4) полную запись информации на внешние накопители.

Перечисление требований к программам позволяет лучше организовать получение решения задачи ОКП.

4.8. Основные направления развития системы ОКП.

1. Проведенные сравнения расчетного оптимального объемно-календарного плана с фактически выполненными объемами работ в течение периода планирования показали наличие существенных расхождений по основным

показателям плана. Это вызывает необходимость, с одной стороны, изучить поток возмущений, вызывающий такого рода отклонения, с целью его прогнозирования, и, с другой стороны, разработать методы принятия оптимальных решений в условиях изменяющейся информации. Для этого необходимо построить следящую за ходом производства систему, осуществляющую своевременную корректировку плановых заданий.

2. Некоторые технико-экономические показатели, участвующие в системе ОКП и существенно влияющие на формирование плана производства, являются интегральными. В их числе ряд сгруппированных заказов по запасным частям и другим видам продукции, фонды рабочего времени некоторых видов оборудования, принадлежащих различным цехам и т. д. Поэтому после получения укрупненного плана производства в системе ОКП необходима его детальная проработка по расшифровке укрупненных показателей. При этом появляется возможность учесть целый ряд особенностей производства, неформализованных в системе ОКП.

3. План производства, получаемый в системе ОКП, позволяет производить более детальную проработку оптимальной загрузки существенно лимитирующих производственную программу видов оборудования. В этом случае необходимо рассматривать задачи оптимального оперативного планирования и управления конкретными производствами с получением суточных или сменных заданий.

$$\mathbf{R}_+ = \{x: x \geq 0\}; \mathbf{R}_- = \mathbf{R} \setminus \mathbf{R}_+;$$

$$\mathbf{R}_+^n = \{x \in \mathbf{R}^n, x \geq 0\};$$

$$x \geq y \text{ означает: } x_i \geq y_i \ (i=1, \dots, n);$$

$$x > y \text{ означает: } x_i > y_i \ (i=1, \dots, n);$$

$$\alpha^+ = \max \{0, \alpha\}, \alpha \in \mathbf{R};$$

$$x^+ = [x_1^+, \dots, x_n^+];$$

$$(x, y) = \sum_{i=1}^n x_i y_i \text{ — скалярное произведение;}$$

x, y — отрезок в \mathbf{R}^n с концами x и y ;

$$\mathbf{N} = \{1, \dots, m, \dots\};$$

$$\mathbf{N}_m = \{1, \dots, m\};$$

$$|x|_0 = \max_{(i=1, \dots, n)} |x_i|; |x|_1 = \sum_{i=1}^n |x_i|;$$

$$|x - B| = \inf \{|x - y|: y \in B \subset \mathbf{R}^n\};$$

$$|A - B| = \sup_{x \in A} \inf_{y \in B} |x - y|;$$

$$|A - B|_+ = \max \{|A - B|, |B - A|\};$$

$$\tilde{\rho}(A, B) = \inf \{|x - y|: x \in A, y \in B\};$$

$|K|$ — число элементов конечного множества K ;

$\text{co } A$ — выпуклая оболочка множества $A \subset \mathbf{R}^n$;

$\text{int } A$ — внутренность множества $A \subset \mathbf{R}^n$;

$\text{graf } g(\cdot) = \{[x, y]: y = g(x)\}, g \in \{\mathbf{R}^n \rightarrow \mathbf{R}^m\}$;

∂A — граница выпуклого замкнутого множества $A \subset \mathbf{R}^n$;

$\text{Arg min } \{f(x): x \in M\} = \{y \in M: f(y) = \inf_{x \in M} f(x)\}$;

$\pi_M(x) = \arg \min \{|x - y|: y \in M\}$;

$$\Phi(A) = \bigcup_{x \in A} \Phi(x), \Phi \in \{\mathbf{R}^n \rightarrow 2^{\mathbf{R}^n}\};$$

$$\Phi(\Psi(x)) = \bigcup_{y \in \Psi(x)} \Phi(y), \Phi \text{ и } \Psi \text{ — из } \{\mathbf{R}^n \rightarrow 2^{\mathbf{R}^n}\};$$

$\{x_i\}'$ — производное множество от $\{x_i\}$;

$\nabla f(x)$ — градиент функции $f(x)$;

$\nabla_x f(x, y)$ — градиент по x функции $f(x, y)$;

$\partial f(x)$ — субдифференциал выпуклой функции $f(x)$;

$\partial_x f(x, y)$ — субдифференциал по x выпуклой по x функции $f(x, y)$;

F_M — класс M -фейеровских отображений;

$X \times Y$ — множество пар $[x, y], x \in X, y \in Y$;

$M^r = \{x: |x - M| \leq r\}, M \subset \mathbf{R}^n, r > 0,$

Введение.

Различным аспектам развития методов решения сложных задач оптимизации, с использованием диалога и других неформальных средств, посвящено большое число работ. Принципиальные вопросы комбинированных процедур оптимизации обсуждаются в работах В. М. Глушкова [11, 13], Н. Н. Моисеева [85, 86], Ю. И. Журавлева [48, 49].

Глава I.

§ 3. В связи с характеристикой нестационарных процессов математического программирования укажем на работы Н. Н. Красовского [66—68, 117], А. Б. Куржанского [70], в которых изучение задач теории управления ведется в таких предположениях, как неполнота априорных данных или сведений о начальном и текущем состояниях системы, их неточность, изменяемость в процессе реализаций этих состояний.

Глава II.

§ 1. Стохастический аналог фейеровских процессов изучался Ю. М. Ермольевым [41, 44].

§ 4. Результаты п. 4.4 принадлежат С. В. Плотникову.

Глава III.

§ 1. Итерационные процессы, порождаемые фейеровским оператором, зависящим от нестационарного параметра, впервые были рассмотрены в работе одного из авторов [23].

§ 4. Нестационарные процессы оптимизации в форме процессов для предельных экстремальных задач рассматривались Ю. М. Ермольевым и Н. А. Нурминским [43, 89].

Глава IV.

§ 1. В списке литературы мы приводим только некоторые работы, непосредственно связанные с содержанием книги. В связи с нашим подходом к моделям и методам классификации, основанным на алгебраических конструкциях теории линейных неравенств и математического программирования, отметим книгу В. А. Ковалевского [61], в которой методы оптимизации играют существенную роль. В статье А. А. Дородницына и др. [15] дан некоторый новый подход к формализации задачи классификации, использующий методы комбинаторной оптимизации; этот подход применяется к решению задачи классификации биологических объектов. В работах Ю. И. Журавлева [45—50] исследуются алгебры алгоритмов решения плохо

формализуемых задач, в частности задач распознавания образов; такой общий подход позволяет строить корректные алгоритмы классификации на базе некорректных эвристических алгоритмов. Математические модели задач классификации и подход к построению общей теории распознавания образов с точки зрения алгебраических систем предложены в книге Н. Г. Загоруйко [51].

§ 3. Понятие комитета однородной строгой системы линейных неравенств появилось в результате работ ряда американских авторов, в том числе К. Эйблау и Д. Кэйлора [112], посвященных многослойным сетям пороговых элементов, настраивающимся на распознавание образов. Некоторые виды таких сетей были названы «ассоциативными машинами»; для них были предложены простые эвристические алгоритмы обучения распознаванию типа методов линейной коррекции. Строгого обоснования таких алгоритмов предложено не было, но на экспериментах выяснилось, что алгоритмы не всегда приводят к цели, в связи с чем предполагались различные рекомендации по выбору начального приближения, числа пороговых элементов в сети и др. К. Эйблау и Д. Кэйлор ввели термин «комитет» («committee solution», «committee»), формализовав это понятие для ситуации, описываемой несовместной системой однородных строгих линейных неравенств. Они также сформулировали достаточное условие существования комитета такой системы, но доказательства не привели. Далее понятие комитета подробно исследовалось в работах одного из авторов настоящей книги [73—75, 78]. Некоторые рекуррентные процедуры построения комитета были предложены В. Н. Фоминым [101].

Глава V.

§ 2. Результаты п. 2.3 получены В. Ф. Князевой.

Глава VI.

§§ 1, 2. В решении перечисленных прикладных задач участвовали В. М. Кисляк, А. С. Садчиков [60], Л. Д. Попов [90], Г. Ф. Корнилова, Н. Н. Глезер, В. Г. Томилова, Л. И. Тягунов [64], В. С. Казанцев, А. И. Смирнов, Н. О. Сачков, А. И. Кривоногов, В. В. Андреев [1].

§ 3. Из работ, посвященных принципам организации и структуры пакетов прикладных программ, отметим статьи Н. Н. Яненко [110, 111], В. М. Глушкова и др. [12], А. А. Самарского и др. [59].

Разработчиками пакета программ «ОКП» являются Г. Ф. Корнилова, А. С. Садчиков, Н. Н. Глезер, В. М. Кисляк, Л. Д. Попов, М. А. Костина, М. С. Хрипун, В. Г. Томилова.

Пакет «КВАЗАР» был разработан В. С. Казанцевым при участии Н. Г. Белецкого, Н. О. Сачкова и В. С. Мезенцева [54, 55].

§ 4. Данный параграф написан В. М. Кисляком и А. С. Садчиковым — основными исполнителями работы по созданию системы объемно-календарного планирования [60]. В разработке этой системы участвовали также Г. Ф. Корнилова, Н. Н. Глезер, Л. Д. Попов и Л. Т. Коротаева [63, 98].

1. Андреев В. В., Мазуров Вл. Д., Мамин В. В. Некоторые математические методы измерения диагностической ценности медицинской информации, их применение в задачах диагностики и при оценке эффективности гормональной терапии при гипотиреозах.— В кн.: Медицинская системотехника, М., 1973.
2. Астафьев Н. Н. Об обратных теоремах обобщенной двойственности в задачах выпуклого программирования. ДАН СССР, 1971, 199, № 3.
3. Астафьев Н. Н. Устойчивость и маргинальные значения задачи выпуклого программирования.— Сибирский матем. журнал, 1978, 19, № 3.
4. Астафьев Н. Н. Двойственная регуляризация задачи линейного программирования, заданной последовательностью реализаций.— ЖВМ и МФ, 1978, 18, № 5.
5. Беляев Е. В., Есипов С. Н., Мазуров Вл. Д., Тягунов Л. И. Экономико-математическая модель рационального подбора шихт для коксования из углей Карагандинского бассейна.— В кн.: Экономика черной металлургии, вып. 1, М., Металлургия, 1972.
6. Булавский В. А. Итеративный метод решения задачи линейного программирования.— ДАН СССР, 1961, 137, № 2.
7. Булавский В. А. Итеративный метод решения общей задачи линейного программирования.— Сибирский матем. журнал, 1962, 3, № 3.
8. Булавский В. А. О расширении области сходимости итеративных методов повышенной точности.— ДАН СССР, 1972, 205, № 2.
9. Булавский В. А. Об одном типе проекционных методов в математическом программировании.— В кн.: Оптимизация, Новосибирск, ИМ СО АН СССР, 1972, № 5.
10. Булатов В. П. Методы погружения в задачах оптимизации.— Новосибирск: Наука, 1978.
11. Глушков В. М. О диалоговом методе решения оптимизационных задач.— Кибернетика, 1975, № 4.
12. Глушков В. М., Вельбицкий И. В., Стогний А. А. Единый автомат-лингвистический подход к технологии разработки пакетов прикладных задач.— Тезисы докладов международной конференции «Структура и организация пакетов программ». Тбилиси, 1976.
13. Глушков В. М., Олейрш Г. Б. Некоторые задачи стохастического программирования, возникающие в диалоговых систе-

- мах планирования в условиях неполной информации.— Кибернетика, 1977, № 3.
14. Деметьев В. Т. Задача выбора типажа оборудования.— В кн.: Методы управления большими системами, т. II, Иркутск, 1970.
 15. Дородницын А. А., Каспшицкая М. Ф., Сергиенко И. В. Об одном подходе к формализации классификации.— Кибернетика, 1976, № 6.
 16. Евтушенко Ю. Г. Численные методы нелинейного программирования.— ДАН СССР, 1977, 221, № 5.
 17. Евтушенко Ю. Г., Жадан В. Г., Релаксационный метод решения задач нелинейного программирования.— ЖВМ и МФ, 1977, 17, № 4.
 18. Еремин И. И. О несовместных системах линейных неравенств.— ДАН СССР, 1961, 138.
 19. Еремин И. И. Итеративный метод для чебышевских приближений несовместных систем линейных неравенств.— ДАН СССР, 1962, 143, № 6.
 20. Еремин И. И. Обобщение релаксационного метода Моцкина — Агмона.— УМН, 1965, 20, вып. 2.
 21. Еремин И. И. Релаксационный метод решения систем неравенств с выпуклыми функциями в левых частях.— ДАН СССР, 1965, 160, № 5.
 22. Еремин И. И. О системах неравенств с выпуклыми функциями в левых частях.— Известия АН СССР (серия матем.), 1966, 30, № 2.
 23. Еремин И. И. О некоторых итерационных методах в выпуклом программировании.— Экономика и матем. методы, 1966, 2, № 6.
 24. Еремин И. И. О методе штрафов в выпуклом программировании.— ДАН СССР, 1967, 173, № 4.
 25. Еремин И. И. О скорости сходимости в методе фейеровских приближений.— Матем. заметки, 1968, 4, № 1.
 26. Еремин И. И. Метод фейеровских приближений в выпуклом программировании.— Матем. заметки, 1968, 3, № 2.
 27. Еремин И. И. Применение метода фейеровских приближений к решению задач выпуклого программирования с негладкими ограничениями.— ЖВМ и МФ, 1969, 9, № 5.
 28. Еремин И. И. Фейеровские отображения и задача выпуклого программирования.— Сибирский матем. журнал, 1969, 10, № 5.
 29. Еремин И. И. О задачах выпуклого программирования с противоречивыми ограничениями.— Кибернетика, 1971, № 4.
 30. Еремин И. И. О задачах последовательного программирования.— Сибирский матем. журнал, 1973, 14, № 1.
 31. Еремин И. И. Процедура обмена информацией о некоторых моделях иерархии управления.— В кн.: Нестационарные процессы математического программирования, Свердловск. ИММ УНЦ АН СССР, 1974, № 4.
 32. Еремин И. И. Дискретные процессы фейеровского типа для негладких задач выпуклого программирования.— Алгебра и логика, 1976, 15, № 6.
 33. Еремин И. И., Астафьев Н. Н. Введение в теорию линейного и выпуклого программирования.— М: Наука, 1976.

34. Еремин И. И., Мазуров Вл. Д. Итерационный метод решения задачи выпуклого программирования. — ДАН СССР, 1966, 170, № 1.
35. Еремин И. И., Мазуров Вл. Д. О нестационарных процессах математического программирования. — В кн.: Нестационарные процессы математического программирования, Свердловск, ИММ УНЦ АН СССР, 1974, вып. 14.
36. Еремин И. И., Мазуров Вл. Д. Нестационарные процессы математического программирования и некоторые приложения. — Системы и управление, 1974, № 3, София.
37. Еремин И. И., Мазуров Вл. Д. Синтез и анализ нестационарных процессов математического программирования, их приложения в экономике и технике. — Экономико-математический обзор, 1976, 12, № 3, Прага.
38. Еремин И. И., Мазуров Вл. Д. Нестационарные процессы математического программирования. — Тезисы докладов международной конференции «Структура и организация пакетов программ», Тбилиси, 1976.
39. Еремин И. И., Мазуров Вл. Д. Итерационный метод обучения дискриминации бесконечных множеств. — Кибернетика, 1977, № 5.
40. Еремин И. И., Мазуров Вл. Д. Автоматизация управления параметрами итерационного процесса для задач математического программирования. — Кибернетика, 1978, № 5.
41. Ермольев Ю. М. О методе обобщенных стохастических градиентов и стохастических квазифейеровских последовательностях. — Кибернетика, 1969, № 2.
42. Ермольев Ю. М. Методы стохастического программирования. — М.: Наука, 1976.
43. Ермольев Ю. М., Нурминский Н. А. Предельные экстремальные задачи. — Кибернетика, 1973, № 4.
44. Ермольев Ю. М., Туниев А. Д. Случайные фейеровские и квазифейеровские последовательности. — В кн.: Теория оптимальных решений, Киев, ИК АН УССР, 1968, № 2.
45. Журавлев Ю. И. Экстремальные алгоритмы в математических моделях для задач распознавания и классификации. — ДАН СССР, 1976, 231, № 3.
46. Журавлев Ю. И. Непараметрические задачи распознавания образов. — Кибернетика, 1976, № 6.
47. Журавлев Ю. И. Алгебры над множествами некорректных (эвристических) алгоритмов. — ДАН СССР, 1977, 235, № 4.
48. Журавлев Ю. И. Корректные алгебры над множествами некорректных (эвристических) алгоритмов. 1. — Кибернетика, 1977, № 4.
49. Журавлев Ю. И. Корректные алгебры над множествами некорректных (эвристических) алгоритмов, 2. — Кибернетика, 1977, № 6.
50. Журавлев Ю. И. Экстремальные алгоритмы в алгебре над некорректными алгоритмами. — ДАН СССР, 1977, 237, № 3.
51. Загоруйко Н. Г. Метод распознавания образов и их применение. — М.: Сов. радио, 1972.
52. Истомин Л. А. О решении некоторых обратных задач математического программирования. — ДАН СССР, 1974, 214, № 5.

53. Истомин Л. А. Модификация метода Хоанг Туя минимизации вогнутой функции на многограннике.— ЖВМ и МФ, 1977, 17, № 6.
54. Казанцев В. С. Пакет «КВАЗАР» прикладных программ распознавания образов.— В кн.: Планирование горно-металлургического производства (программы оптимизации). Свердловск, ИММ УНЦ АН СССР, 1977, вып. 7.
55. Казанцев В. С., Белецкий Н. Г., Мезенцев С. В., Сачков Н. О. Пакет «КВАЗАР» прикладных программ распознавания образов (версия 2).— Свердловск: ИММ УНЦ АН СССР, 1978.— Научный отчет.
56. Казанцев В. С., Тягунов Л. И. Алгоритм комитетного распознавания.— В кн.: Программы оптимизации (оптимальное планирование), Свердловск, ИММ УНЦ АН СССР, 1974, вып. 5.
57. Карманов В. Г. Математическое программирование.— М.: Наука, 1975.
58. Карпов В. Я. Оформление программ, написанных на ФОРТРАНе.— М.: ИПМ АН СССР, 1976.
59. Карпов В. Я., Корягин Д. А., Самарский А. А. Принципы разработки пакетов прикладных программ для задач математической физики.— ЖВМ и МФ, 1978, 18, № 2.
60. Кисляк В. М., Садчиков А. С. Оптимизация объемно-календарного планирования производства Уралмашзавода.— В кн.: Математические методы в планировании промышленного производства, Свердловск, ИММ УНЦ АН СССР, 1977, вып. 22.
61. Ковалевский В. А. Методы оптимальных решений в распознавании изображений.— М.: Наука, 1976.
62. Коновалов А. Н., Яненко Н. Н. Модульный принцип построения программ как основа создания пакета прикладных программ решения задач механики сплошной среды.— В кн.: Комплексы программ математической физики, Новосибирск, 1972.
63. Корнилова Г. Ф. и др. Пакет «Оптим-П» решения задач оптимального планирования для машин БЭСМ-6.— В кн.: Программы оптимизации. Свердловск, ИММ УНЦ АН СССР, 1974, вып. 5.
64. Костина М. А. Об организации итерационных процедур в методах фейеровского типа и штрафных функций.— В кн.: Нестационарные процессы математического программирования, Свердловск, ИММ УНЦ АН СССР, 1974.
65. Костина М. А., Хрипун М. С. Фейеровские процессы и приемы ускорения их сходимости.— В кн.: Математические методы в некоторых задачах оптимального планирования. Свердловск, ИММ УНЦ АН СССР, 1971.
66. Красовский Н. Н. К задаче управления с неполной информацией.— Изв. АН СССР (Техн. киберн.), 1976, № 2.
67. Красовский Н. Н. Об управлении при неполной информации.— Прикл. мат. и мех., 1976, 40, № 2.
68. Красовский Н. Н. О дифференциальных эволюционных системах.— Прикл. мат. и мех., 1977, 41, № 5.
69. Кривоногов А. И. Итерационные алгоритмы комитетного распознавания.— В кн.: Оптимизация планирования в больших горно-металлургических системах, Свердловск, ИММ УНЦ АН СССР, 1976.— ВИНТИ, № 3198—76 ДЕП.

70. Куржанский А. Б. Управление и наблюдение в условиях неопределенности.— М.: Наука, 1977.
71. Ляпунов А. А., Яблонский С. В. Теоретические проблемы кибернетики.— В кн.: Проблемы кибернетики, вып. 9, М., Физматгиз, 1963.
72. Мамин В. В., Мазуров Вл. Д., Смирнов А. И. Радиологическое определение степени гипотиреоидного состояния с помощью метода комитетов.—Труды Всесоюзного радиологического съезда, Ереван, 1977.
73. Мазуров В. Д. О построении комитета системы выпуклых неравенств.— Кибернетика, 1967, № 2.
74. Мазуров Вл. Д. Распознавание образов как средство автоматического выбора процедуры в вычислительных методах.— ЖВМ и МФ, 1970, 10, № 6.
75. Мазуров Вл. Д. Комитеты систем неравенств и задача распознавания.— Кибернетика, 1971, № 3.
76. Мазуров Вл. Д. Применение методов теории распознавания образов в оптимальном планировании и управлении.—Труды I Всесоюзной конференции по оптимальному планированию и управлению народным хозяйством, М., ЦЭМИ, 1971.
77. Мазуров Вл. Д. Об одном итерационном методе планирования, использующем распознавание образов для учета плохо формализуемых факторов.—Изв. АН СССР (Техн. киберн.), 1973, № 3.
78. Мазуров Вл. Д. О системах неравенств сопряженного вида в задачах распознавания.—В кн.: Матем. методы в некоторых задачах оптим. планир., Свердловск, ИММ УНЦ АН СССР, 1973, вып. 4.
79. Мазуров Вл. Д. Теория линейных неравенств и распознавание образов.—В кн.: Методы фейеровского типа в выпуклом программировании, Свердловск, ИММ УНЦ АН СССР, 1975, вып. 18.
80. Мазуров Вл. Д. Методы математического программирования и распознавания образов в планировании производства.—В кн.: Математические методы планирования промышленного производства, Свердловск, ИММ УНЦ АН СССР, 1977, вып. 22.
81. Мазуров Вл. Д., Мамин В. В., Мезенцев А. И., Пасынкова Т. Г. Применение математических методов распознавания образов к диагностике гипотиреоидных состояний.— В кн.: Техника и методология в радиологии, Свердловск, НТО, 1974.
82. Макаров В. Л. Модели оптимального роста экономики.— Экономика и матем. методы, 1968, 5, № 4.
83. Мерзляков Ю. И. Об одном релаксационном методе решения системы линейных неравенств.— ЖВМ и МФ, 1962, 2, № 3.
84. Метод комитетов в распознавании образов.—Свердловск: ИММ УНЦ АН СССР, 1974, вып. 6.
85. Мойсеев Н. Н. Элементы теории оптимальных систем.— М.: Наука, 1975.
86. Мойсеев Н. Н. Кибернетическое описание эколого-экономических систем.— Кибернетика, 1977, № 6.
87. Никайдо Х. Выпуклые структуры и математическая экономика.— М.: Мир, 1972.

88. Нурминский Е. А. Условия сходимости алгоритмов нелинейного программирования.— Кибернетика, 1972, № 6.
89. Нурминский Е. А. Об одной задаче нестационарной оптимизации.— Кибернетика, 1977, № 2.
90. Попов Л. Д. Задача таксономии при агрегировании в линейных программах большой размерности.— В кн.: Математические методы планирования промышленного производства. Свердловск, ИММ УНЦ АН СССР, 1977, вып. 2.
91. Пшеничный Б. Н. Необходимые условия экстремума.— М.: Наука, 1969.
92. Пшеничный Б. Н., Данилин Ю. М. Численные методы в экстремальных задачах.— М.: Наука, 1975.
93. Самарский А. А. Общие вопросы модульного программирования для задач математической физики.— Тезисы докладов международной конференции «Структура и организация пакетов программ», Тбилиси, 1976.
94. Скарин В. Д. О методе штрафных функций для задач нелинейного программирования.— ДАН СССР, 1973, 209, № 6.
95. Скарин В. Д. К регуляризации минимаксных задач, возникающих в выпуклом программировании.— ЖВМ и МФ, 1977, 17, № 6.
96. Скарин В. Д. О минимизации выпуклых функций.— В кн.: Нестационарные процессы матем. программирования, Свердловск, ИММ УНЦ АН СССР, 1974, вып. 14.
97. Смирнов А. И. Динамика возрастно-генетического состава биологической популяции в одной математической модели.— В кн.: Математические методы в планировании промышленного производства, Свердловск, ИММ УНЦ АН СССР, 1977, вып. 22.
98. Соловьева А. А., Корнилова Г. Ф. Обзор способов организации пакетов прикладных программ.— Свердловск, ИММ УНЦ АН СССР, 1977.
99. Тихонов А. Н. О многоцелевых системах обработки данных физических экспериментов.— Тезисы докладов международной конференции «Структура и организация пакетов программ», Тбилиси, 1976.
100. Хрипун М. С. О некоторых модификациях фейеровских методов решения систем линейных неравенств.— В кн.: Методы выпуклого программирования и приложения, Свердловск, ИММ УНЦ АН СССР, 1973.
101. Фомин В. Н., Математическая теория обучаемых опознающих систем.— Л.: Изд-во ЛГУ, 1976.
102. Чарин В. С. Линейные преобразования и выпуклые множества.— Киев: Вища школа, 1979.
103. Черников С. Н. Свертывание систем линейных неравенств.— ДАН СССР, 1960, 131, № 3.
104. Черников С. Н. Линейные неравенства.— М.: Наука, 1968.
105. Черников С. Н. Свертывание конечных систем линейных неравенств.— Доповіді АН УССР, 1969, № 1.
106. Черников С. Н. Системы линейных неравенств и некоторые их приложения.— В кн.: Математизация знаний и научно-технический прогресс, Киев, Наукова думка, 1975.

107. Черникова Н. В. Алгоритм для нахождения общей формулы неотрицательных решений системы линейных неравенств.— ЖВМ и МФ, 1965, 5, № 2.
108. Эльстер К. Х., Гроссман Х. Решение нелинейных оптимизационных задач с помощью штрафных и барьерных функций.— В кн.: Применение исследования операций в экономике, М., Экономика, 1977.
109. Яблонский С. В. Основные понятия кибернетики.— В кн.: Проблемы кибернетики, вып. 2, М., Физматгиз, 1959.
110. Яненко Н. Н., Коновалов А. Н., Технологические аспекты численных методов математической физики.— Acta Univ. carol. Math. et Phys., 1974, 15, № 1, 2.
111. Яненко Н. Н. Проблемы математической технологии.— Тезисы докладов международной конференции «Структура и организация пакетов программ», Тбилиси, 1976.
112. Ablow C. M., Kaylor D. J. Inconsistent homogeneous linear inequalities.— Bul. Amer. Math. Soc., 1965, 71, № 1.
113. Agmon S. The relaxation method for linear inequalities.— Canad. J. Mathem., 1954, 6, № 3.
114. Babikov G. V., Mazurov V. I. D. Some methods of evaluation and construction of optimal systems with education functions.— Prace Naukowe Osrodka Badan Prognostycznych Politechniki Wroclawskiej 4, Seria: Konferencje 2, Wydawn. Polit. Wroclawskiej, 1975.
115. Beer K. Lösung grosser linearer Optimierungsaufgaben.— Berlin: VEB Deutschen Verlag der Wissenschaften, 1977.
116. Eremin I. I., Mazurov V. I. D. Nonstationary processes for mathematical optimization problems under the conditions of poorly formalized constraints and incomplete defining information.— Berlin: Springer—Verlag, 1975.
117. Krasovskii N. N. Optimal control under conditions of conflict or uncertainty.— In: Proc. IFSC 6th World Congr., Boston — Cambridge, Mass., Pt4, 1975.— Pittsburgh, Pa, 28.5/1—8, 1975.
118. Motzkin T. S., Schoenberg J. J. The relaxation method for linear inequalities.— Canad. J. Mathem., 1954, 6, № 3.
119. Nozicka F., Guddat J., Hollatz H., Bank B. Theorie de linearen parametrischen Optimierung.— Berlin: Akademier Verlag, 1974.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Агрегирование ограничений и переменных 219
— — — — —, замкнутая схема 219
— — — — —, открытая схема 219
Анализ дискриминантный аффинный 137
Аппроксимация симметрическая 97
- Диагностика 218
— медицинская 232
— технологического процесса 224
- Задача дискриминантного анализа 17
— коррекции материала наблюдений 135
— несобственная 95
— обучения 8
— объемно-календарного планирования 221
— оценки информативности 18
— разбиения материала наблюдений 135
— собственная 95
— таксономии 18
— управления процессом лечения 232
— — технологическим процессом 223
- Индекс неравенства свертки 168
— элемента конуса $C^*(U)$ 168
- Комитет системы неравенств 151
Конструкция комитетная 152
Коэффициент коррекции 139
Критерий близости 133
— качества классификации 138
К.—Т. условие 17
- Метод дискриминантного анализа непараметрический 138
— линейной коррекции 139
— свертывания 168
Множество звездное 207
— информационное 29
Модель нестационарная 26
— объекта (в распознавании образов)
- Отображение замкнутое 49, 102
— итерационное 27
— М-фейеровское 48
Оценки двойственные оптимальные 17
- Пакет прикладных программ 233
Подсистема системы множеств 151
— — неравенств 159
Помехоустойчивость 85
Последовательность, монотонная по Фейеру 10
— поточечно-монотонная 10
— Q-фейеровская 48
Правило распознавания 132
Приближение дискретное 141
Признак объекта 134
Прогнозирование технико-экономических показателей 226
Процесс нестационарный 26
— следящий 32
— z-нормализованный 94
- Распознавание образов 127
- Свертка полная 168
Свойство дуальности для систем неравенств 160
Сигнатура элемента 165
Система множеств 151
— неравенств 14
— представителей 151
Составляющая информационная 30
Субградиент 40
Субдифференциал 40
- Таксон 18, 133
Типология 218
Точка седловая 17
- Формализуемость плохая 13, 183
Функция дискриминантная 132
— разделяющая 132
- Эксперт 26
- Неравенство несовместное 170

