
Н. С. Бахвалов, А. В. Лапин, Е. В. Чижонков

ЧИСЛЕННЫЕ МЕТОДЫ В ЗАДАЧАХ И УПРАЖНЕНИЯХ

Учебное пособие

4-е издание (электронное)

Рекомендовано

*УМО по классическому университетскому образованию
в качестве учебного пособия для студентов высших
учебных заведений, обучающихся по специальностям
высшего профессионального образования
010101 «Математика» и 010901 «Механика»*



Москва
БИНОМ. Лаборатория знаний
2015

УДК 519.6(07)
ББК 22.193я73
Б30

Рецензенты:

кафедра математического моделирования МЭИ (ТУ)
(зав. кафедрой д-р физ.-мат. наук, проф. А. А. Амосов);
д-р физ.-мат. наук, проф. В. И. Лебедев
(РНЦ «Курчатовский институт»)

Бахвалов Н. С.

Б30 Численные методы в задачах и упражнениях [Электронный ресурс] : учебное пособие / Н. С. Бахвалов, А. В. Лапин, Е. В. Чижонков ; под ред. В. А. Садовниченко. — 4-е изд. (эл.). — Электрон. текстовые дан. (1 файл pdf : 243 с.). — М. : БИНОМ. Лаборатория знаний, 2015. — Систем. требования: Adobe Reader XI ; экран 10".

ISBN 978-5-9963-2980-9

Материал учебного пособия полностью соответствует требованиям Государственного образовательного стандарта по математике. В книге содержатся элементы теории, примеры решений задач и упражнения для самостоятельной работы. Отличительная особенность пособия состоит в том, что представленные задачи и упражнения (их около 700) разбиты по рекомендуемым темам семинарских занятий, а их подбор призван способствовать закреплению материала, излагаемого в теоретическом курсе. При этом типовые задачи снабжены решениями (числом около 200) и могут быть использованы студентами для самостоятельного изучения предмета, а приведенные ответы и указания помогут преподавателям в выборе содержательных и интересных задач в соответствии со спецификой вуза.

Для студентов университетов, педагогических вузов и вузов с углубленным изучением математики, а также для студентов технических вузов, аспирантов и преподавателей, инженеров и научных работников, использующих в своей практической деятельности численные методы.

УДК 519.6(07)
ББК 22.193я73

Деривативное электронное издание на основе печатного аналога: Численные методы в задачах и упражнениях : учебное пособие / Н. С. Бахвалов, А. В. Лапин, Е. В. Чижонков ; под ред. В. А. Садовниченко. — 2-е изд., перераб. и доп. — М. : БИНОМ. Лаборатория знаний, 2010. — 240 с. : ил. — ISBN 978-5-9963-0333-5.

В соответствии со ст. 1299 и 1301 ГК РФ при устранении ограничений, установленных техническими средствами защиты авторских прав, правообладатель вправе требовать от нарушителя возмещения убытков или выплаты компенсации

ISBN 978-5-9963-2980-9 © БИНОМ. Лаборатория знаний, 2010

Предисловие ко второму изданию



Первое издание этой книги вышло в свет в серии «Высшая математика» издательства «Высшая школа» под общей редакцией академика РАН В. А. Садовниченко в 2000 г. За это время стало ясно, что она нашла своего читателя. Книга активно используется студентами, аспирантами, преподавателями не только Московского государственного университета имени М. В. Ломоносова, но и большого количества других вузов страны. В то же время обнаружился ряд опечаток и неточностей в изложении и появились свежие идеи по улучшению подачи материала. Поэтому авторы с благодарностью приняли любезное предложение издательства «БИНОМ. Лаборатория знаний» о переиздании учебного пособия.

Во втором издании исправлены замеченные недостатки и добавлена новая глава про интегральные уравнения. Кроме того, из методических соображений был уточнен (в основном расширен) набор задач и упражнений, а также порядок их следования внутри параграфов; добавлено значительное количество решений, указаний и ответов.

К сожалению, в 2005 г. скоропостижно ушел из жизни один из авторов книги — Николай Сергеевич Бахвалов, выдающийся ученый — математик и педагог, доктор физико-математических наук, профессор, заслуженный деятель науки РФ, дважды лауреат Государственной премии, действительный член РАН. Однако подготовка второго издания к печати была выполнена на основании его многочисленных замечаний и рекомендаций.

Авторы искренне верят, что новая версия пособия окажется столь же полезной, как и предыдущая, для студентов и аспирантов, изучающих и применяющих численные алгоритмы, преподавателей, проводящих учебные занятия, а также для инженеров и исследователей, использующих в своей деятельности методы вычислительной математики.

А. В. Лапин, Е. В. Чижевков

Предисловие к первому изданию



Математика как наука возникла в связи с необходимостью решения практических задач: измерений на местности, навигации и т. д. Вследствие этого математика всегда была численной математикой, ее целью являлось получение решения в виде числа.

Крупнейшие ученые прошлого сочетали в своих трудах как построение математического описания явления природы (математической модели), так и его исследование. Анализ усложненных моделей требовал создания новых, как правило, численных или асимптотических методов решения задач. Названия некоторых из таких методов — методы Ньютона, Эйлера, Гаусса, Чебышёва — свидетельствуют о том, что их разработкой занимались крупнейшие ученые своего времени.

Последние полвека характерны бурным развитием вычислительной техники и теории численных методов. В результате происходит быстрое изменение взглядов на весь комплекс вопросов, связанных с применением компьютеров, в частности, на требования к численным методам. Поэтому нельзя предложить пособия по численным методам, содержащего рецепты решения всех реально встречающихся проблем. При выборе способа решения конкретной задачи всякое пособие играет роль лишь общего руководства, отталкиваясь от которого исследователь анализирует возникающие проблемы.

Настоящее пособие написано на основе опыта преподавания курса «Численные методы» на механико-математическом факультете и факультете вычислительной математики и кибернетики МГУ им. М. В. Ломоносова и полностью соответствует требованиям Государственного образовательного стандарта по математике, рекомендованного Министерством образования и науки Российской Федерации. Каждый раздел начинается с изложения базовых определений и теоретических результатов; далее рассматриваются типовые задачи, как правило, снабженные подробными решениями, а также приводятся упражнения для самостоятельных занятий, снабженные указаниями и ответами.

В процессе написания использовалась литература, список которой приведен в конце книги. Поскольку многие задачи встречаются в различных изданиях, установить авторство оказалось практически невозможно, поэтому было решено не делать в тексте ссылки на литературу.

Пособие охватывает традиционный материал по разностным уравнениям, приближению функций, численному интегрированию и дифференцированию, интегральным уравнениям, задачам алгебры и решению нелинейных уравнений, приближенным методам решения дифференциальных уравнений как обыкновенных, так и с частными производными, а также по влиянию вычислительной погрешности в различных алгоритмах. Оно является хорошим дополнением к классическому университетскому учебнику Бахвалова Н. С., Жидкова Н. П., Кобелькова Г. М. «Численные методы» и недавно вышедшему учебному пособию Бахвалова Н. С., Корнева А. А., Чижонкова Е. В. «Численные методы. Решения задач и упражнения», носящему справочный характер.

Авторы

Погрешность решения задачи



Если a — точное значение некоторой величины, a^* — известное приближение к нему, то *абсолютной погрешностью* приближенного значения a^* обычно называют некоторую величину $\Delta(a^*)$, про которую известно, что

$$|a^* - a| \leq \Delta(a^*).$$

Относительной погрешностью приближенного значения называют некоторую величину $\delta(a^*)$, про которую известно, что

$$\left| \frac{a^* - a}{a^*} \right| \leq \delta(a^*).$$

Относительную погрешность часто выражают в процентах.

§ 1. Вычислительная погрешность

Наиболее распространенная форма представления действительных чисел в компьютерах — *числа с плавающей запятой*. Множество F чисел с плавающей запятой характеризуется четырьмя параметрами: основанием системы счисления p , разрядностью t и интервалом показателей $[L, U]$. Каждое число x , принадлежащее F , представимо в виде

$$x = \pm \left(\frac{d_1}{p} + \frac{d_2}{p^2} + \dots + \frac{d_t}{p^t} \right) p^\alpha,$$

где целые числа $p, \alpha, d_1, \dots, d_t$ удовлетворяют неравенствам

$$0 \leq d_i \leq p - 1; \quad i = 1, \dots, t; \quad L \leq \alpha \leq U.$$

Часто d_i называют *разрядами*, t — *длиной мантииссы*, α — *порядком числа*. *Мантииссой* m (дробной частью) x называют число в скобках. Множество F называют *нормализованным*, если для каждого $x \neq 0$ справедливо $d_1 \neq 0$.

Удобно определить, что округление с точностью ε — это некоторое отображение fl действительных чисел \mathbb{R} на множество F чисел с плавающей запятой, удовлетворяющее следующим аксиомам.

- 1) Для произвольного $y \in \mathbb{R}$, такого, что результат отображения $fl(y) \in F$, $fl(y) \neq 0$, имеет место равенство $fl(y) = y(1 + \eta)$, $|\eta| \leq \varepsilon$.
- 2) Обозначим результат арифметической операции $*$ с числами $a, b \in F$ через $fl(a*b)$. Если $fl(a*b) \neq 0$, то $fl(a*b) = (a*b)(1 + \eta)$, $|\eta| \leq \varepsilon$.

Приведенные соотношения позволяют изучать влияние ошибок округления в различных алгоритмах.

Если результат округления не принадлежит F , то его обычно называют *переполнением* и обозначают ∞ .

Будем считать, что ε — точная верхняя грань для $|\eta|$. При традиционном способе округления чисел имеем $\varepsilon = \frac{1}{2}p^{1-t}$, при округлении отбрасыванием разрядов $\varepsilon = p^{1-t}$. Величину ε часто называют *машинной точностью*.

1.1. Построить нормализованное множество F с параметрами $p = 2$, $t = 3$, $L = -1$, $U = 2$.

1.2. Каков результат операций при использовании множества F из 1.1?

$$\begin{array}{lll}
 1) x = fl\left(\frac{23}{32}\right), & 2) x = fl\left(\frac{1}{8}\right), & 3) x = fl(4), \\
 4) x = fl\left(\frac{1}{2} + \frac{3}{4}\right), & 5) x = fl\left(\frac{3}{8} + \frac{5}{4}\right), & 6) x = fl\left(3 + \frac{7}{2}\right), \\
 7) x = fl\left(\frac{7}{16} - \frac{3}{8}\right), & 8) x = fl\left(\frac{1}{4} \cdot \frac{5}{16}\right).
 \end{array}$$

1.3. Верно ли, что всегда $fl\left(\frac{a+b}{2}\right) \in [a, b]$?

1.4. Пусть отыскивается наименьший корень уравнения

$$y^2 - 140y + 1 = 0.$$

Вычисления производятся в десятичной системе счисления, причем в мантиссе числа после округления удерживается четыре разряда. Какая из формул $y = 70 - \sqrt{4899}$ или $y = \frac{1}{70 + \sqrt{4899}}$ дает более точный результат?

1.5. Пусть приближенное значение производной для гладкой функции $f(x)$ определяется при $h \ll 1$ по формуле $f'(x) \approx \approx (f(x+h) - f(x))/h$, а сами значения $f(x)$ вычисляются с абсолютной погрешностью Δ . Получить оценку полной погрешности этой формулы при условии $|f''(x)| \leq M_2$. Найти

оптимальный шаг h_0 , при котором минимизируется величина оценки полной погрешности.

- 1.6. Найти абсолютную погрешность вычисления суммы $S = \sum_{j=1}^n x_j$, где все x_j — числа одного знака.
- 1.7. Пусть вычисляется сумма $\sum_{j=1}^{10^6} \frac{1}{j^2}$. Какой алгоритм $S_0 = 0$, $S_n = S_{n-1} + \frac{1}{n^2}$, $n = 1, \dots, 10^6$, или $R_{10^6+1} = 0$, $R_{n-1} = R_n + \frac{1}{n^2}$, $n = 10^6, \dots, 1$, $\tilde{S}_{10^6} = R_0$, следует использовать, чтобы вычислительная погрешность суммы была меньше?
- 1.8. Предложить способ вычисления суммы, состоящей из слагаемых одного знака, минимизирующий влияние вычислительной погрешности.
- 1.9. Предложить способ вычисления знакопеременной суммы, минимизирующий влияние вычислительной погрешности.
- 1.10. Пусть значение многочлена $P_n(x) = a_0 + a_1x + \dots + a_nx^n$ вычисляется в точке $x = 1$ по *схеме Горнера*:

$$P_n(x) = a_0 + x(a_1 + x(\dots(a_{n-1} + a_nx)\dots)).$$

Какую погрешность можно ожидать в результате, если коэффициенты округлены с погрешностью η ?

- 1.11. Оценить погрешность вычисления скалярного произведения двух векторов $S = \sum_{j=1}^n x_j y_j$, если их компоненты округлены с погрешностью η .
- 1.12. Для элементов последовательности $I_n = \int_0^1 x^n e^{x-1} dx$, $n = 1, 2, \dots$, справедливо точное рекуррентное соотношение $I_n = 1 - n I_{n-1}$, $I_1 = 1/e$. Можно ли его использовать для приближенного вычисления интегралов, считая, что ошибка округления допускается только при вычислении I_1 ?
- 1.13. Пусть вычисления ведутся по формуле

$$y_{n+1} = 2y_n - y_{n-1} + h^2 f_n, \quad n = 1, 2, \dots,$$

где y_0, y_1 заданы точно, $|f_n| \leq M$ и $h \ll 1$. Какую вычислительную погрешность можно ожидать при вычислении y_n для

больших значений n ? Улучшится ли ситуация, если вычисления вести по формулам $\frac{z_{n+1} - z_n}{h} = f_n$, $\frac{y_n - y_{n-1}}{h} = z_n$?

- 1.14. Найти приближенное значение интеграла $\int_0^{2\pi} \sin^{100} x \, dx$ с относительной погрешностью не более 10 %.

§ 2. Погрешность функции

Пусть искомая величина y является функцией параметров x_j , $j = 1, 2, \dots, n$: $y = y(x_1, x_2, \dots, x_n)$. Область G допустимого изменения параметров x_j известна, требуется получить приближение к y и оценить его погрешность. Если y^* — приближенное значение величины y , то *предельной абсолютной погрешностью* называют величину

$$A(y^*) = \sup_{(x_1, x_2, \dots, x_n) \in G} |y(x_1, x_2, \dots, x_n) - y^*|;$$

при этом *предельной относительной погрешностью* называют величину $R(y^*) = A(y^*)/|y^*|$.

- 2.1. Доказать, что предельная абсолютная погрешность $A(y^*)$ минимальна при $y^* = (y_1 + y_2)/2$, где

$$y_1 = \inf_G y(x_1, x_2, \dots, x_n), \quad y_2 = \sup_G y(x_1, x_2, \dots, x_n).$$

- 2.2. Показать, что предельная абсолютная погрешность суммы или разности равна сумме их предельных абсолютных погрешностей.
- 2.3. Показать, что предельная относительная погрешность произведения или частного с точностью до членов второго порядка малости равна сумме предельных относительных погрешностей.
- 2.4. Пусть $y = y(x_1, x_2, \dots, x_n)$ — непрерывно дифференцируемая функция. Положим

$$A_{\text{sup}}(y^*) = \sum_{j=1}^n B_j \Delta(x_j^*), \quad \text{где } B_j = \sup_G \left| \frac{\partial y(x_1, x_2, \dots, x_n)}{\partial x_j} \right|,$$

$$A_{\text{lin}}(y^*) = \sum_{j=1}^n |b_j| \Delta(x_j^*), \quad \text{где } b_j = \left. \frac{\partial y(x_1, x_2, \dots, x_n)}{\partial x_j} \right|_{\mathbf{x}=(x_1^*, x_2^*, \dots, x_n^*)}.$$

Доказать, что $A(y^*) \leq A_{\text{sup}}(y^*)$ и, если величина $\rho = \left(\sum_{j=1}^n \Delta^2(x_j^*)\right)^{1/2}$ мала, то справедливо равенство: $A_{\text{sup}}(y^*) = A_{\text{lin}}(y^*) + o(\rho)$.

2.5. Пусть $y = x^{10}$, $x^* = 1$ и задана абсолютная погрешность

$$1) \Delta(x^*) = 0,001; \quad 2) \Delta(x^*) = 0,1.$$

Вычислить величины $A_{\text{sup}}(y^*)$, $A_{\text{lin}}(y^*)$, $A(y^*)$.

2.6. Получить линейную оценку погрешности функции, заданной неявно уравнением $F(y, x_1, \dots, x_n) = 0$.

2.7. Пусть y^* — простой (не кратный!) корень уравнения $y^2 + by + c = 0$, вычисленный при заданных приближенных значениях коэффициентов b^* , c^* , и известны погрешности $\Delta(b^*)$, $\Delta(c^*)$. Доказать, что

$$A_{\text{lin}}(y^*) = \frac{|y^*| \Delta(b^*) + \Delta(c^*)}{|2y^* + b^*|}.$$

2.8. Показать, что если уравнение из **2.7** имеет кратный корень, то погрешность приближенного значения корня имеет порядок $O(\sqrt{\rho})$, где $\rho = (\Delta^2(b^*) + \Delta^2(c^*))^{1/2} \ll 1$.


2.9. Имеется приближение y^* к простому корню уравнения $f(y) = 0$. Вывести приближенное равенство $y - y^* \approx -f(y^*)/f'(y^*)$.

2.10. С каким минимальным числом верных знаков надо взять $\lg 2$ для того, чтобы вычислить корни уравнения $y^2 - 2y + \lg 2 = 0$ с четырьмя верными знаками?

2.11. Пусть ограниченные по модулю величиной M коэффициенты уравнения $ay^2 + by + c = 0$ заданы с одинаковой относительной погрешностью δ . Найти максимальную абсолютную (относительную) погрешность, с которой могут вычисляться корни уравнения.

Глава II

Приближение функций и производных



§ 3. Полиномиальная интерполяция

Пусть $a = x_1 < x_2 < \dots < x_n = b$ — набор различных точек (узлов) на отрезке $[a, b]$, в которых заданы значения функции $f(x)$ так, что $f_i = f(x_i)$, $i = 1, \dots, n$. Требуется построить многочлен наименьшей степени, принимающий в точках x_i значения f_i , и оценить погрешность приближения достаточно гладкой функции $f(x)$ этим многочленом на всем отрезке $[a, b]$.

Здесь и далее под многочленом всюду будет пониматься алгебраический многочлен; в случае использования тригонометрического многочлена, это будет специально уточнено.

Приведем в явном виде вспомогательные многочлены $\Phi_i(x)$ степени $n - 1$, удовлетворяющие условиям $\Phi_i(x_i) = 1$, $\Phi_i(x_j) = 0$ при $j \neq i$. Имеем $\Phi_i(x) = \prod_{\substack{j=1 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}$. Запишем с их помощью формулу

для искомого многочлена Лагранжа $L_n(x) = \sum_{i=1}^n f_i \Phi_i(x)$. Так как существует единственный многочлен степени $n - 1$, принимающий в n различных точках заданные значения, то многочлен $L_n(x)$ есть решение поставленной задачи.

Теорема. Пусть n -я производная функции $f(x)$ непрерывна на отрезке $[a, b]$. Тогда для любой точки $x \in [a, b]$ существует точка $\xi \in [a, b]$ такая, что справедливо равенство

$$f(x) - L_n(x) = \frac{f^{(n)}(\xi)}{n!} \omega_n(x), \quad \text{где } \omega_n(x) = \prod_{i=1}^n (x - x_i).$$

Следствием этого представления является оценка погрешности в равномерной норме

$$\varepsilon_n = \|f(x) - L_n(x)\| \leq \frac{\|f^{(n)}(x)\|}{n!} \|\omega_n(x)\|, \quad \text{где } \|g(x)\| = \max_{x \in [a, b]} |g(x)|.$$

Величина $\lambda_n = \max_{x \in [a, b]} \sum_{i=1}^n |\Phi_i(x)|$ называется *константой Лебега интерполяционного процесса*. Скорость ее роста в зависимости от величины n существенно влияет как на сходимость $L_n(x)$ к $f(x)$, так и на оценку вычислительной погрешности интерполяции.

В задачах параграфа узлы $x_i = a + (i - 1)(b - a)/(n - 1)$, $i = 1, 2, \dots, n$, $n > 1$, на отрезке $[a, b]$ будем называть *равноотстоящими*.

3.1. Построить многочлен Лагранжа при $n = 3$ для следующих случаев:

$$1) \quad x_1 = -1, \quad x_2 = 0, \quad x_3 = 1, \quad 2) \quad x_1 = 1, \quad x_2 = 2, \quad x_3 = 4, \\ f_1 = 3, \quad f_2 = 2, \quad f_3 = 5; \quad f_1 = 3, \quad f_2 = 4, \quad f_3 = 6.$$

3.2. Построение многочлена Лагранжа $L_n(x)$ эквивалентно задаче нахождения коэффициентов c_i из системы уравнений

$$\sum_{i=0}^{n-1} c_i x_j^i = f_j \quad \text{при } j = 1, \dots, n. \text{ Показать, что эта система при}$$

больших n может быть близка к вырожденной.

3.3. Найти $\sum_{i=1}^n x_i^p \Phi_i(x)$ при $p = 0, \dots, n$.

3.4. На отрезке $[a, b]$ заданы равноотстоящие узлы. Вычислить $\|\omega_n(x)\|$ при $n = 2, 3, 4$.

3.5. Функция $f(x)$ приближается на $[a, b]$ по n равноотстоящим узлам. Найти наибольшее целое p в оценке погрешности $\varepsilon_n \leq 10^{-p}$ для следующих случаев:

$$1) \quad [0, 0, 1], \quad f(x) = \sin 2x, \quad n = 2; \quad 2) \quad [-1, 0], \quad f(x) = e^x, \quad n = 3.$$

3.6. Приближение к числу $\ln 15,2$ вычислено следующим образом. Найдены точные значения $\ln 15$ и $\ln 16$ и построена линейная интерполяция между этими числами. Показать, что если a и a^* — соответственно точное и интерполированное значения $\ln 15,2$, то справедлива оценка $0 < a - a^* < 4 \cdot 10^{-4}$.

3.7. Функция $f(x) = \frac{1}{A^2 - x}$ приближается на $[-4, -1]$ многочленом Лагранжа по узлам $-4, -3, -2, -1$. При каких значениях A оценка погрешности в равномерной норме не превосходит 10^{-5} ?

3.8. Доказать, что если узлы интерполяции расположены симметрично относительно некоторой точки c , а значения интерполируемой функции в симметричных узлах равны, то интерполяционный многочлен Лагранжа — функция, четная относительно точки c .

3.9. Показать, что многочлен Лагранжа может быть построен рекуррентным способом:

$$L_1(x) = f(x_1), L_n(x) = L_{n-1}(x) + (f(x_n) - L_{n-1}(x_n)) \frac{\omega_{n-1}(x)}{\omega_{n-1}(x_n)}, \quad n \geq 2,$$

$$\text{где } \omega_1(x) = (x - x_1), \omega_n(x) = \omega_{n-1}(x)(x - x_n).$$

3.10. Построить многочлен Лагранжа $L_n(x)$ степени $n - 1$, удовлетворяющий условиям $L_n(x_k) = y_k$:

1) $n = 4, x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 4; y_1 = 2, y_2 = 3, y_3 = 4, y_4 = 6;$

2) $n = 3, x_k = 2k - 1, y_k = 8 \sin \frac{\pi}{6}(2k - 1), k = 1, 2, 3.$

3.11. Построить интерполяционный многочлен для функции $f(x) = |x|$ по узлам $-1, 0, 1$.

3.12. Построить интерполяционный многочлен для функции $f(x) = x^2$ по узлам $0, 1, 2, 3$.

3.13. Построить многочлен Лагранжа $L_4(x)$ третьей степени, удовлетворяющий условиям $L_4(x_k) = y_k: x_k = k - 5, y_k = 3k^3 + 2k^2 + k + 1, k = 1, 2, 3, 4$.

3.14. Функция $f(x)$ приближается интерполяционным многочленом Лагранжа на $[a, b]$ по n равноотстоящим узлам. Найти наибольшее целое p в оценке погрешности вида $\varepsilon_n \leq 10^{-p}$ для следующих случаев:

1) $f(x) = \frac{1}{\pi} \int_0^{\pi} \cos(x \sin t) dt, [0, 1], n = 3;$

2) $f(x) = \ln x, [1, 2], n = 4.$

3.15. Оценить погрешность приближения функции $\exp(x)$ интерполяционным многочленом Лагранжа $L_3(x)$, построенным по узлам $0, 0, 0, 1, 0, 2$, в точке: 1) $x = 0,05;$ 2) $x = 0,15$.

- 3.16.** Функция $\sin x$ приближается на отрезке $[0, \pi/4]$ интерполяционным многочленом по значениям в точках $0, \pi/8, \pi/4$. Оценить погрешность интерполяции на этом отрезке.
- 3.17.** Функция $\ln(x)$ приближается на отрезке $[1, 2]$ интерполяционным многочленом третьей степени по узлам $1, 4/3, 5/3, 2$. Доказать, что погрешность интерполяции не превосходит величины $1/300$.
- 3.18.** Функция $f(x) = \exp(2x)$ приближается на отрезке $[-1/2, 1/2]$ интерполяционным многочленом второй степени по узлам: $-1/2, 0, 1/2$. Доказать, что погрешность интерполяции не превосходит величины $\sqrt{3}/9$.
- 3.19.** Оценить погрешность интерполяции функции $f(x) = \operatorname{arctg} x$ на отрезке $[0, 1]$ многочленом Лагранжа пятой степени, построенным по равноотстоящим узлам.
- 3.20.** Оценить число равноотстоящих узлов интерполяции на отрезке $[0, \pi/4]$, обеспечивающее точность $\varepsilon \leq 10^{-2}$ приближения функции $f(x) = \sin x$.
- 3.21.** Определить степень многочлена Лагранжа, построенного по равноотстоящим узлам, обеспечивающую точность $\varepsilon \leq 10^{-3}$ приближения функции $\exp(x)$ на отрезке $[0, 1]$.
- 3.22.** Пусть функция $f(x) = \sin x$ задана на отрезке $[0, b]$. При каком b многочлен Лагранжа $L_3(x)$, построенный по равноотстоящим узлам, приближает эту функцию с погрешностью $\varepsilon_3 \leq 10^{-3}$?
- 3.23.** Привести пример непрерывной на отрезке $[-1, 1]$ функции, для которой интерполяционный процесс Лагранжа по равноотстоящим узлам расходится.
- 3.24.** Пусть гладкая функция $f(x)$ задана на отрезке $[a, b]$ и $\max_{x \in [a, b]} |f''(x)| \leq 1$. Оценить погрешность приближения $f(x)$ кусочно-линейным интерполянтном, построенным на равномерной сетке с шагом h : $x_i = a + ih, 0 \leq i \leq N, Nh = b - a$.
- 3.25.** С каким шагом следует составлять таблицу функции $\sin x$ на отрезке $[0, \pi/2]$, чтобы погрешность кусочно-линейной интерполяции не превосходила величины $0,5 \cdot 10^{-6}$?

3.26. Пусть $f(x)$ имеет непрерывную производную на отрезке $[a, b]$ и $p(x)$ — многочлен, приближающий $f'(x)$ с точностью ε в равномерной норме. Доказать, что многочлен $q(x) = f(a) + \int_a^x p(t) dt$ приближает $f(x)$ с точностью $\varepsilon(b-a)$ в той же норме.

Построить многочлен $P_3(x) = a_0 + a_1x + a_2x^2 + a_3x^3$, удовлетворяющий условиям:

3.27. $P_3(-1) = 0, P_3(1) = 1, P_3(2) = 2, a_3 = 1.$

3.28. $P_3(0) = P_3(-1) = P_3(1) = 0, a_2 = 1.$

3.29. $P_3(-1) = 0, P_3(1) = 1, P_3(2) = 2, a_1 = 1.$

3.30. $P_3(-2) = P_3(-1) = P_3(1) = 0, a_0 = 1.$

Построить многочлен $P_4(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$, удовлетворяющий условиям:

3.31. $P_4(0) = 0, P_4(-1) = 1, P_4(2) = 2, P_4(3) = 3, \sum_{i=0}^4 a_i = 0.$

3.32. $P_4(1) = P_4(-1) = P_4'(0) = P_4''(0) = 0, P_4(0) = 1.$

3.33. $P_4(0) = 0, P_4(1) = 1, P_4(2) = 2, P_4(3) = 3, \sum_{i=0}^4 a_i = 0.$

3.34. Пусть $x \in [a, b], y \in [-1, 1]$ и узлы интерполяции x_i, y_i при $i = 1, 2, \dots, n$, связаны линейным соотношением $x_i = x(y_i) = \frac{a+b}{2} + \frac{b-a}{2}y_i$. Доказать, что константы Лебега $\lambda_n^{[a,b]}$ и $\lambda_n^{[-1,1]}$, соответствующие этим отрезкам, совпадают.

3.35. Показать, что для равноотстоящих узлов при $n \geq 2$ справедлива оценка снизу для константы Лебега $\lambda_n \geq K_1 2^n/n^{3/2}$ с постоянной K_1 , не зависящей от n .

3.36. Показать, что для равноотстоящих узлов при $n \geq 2$ справедлива оценка сверху для константы Лебега $\lambda_n \leq K_2 2^n$ с постоянной K_2 , не зависящей от n .

3.37. Определить узлы интерполяции, при которых константа Лебега λ_3 минимальна.

3.38. Показать, что если x_1, \dots, x_{2n} — вещественные, то функция

$T(x) = \prod_{k=1}^{2n} \sin \frac{x - x_k}{2}$ является тригонометрическим многочленом вида $T(x) = a_0/2 + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx)$ с вещественными коэффициентами a_k, b_k .

3.39. Доказать, что интерполяционный тригонометрический многочлен $T(x)$, удовлетворяющий условиям $T(x_j) = y_j$ при $j = 0, 1, \dots, 2n$, где $0 \leq x_0 < x_1 < \dots < x_{2n} < 2\pi$, может быть записан в виде

$$T(x) = \sum_{k=0}^{2n} y_k t_k(x), \quad t_k(x) = \prod_{\substack{s=0 \\ s \neq k}}^{2n} \sin \frac{x - x_s}{2} / \sin \frac{x_k - x_s}{2}.$$

3.40. Доказать, что для любых x_0, x_1, \dots, x_{2n} , удовлетворяющих условиям $0 \leq x_0 < x_1 < \dots < x_{2n} < 2\pi$, и для любых y_0, y_1, \dots, y_{2n} существует единственный тригонометрический многочлен $T(x) = a_0/2 + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx)$, удовлетворяющий условиям $T(x_j) = y_j, j = 0, 1, 2, \dots, 2n$. Если при этом y_0, y_1, \dots, y_{2n} — вещественные, то и коэффициенты a_k, b_k являются вещественными.

3.41. Доказать, что для любых x_0, x_1, \dots, x_n , удовлетворяющих условиям $0 \leq x_0 < x_1 < \dots < x_n < \pi$, и для любых y_0, y_1, \dots, y_n существует единственный тригонометрический многочлен

$$C(x) = \sum_{k=0}^n a_k \cos kx, \text{ удовлетворяющий условиям } C(x_j) = y_j, \\ j = 0, 1, 2, \dots, n.$$

3.42. Построить интерполяционный тригонометрический многочлен на отрезке $[0, 1]$ по заданным значениям $f(0), f(h), f(2h), f(3h)$ при $h = 1/3$.

3.43. Построить интерполяционный тригонометрический многочлен второй степени $T_2(x) = a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x$, удовлетворяющий следующим условиям:

$$T_2(0) = 0, T_2(\pi/4) = 1, T_2(\pi/2) = 1, T_2(3\pi/4) = 1, T_2(\pi) = 1.$$

- 3.44.** Построить интерполяционный тригонометрический многочлен минимальной степени по заданным значениям $f(-\pi) = 0$, $f(-\pi/2) = 0$ и $f(\pi/2) = 1$.
- 3.45.** Доказать, что тригонометрический многочлен $T_n(z)$ степени n имеет в любой полосе $\operatorname{Re}(z) \in [a, a + 2\pi)$ ровно $2n$ корней.
- 3.46.** Пусть $T_n(x)$ — интерполяционный тригонометрический многочлен степени n , построенный по равноотстоящим узлам на $[0, 2\pi]$ для функции $f(x) \in C^{(\alpha)}$, $\alpha > 0$. Доказать, что в равномерной норме

$$\lim_{n \rightarrow \infty} \|T_n - f\| = 0.$$

- 3.47.** Функция двух переменных $f(x_1, x_2)$ аппроксимируется интерполяционным многочленом $P(x_1, x_2) = a_0 + a_1x_1 + a_2x_2 + a_3x_1x_2$. При этом $f(0, 0) = 1$, $f(1, 0) = 2$, $f(0, 1) = 4$, $f(1, 1) = 3$. Найти $P(1/2, 1/2)$.
- 3.48.** Пусть $P(x_1, x_2)$ — многочлен от двух переменных степени не выше n по каждой переменной и $P(k/n, m/n) = 0$, $k, m = 0, 1, \dots, n$. Доказать, что $P(x_1, x_2) \equiv 0$.

§ 4. Многочлены Чебышёва

Имеется несколько способов определения последовательности многочленов Чебышёва первого рода. Рассмотрим некоторые из них.

а) *Рекуррентное соотношение:*

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n \geq 1.$$

б) *Тригонометрическая форма.* При любом η имеем

$$\cos((n+1)\eta) = 2 \cos \eta \cos(n\eta) - \cos((n-1)\eta).$$

Полагая $\eta = \arccos x$, получаем

$$T_n(x) = \cos(n \arccos x), \quad n \geq 0.$$

Простое следствие: $|T_n(x)| \leq 1$ при $|x| \leq 1$.

в) *Разностное уравнение.* Рекуррентное соотношение является разностным уравнением по переменной n . Ему соответствует характеристическое уравнение $\mu^2 - 2x\mu + 1 = 0$. Следовательно,

$\mu_{1,2} = x \pm \sqrt{x^2 - 1}$, $T_n(x) = C_1 \mu_1^n + C_2 \mu_2^n$. Из начальных условий получаем $C_1 = C_2 = 1/2$, что приводит к формуле

$$T_n(x) = \frac{1}{2} \left(\left(x + \sqrt{x^2 - 1} \right)^n + \left(x - \sqrt{x^2 - 1} \right)^n \right), \quad n \geq 0.$$

Отметим, что все многочлены $T_{2n}(x)$ — четные, $T_{2n+1}(x)$ — нечетные, т. е. $T_k(-x) = (-1)^k T_k(x)$, а коэффициент при старшем члене $T_n(x)$ равен 2^{n-1} для $n \geq 1$.

4.1. Доказать следующие свойства многочленов Чебышёва:

1) $T_{2n}(x) = 2T_n^2(x) - 1$;

2) $I_{mn} = \int_{-1}^1 \frac{T_n(x) T_m(x)}{\sqrt{1-x^2}} dx = \begin{cases} 0 & \text{при } n \neq m, \\ \pi/2 & \text{при } n = m \neq 0, \\ \pi & \text{при } n = m = 0; \end{cases}$

3) $\int_{-1}^x T_n(y) dy = \frac{1}{2} \left(\frac{1}{n+1} T_{n+1}(x) - \frac{1}{n-1} T_{n-1}(x) \right) - \frac{(-1)^n}{n^2-1}$,
 $n \geq 2$;

4) $(1-x^2) T_n''(x) - x T_n'(x) + n^2 T_n(x) = 0$, $n \geq 0$.

4.2. Пусть $x^2 + y^2 = 1$. Доказать, что $T_{2n}(y) = (-1)^n T_{2n}(x)$.

4.3. Найти все нули многочлена Чебышёва $T_n(x)$.

4.4. Найти все экстремумы многочлена Чебышёва $T_n(x)$ на отрезке $[-1, 1]$.

4.5. Доказать, что *приведенный* многочлен Чебышёва $\overline{T}_n(x) = 2^{1-n} T_n(x)$ наименее уклоняется от нуля на отрезке $[-1, 1]$ среди всех многочленов $P_n(x)$ со старшим коэффициентом 1, т. е.

$$\|P_n(x)\| = \max_{[-1,1]} |P_n(x)| \geq \max_{[-1,1]} |\overline{T}_n(x)| = 2^{1-n}, \quad n \geq 1.$$

4.6. Доказать единственность многочлена, наименее уклоняющегося от нуля на отрезке $[-1, 1]$, среди всех многочленов со старшим коэффициентом 1.

4.7. Найти многочлен, наименее уклоняющийся от нуля на отрезке $[a, b]$, среди всех многочленов со старшим коэффициентом 1.

- 4.8. Пусть $\omega_n(x) = \prod_{i=1}^n (x - x_i)$. Показать, что при любом выборе узлов x_i имеет место неравенство $\|\omega_n(x)\| \geq (b - a)^n 2^{1-2n}$. Сравнить полученный результат с аналогичным для равномерного распределения узлов.
- 4.9. В классе многочленов степени n , принимающих в точке a ($|a| \geq 1$) значение $b \neq 0$, найти наименее уклоняющийся от нуля на отрезке $[-1, 1]$.
- 4.10. Пусть $k \leq n$, $0 \leq a < b$. В классе многочленов $P_n(x)$ степени n , удовлетворяющих условию $P_n^{(k)}(0) = c \neq 0$, найти наименее уклоняющийся от нуля на отрезке $[a, b]$.
- 4.11. Среди всех многочленов $P_n(x) = x^n + \dots$ степени $n \geq 2$, удовлетворяющих условиям $P_n(-1) = P_n(1) = 0$, найти наименее уклоняющийся от нуля на отрезке $[-1, 1]$.
- 4.12. Пусть $P_n(x)$ — многочлен степени n и $\max_{x \in [-1, 1]} |P_n(x)| = M$. Доказать, что для всех x таких, что $|x| \geq 1$, выполняется неравенство $|P_n(x)| \leq M |T_n(x)|$, где $T_n(x)$ — многочлен Чебышёва степени n .
- 4.13. Для производных многочленов Чебышёва доказать равенства $\frac{T'_{2n}}{2n} = 2(T_{2n-1} + T_{2n-3} + \dots + T_1)$, $\frac{T'_{2n+1}}{2n+1} = 2(T_{2n} + T_{2n-2} + \dots + T_2) + 1$.
- 4.14. Пусть функция $f(x)$ представима при $|x| \leq 1$ в виде $f(x) = \sum_{k=0}^{\infty} a_k T_k(x)$, где $\sum_{k=0}^{\infty} |a_k| < \infty$, $T_k(x)$ — многочлены Чебышёва. Доказать, что для всех $x \in [-1, 1]$ справедливо равенство
- $$\int_{-1}^x f(t) dt = \frac{a_0}{2} x + \sum_{k=1}^{\infty} \frac{1}{2k} (a_{k-1} - a_{k+1}) T_k(x) + a_0 - \frac{a_1}{4} + \sum_{k=2}^{\infty} \frac{(-1)^{k+1} a_k}{k^2 - 1}.$$
- 4.15. Вычислить значение многочлена Чебышёва n -й степени в точке: 1) $x = 1/2$; 2) $x = -1/2$.
- 4.16. Вычислить значение первой производной многочлена Чебышёва n -й степени в точке: 1) $x = 1$; 2) $x = -1$.

- 4.17. Функция $f(x) = \sin 2x$ приближается многочленом Лагранжа на отрезке $[0, 2]$ по чебышёвским узлам $x_i = 1 + \cos \frac{2i-1}{12} \pi$, $1 \leq i \leq 6$. Найти наибольшее целое p в оценке погрешности вида $\varepsilon_6 \leq \frac{1}{3} 10^{-p}$.
- 4.18. Функция $f(x) = \cos x$ приближается многочленом Лагранжа на отрезке $[-1, 1]$ по чебышёвским узлам $x_i = \cos \frac{2i-1}{10} \pi$, $1 \leq i \leq 5$. Найти наибольшее целое p в оценке погрешности вида $\varepsilon_5 \leq 10^{-p}$.
- 4.19. Функция e^x приближается многочленом Лагранжа на отрезке $[0, 1]$ по чебышёвским узлам $x_i = \frac{1}{2} + \frac{1}{2} \cos \frac{(2i-1)\pi}{8}$, $1 \leq i \leq 4$. Доказать, что погрешность интерполяции ε_4 не превосходит величины $e \cdot 10^{-3}$.
- 4.20. Среди всех многочленов вида $a_3x^3 + 2x^2 + a_1x + a_0$ найти наименее уклоняющийся от нуля на отрезке $[3, 5]$.
- 4.21. Среди всех многочленов вида $a_2x^2 + x + a_0$ найти наименее уклоняющийся от нуля на отрезке $[-1, 1]$.
- 4.22. Среди всех многочленов вида $5x^3 + a_2x^2 + a_1x + a_0$ найти наименее уклоняющийся от нуля на отрезке $[1, 2]$.
- 4.23. Среди всех многочленов вида $a_3x^3 + a_2x^2 + a_1x + 4$ найти наименее уклоняющийся от нуля на отрезке $[1, 3]$.
- 4.24. Среди всех многочленов вида $a_3x^3 + a_2x^2 + 3x + a_0$ найти наименее уклоняющийся от нуля на отрезке $[2, 4]$.
- 4.25. Показать, что если для интерполяции многочленом Лагранжа на отрезке $[-1, 1]$ взяты чебышёвские узлы $x_i = \cos \frac{2i-1}{2n} \pi$, $i = 1, 2, \dots, n$, то справедлива асимптотическая оценка сверху для константы Лебега $\lambda_n \leq K_2 \ln n$ с постоянной K_2 , не зависящей от n .
- 4.26. Показать, что если для интерполяции многочленом Лагранжа на отрезке $[-1, 1]$ взяты чебышёвские узлы $x_i = \cos \frac{2i-1}{2n} \pi$, $i = 1, 2, \dots, n$, то справедлива асимптотическая оценка снизу для константы Лебега $\lambda_n \geq K_1 \ln n$ с постоянной K_1 , не зависящей от n .

4.27. Определить константу Лебега λ_3 на отрезке $[-1, 1]$ для чебышёвских узлов интерполяции — нулей многочлена Чебышёва $T_3(x)$.

В приложениях встречаются также многочлены Чебышёва *второго рода* $U_n(x)$. Они удовлетворяют рекуррентному соотношению и начальным условиям: $U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x)$, $U_0(x) = 1$, $U_1(x) = 2x$.

4.28. Показать, что над полем действительных чисел для $U_n(x)$ справедливо представление

$$U_n(x) = \begin{cases} \frac{\sin((n+1)\arccos x)}{\sin(\arccos x)} & \text{при } |x| \leq 1, \\ \frac{1}{2\sqrt{x^2-1}} \left(\left(x + \sqrt{x^2-1}\right)^{n+1} - \left(x - \sqrt{x^2-1}\right)^{n+1} \right) & \text{при } |x| \geq 1. \end{cases}$$

4.29. Показать, что общее решение разностного уравнения $y_{n+1}(x) - 2xy_n(x) + y_{n-1}(x) = 0$ представимо в виде $y_n = C_1(x)T_n(x) + C_2(x)U_{n-1}(x)$.

4.30. Проверить соотношения для $T_n(x)$ и $U_n(x)$ при $n \geq i \geq 1$:

$$1) T_{n-1}(x) - xT_n(x) = (1-x^2)U_{n-1}(x),$$

$$2) U_{n-1}(x) - xU_n(x) = -T_{n+1}(x),$$

$$3) U_{n+i}(x) + U_{n-i}(x) = 2T_i(x)U_n(x),$$

$$4) U_{in-1}(x) = 2U_{i-1}(T_n(x)).$$

4.31. Показать, что $\max_{|x| \leq 1} |U_n(x)| = U_n(1) = n+1$.

4.32. Вычислить

$$I_{mn} = \int_{-1}^1 \sqrt{1-x^2} U_n(x) U_m(x) dx.$$

§ 5. Численное дифференцирование

Пусть известны значения функции $f(x)$ в точках x_1, x_2, \dots, x_n и требуется приближенно определить производную $f^{(k)}(x_0)$ для некоторого $0 \leq k \leq n-1$. Построим интерполяционный многочлен $L_n(x)$

и положим $f^{(k)}(x_0) \approx L_n^{(k)}(x_0)$. Тогда для погрешности справедливо представление

$$f^{(k)}(x) - L_n^{(k)}(x) = \sum_{j=0}^k \frac{k!}{(k-j)!(n+j)!} f^{(n+j)}(\xi_j) \omega_n^{(k-j)}(x).$$

Для системы равноотстоящих узлов ($x_{i+1} - x_i = h$) часто используют другой подход, основанный на получении приближений для старших производных через приближения для младших, аналогично последовательному дифференцированию. Базовыми являются следующие выражения:

$$\begin{aligned} \partial f(x) &= \frac{f(x+h) - f(x)}{h}, & \bar{\partial} f(x) &= \frac{f(x) - f(x-h)}{h}, \\ \tilde{\partial} f(x) &= \frac{\partial f(x) + \bar{\partial} f(x)}{2}, \end{aligned}$$

которые являются простейшими аналогами первой производной функции $f(x)$. Их называют *разностями вперед, назад и центральной* соответственно. Для вывода оценок погрешностей при данном подходе удобно использовать разложения в ряд Тейлора.

Для получения формул численного дифференцирования на практике также используют *метод неопределенных коэффициентов*. Он заключается в следующем: искомую формулу записывают в виде

$f^{(k)}(x_0) = \sum_{i=1}^n c_i f(x_i) + R(f)$, и коэффициенты c_i определяют из системы линейных уравнений при $R(f) = 0$, для получения которой последовательно полагают $f(x)$ равной $1, x, x^2, \dots, x^{n-1}$.

В задачах параграфа будем считать функцию $f(x)$ достаточно гладкой, т. е. имеющей все непрерывные производные требуемого порядка.

5.1. Показать, что в точке $x = x_i$ (одном из узлов интерполяции) справедлива оценка погрешности

$$|f'(x_i) - L'_n(x_i)| \leq \frac{\max_x |f^{(n)}(x)|}{n!} \prod_{\substack{j=1 \\ j \neq i}}^n |x_i - x_j|.$$

5.2. Доказать равенства:

- 1) $\partial f(x) - f'(x) = h f''(\xi)/2, x < \xi < x + h;$
- 2) $\tilde{\partial} f(x) - f'(x) = h^2 f'''(\xi)/6, x - h < \xi < x + h.$

5.3. Получить явные формулы для приближений старших производных, используя разности вперед, назад и центральную:
 $f''(x) \approx \bar{\partial}\partial f(x)$, $f'''(x) \approx \bar{\partial}\bar{\partial}\partial f(x)$, $f^{(4)}(x) \approx \bar{\partial}^2\partial^2 f(x)$.

5.4. Найти величину $K_i = K_i(h)$ в следующих равенствах:

- 1) $\bar{\partial}\partial f(x) - f''(x) = K_2 f^{(4)}(\xi)$, $x - h < \xi < x + h$;
- 2) $\bar{\partial}\bar{\partial}\partial f(x) - f'''(x) = K_3 f^{(5)}(\xi)$, $x - 2h < \xi < x + 2h$;
- 3) $\bar{\partial}^2\partial^2 f(x) - f^{(4)}(x) = K_4 f^{(6)}(\xi)$, $x - 2h < \xi < x + 2h$.

5.5. Считая, что значения функции в формулах численного дифференцирования $\bar{\partial}\partial f(x)$, $\bar{\partial}^2\partial^2 f(x)$ заданы с абсолютной погрешностью Δ , получить оценки полной погрешности этих формул как суммы погрешности метода и вычислительной погрешности. Найти оптимальный шаг h_0 , при котором минимизируется величина оценки полной погрешности.

5.6. Методом неопределенных коэффициентов построить формулы численного дифференцирования наиболее высокого порядка точности по h :

- 1) $f'(0) \approx [a f(-2h) + b f(0) + c f(h)]/h$;
- 2) $f''(0) \approx [a f(-h) + b f(h) + c f(2h) + d f(3h)]/h^2$.

5.7. Доказать, что $\bar{\partial}f(0) - f'(0) = \frac{1}{4h} \int_{-h}^h (h - |x|)^2 f'''(x) dx$.

5.8. Используя формулу Тейлора с остаточным членом в интегральной форме $f(b) = f(a) + (b - a)f'(a) + \dots + \frac{(b - a)^n}{n!} f^{(n)}(a) + \frac{1}{n!} \int_a^b (b - \xi)^n f^{(n+1)}(\xi) d\xi$, получить оценки погрешности формул численного дифференцирования (постоянные C_1, C_2 не зависят от f и h):

- 1) $|\bar{\partial}f(x) - f'(x)| \leq C_1 \int_{x-h}^x |f''(\xi)| d\xi$;
- 2) $|\bar{\partial}\bar{\partial}f(x) - f''(x)| \leq C_2 h \int_{x-h}^{x+h} |f^{(4)}(\xi)| d\xi$.

5.9. Доказать справедливость следующих равенств:

$$\partial(fg) = f\partial g + g\partial f + h\partial f\partial g, \quad \bar{\partial}(f/g) = \frac{g\bar{\partial}f - f\bar{\partial}g}{g(g - h\bar{\partial}g)}.$$

5.10. Пусть вычислены точное $f''(x_0)$ и приближенное $L_n''(x_0)$ значения при заданных узлах интерполяции $x_{-l}, \dots, x_0, \dots, x_l$, $x_i - x_{i-1} = h$. Показать, что справедливо представление

$$f''(x_0) - L_n''(x_0) = \frac{2(-1)^l(l!)^2}{(2l+2)!} f^{(2l+2)}(\xi) h^{2l}, \quad x_{-l} \leq \xi \leq x_l.$$

5.11. Используя формулу Тейлора с остаточным членом в интегральной форме, получить оценки погрешности следующих формул численного дифференцирования (постоянные C_i , $1 \leq i \leq 5$, не зависят от f и h):

$$1) \quad |\partial f(x) - f'(x)| \leq C_1 \int_x^{x+h} |f''(\xi)| d\xi;$$

$$2) \quad |\tilde{\partial} f(x) - f'(x)| \leq C_2 h \int_{x-h}^{x+h} |f'''(\xi)| d\xi;$$

$$3) \quad |2\partial f(x) - \tilde{\partial} f(x+h) - f'(x)| \leq C_3 h \int_x^{x+2h} |f'''(\xi)| d\xi;$$

$$4) \quad |2\tilde{\partial} f(x) - \tilde{\partial} f(x-h) - f'(x)| \leq C_4 h \int_{x-2h}^x |f'''(\xi)| d\xi;$$

$$5) \quad |\bar{\partial}^2 \partial^2 f(x) - f^{(4)}(x)| \leq C_5 h \int_{x-2h}^{x+2h} |f^{(6)}(\xi)| d\xi.$$

5.12. Доказать справедливость следующих равенств:

$$1) \quad \bar{\partial}(fg) = f \bar{\partial}g + g \bar{\partial}f - h \bar{\partial}f \bar{\partial}g;$$

$$2) \quad \tilde{\partial}(fg) = f \tilde{\partial}g + g \tilde{\partial}f + \frac{h^2}{2} (\bar{\partial}\partial f \tilde{\partial}g + \bar{\partial}\partial g \tilde{\partial}f);$$

$$3) \quad \partial(f/g) = [g \partial f - f \partial g] / [g(g + h \partial g)].$$

5.13. Получить приближение к $f'(0)$ наиболее высокого порядка точности по h следующего вида:

$$1) \quad h^{-1}[a f(0) + b f(h) + c f(2h)];$$

$$2) \quad h^{-1}[a f(0) + b f(-h) + c f(2h)];$$

$$3) \quad h^{-1}[a f(0) + b f(-h) + c f(-2h)];$$

$$4) \quad h^{-1}[a f(0) + b f(2h) + c f(3h)].$$

Найти h , при котором достигается минимум оценки полной погрешности, если $\max_x |f^{(k)}(x)| \leq A_k$, и абсолютная вы-

числительная погрешность функции не превосходит Δ , т. е.

$$\max_x |f(x) - f^*(x)| \leq \Delta.$$

5.14. Пусть числа α_i , не зависящие от h , порождают формулу численного дифференцирования максимального порядка точности среди формул вида $f^{(k)}(x) \approx h^{-k} \sum_{i=-n}^n \alpha_i f(x + ih)$. Доказать, что справедливы утверждения:

- 1) $\alpha_i = \alpha_{-i}$, если k четно, $\alpha_i = -\alpha_{-i}$, если k нечетно;
- 2) формула с дополнительным слагаемым

$$f^{(k)}(x) \approx h^{-k} \sum_{i=-n}^{n+1} \beta_i f(x + ih)$$

не может иметь больший порядок точности; причем она имеет тот же порядок точности тогда и только тогда, когда $\beta_{n+1} = 0$, $\beta_i = \alpha_i$, $i = -n, -n + 1, \dots, n - 1, n$.

5.15. Доказать, что если все точки x_i различны и удалены от точки x_0 на расстояние $O(h)$, где h — малая величина, то при гладкой $f(x)$ приближенная формула численного дифференцирования

$f^{(k)}(x_0) \approx \sum_{i=1}^n c_i f(x_i)$ имеет порядок погрешности $O(h^m)$. Здесь $m \geq j + 1 - k$, j — максимальная степень многочленов, для которых эта формула точна.

5.16. Построить формулу численного дифференцирования для $f''(x_i)$ с равноотстоящими (т. е. $x_{i+1} - x_i = h$) узлами $x_i, x_{i\pm 1}, x_{i\pm 2}$, имеющую максимально возможный порядок точности по h .

5.17. Найти коэффициенты формул численного дифференцирования, приводящие к максимально возможному порядку точности по h :

- 1) $f'(x) \approx (af(x) + bf(x+h) + cf(x-h))/h$;
- 2) $f'(x) \approx (af(x) + bf(x+h) + cf(x-2h))/h$;
- 3) $f''(x) \approx (af(x) + bf(x+h) + cf(x+2h))/h^2$;
- 4) $f''(x) \approx (af(x) + bf(x+h) + cf(x-h))/h^2$;
- 5) $f''(x) \approx (af(x) + bf(x-h) + cf(x-2h))/h^2$.

§ 6. Многочлен наилучшего равномерного приближения

Пусть \mathbb{R} — пространство ограниченных вещественных функций, определенных на отрезке $[a, b]$ вещественной оси, с нормой $\|f(x)\| = \sup_{x \in [a, b]} |f(x)|$. Для элемента $f \in \mathbb{R}$ отыскивается наилучшее

приближение вида $Q_n(x) = \sum_{j=0}^n a_j x^j$. Многочлен $Q_n^0(x)$ называется

многочленом наилучшего равномерного приближения для функции $f(x)$, если для любого многочлена $Q_n(x)$ степени n справедливо неравенство $\|f - Q_n^0\| \leq \|f - Q_n\|$. Такой многочлен существует всегда, а его единственность имеет место при дополнительном предположении о непрерывности $f(x)$.

Теорема Чебышёва. *Чтобы многочлен $Q_n(x)$ был многочленом наилучшего равномерного приближения непрерывной функции $f(x)$, необходимо и достаточно существования на $[a, b]$ по крайней мере $n + 2$ точек $x_0 < x_1 < \dots < x_{n+1}$, таких, что*

$$f(x_i) - Q_n(x_i) = \alpha(-1)^i \|f - Q_n\|,$$

где $i = 0, 1, \dots, n + 1$ и $\alpha = 1$ (или $\alpha = -1$) одновременно для всех i .

Точки x_0, \dots, x_{n+1} , удовлетворяющие условию теоремы, называются *точками чебышёвского альтернанса*.

- 6.1.** Построить многочлен наилучшего равномерного приближения степени $n = 50$ для $f(x) = \sin 100x$ на отрезке $[0, \pi]$.
- 6.2.** Пусть $f(x)$ — выпуклая непрерывная функция на отрезке $[a, b]$ и $Q_1(x)$ — ее многочлен наилучшего равномерного приближения первой степени. Доказать, что концы отрезка a и b входят в альтернанс.
- 6.3.** Построить многочлен наилучшего равномерного приближения степени $n = 1$ для $f(x) = x^3$ на отрезке $[1, 2]$.
- 6.4.** Построить многочлен наилучшего равномерного приближения степени $n = 1$ для $f(x) = |x|$ на отрезке $[-1, 5]$.
- 6.5.** Построить многочлен наилучшего равномерного приближения $Q_n(x)$ степени n для $P_{n+1}(x) = a_{n+1}x^{n+1} + \dots + a_0$ на отрезке $[a, b]$.

- 6.6.** Пусть $f^{(n+1)}(x)$ не меняет знак на отрезке $[a, b]$ и $Q_n(x)$ — многочлен наилучшего равномерного приближения степени n для $f(x)$. Оценить величины C_1 и C_2 в неравенстве $C_1 \leq \|f(x) - Q_n(x)\| \leq C_2$.
- 6.7.** Пусть $f(x)$ — непрерывная нечетная функция на отрезке $[-1, 1]$. Показать, что ее многочлен наилучшего равномерного приближения произвольной степени n — также нечетная функция.
- 6.8.** Получить оценку вида $C_n \leq \|\sin x - Q_n(x)\| \leq 2C_n$ для многочлена наилучшего равномерного приближения степени n на отрезке $[-\pi/3, \pi/3]$.
- 6.9.** Построить функцию $f(x)$ и ее многочлен наилучшего равномерного приближения $Q_n(x)$, не удовлетворяющие теоремам Чебышёва и единственности.
- 6.10.** Построить многочлен наилучшего равномерного приближения степени n для функции $f(x)$ на отрезке $[a, b]$:
- 1) $n = 1, f(x) = x^3, [-1, 1]$;
 - 2) $n = 3, f(x) = \exp(x^2), [-1, 1]$;
 - 3) $n = 3, f(x) = 3 \sin^2 10x + |x^2 - 7x + 10|, [3, 4]$.
- 6.11.** Построить многочлен наилучшего равномерного приближения первой степени для функции $f(x) = \sqrt{x^2 + 1}$ на отрезке $[0, 1]$.
- 6.12.** Привести пример непрерывной функции и ее многочлена наилучшего равномерного приближения шестой степени, для которых имеется 99 точек альтернанса.
- 6.13.** Построить многочлен наилучшего равномерного приближения степени n для функции $f(x)$ на отрезке $[a, b]$:
- 1) $n = 2, f(x) = x^3, a = 0, b = 1$;
 - 2) $n = 2, f(x) = x^4, a = -1, b = 1$;
 - 3) $n = 1, f(x) = \sin x, a = -\pi, b = \pi$;
 - 4) $n = 3, f(x) = |x^2 - 7x + 10|, a = 3, b = 4$;
 - 5) $n = 30, f(x) = 2x^2 + 3x + \cos 50x, a = 0, b = \pi$;
 - 6) $n = 1, f(x) = 1 + x^p, p > 0, a = 0, b = 1$;
 - 7) $n = 4, f(x) = \sin(6\pi x), a = 0, b = \pi$;
 - 8) $n = 2, f(x) = 2x^2 + 3x + 5, a = 1, b = 7$.

- 6.14.** Получить оценку $C_4/2 \leq \| \cos x - Q_4(x) \| \leq C_4$ с явным выражением для C_4 , где $Q_4(x)$ — многочлен наилучшего равномерного приближения четвертой степени на отрезке $[\pi/6, \pi/2]$.
- 6.15.** Пусть $Q_4(x)$ — многочлен наилучшего равномерного приближения четвертой степени для функции $\exp(x)$ на отрезке $[0, 1]$. Показать, что $\| \exp(x) - Q_4(x) \| \geq 1/64000$.
- 6.16.** Пусть $Q_6(x)$ — многочлен наилучшего равномерного приближения шестой степени для функции $\exp(x)$ на отрезке $[-1, 1]$. Показать, что $10^{-6} \leq \| \exp(x) - Q_6(x) \| \leq 10^{-5}$.
- 6.17.** Показать, что чебышёвский альтернанс для функции $\exp(x)$ всегда содержит крайние точки отрезка, на котором решается задача наилучшего равномерного приближения.
- 6.18.** Привести пример функции и соответствующего ей многочлена наилучшего равномерного приближения, для которых среди точек чебышёвского альтернанса нет граничных точек отрезка, на котором решается задача приближения.
- 6.19.** Пусть $\sum_{k=0}^{\infty} a_k T_k(x)$ — некоторый сходящийся ряд, где $T_k(x)$ есть многочлен Чебышёва k -й степени. Доказать, что каждая частичная сумма ряда $S_n(x) = \sum_{k=0}^n a_k T_k(x)$ — многочлен наилучшего равномерного приближения степени n на отрезке $[-1, 1]$ для функции $S_{n+1}(x)$.
- 6.20.** Функция $f(x) = 1/(x + 9)$ приближается на отрезке $[-1, 1]$ многочленом первой степени следующими способами:
- 1) наилучшее равномерное приближение;
 - 2) отрезок ряда Тейлора в точке $x = 0$;
 - 3) интерполяция с чебышёвскими узлами $x_{1,2} = \pm 2^{-1/2}$.
- Построить эти многочлены и вычислить равномерные нормы погрешностей.
- 6.21.** Функция $f(x) = \exp(-x)$ приближается на $[-1, 1]$ многочленом первой степени следующими способами:
- 1) наилучшее равномерное приближение;
 - 2) наилучшее приближение в $L_2(-1, 1)$, где $\|f\|_{L_2(-1,1)}^2 = \int_{-1}^1 |f(x)|^2 dx$;

- 3) отрезок ряда Тейлора в точке $x = 0$;
- 4) интерполяция с узлами $x_1 = -1, x_2 = 1$;
- 5) интерполяция с чебышёвскими узлами $x_{1,2} = \pm 2^{-1/2}$.

Построить эти многочлены и вычислить равномерные нормы погрешностей.

- 6.22.** Найти многочлен наилучшего равномерного приближения степени $n = 1$ для функции $f(x) = 1 + \sqrt{x}$ на отрезке $[0, 1]$.
- 6.23.** Найти многочлен наилучшего равномерного приближения степени $n = 3$ для функции $f(x) = \sin x^2$ на отрезке $[-\sqrt{\pi}, \sqrt{\pi}]$.
- 6.24.** Найти многочлен наилучшего равномерного приближения степени $n = 1$ для функции $f(x) = |x|$ на отрезке $[-1, 2]$.
- 6.25.** Найти многочлен наилучшего равномерного приближения степени $n = 2$ для функции $f(x) = x^3$ на отрезке $[-1, 1]$.
- 6.26.** Найти для функции $\exp(x)$ наилучшее приближение многочленом нулевой степени в $L_1(0, 1)$, где $\|f\|_{L_1(0,1)} = \int_0^1 |f(x)| dx$.
- 6.27.** Пусть P_2 — пространство алгебраических многочленов второй степени на отрезке $[-1, 1]$ с нормой $\|p(x)\| = |p(-1)| + |p(0)| + |p(1)|$. Найти наилучшее приближение константой для функции $p(x) = x^2 \in P_2$.
- 6.28.** Пусть $n \geq 1$ и заданы $(x_k, y_k), k = 0, 1, \dots, n$. Найти линейную функцию $p(x) = ax + b$, минимизирующую функционал $\sum_{k=0}^n (y_k - ax_k - b)^2$.
- 6.29.** Пусть A и \mathbf{x} — вещественные симметричная матрица размерности $n \times n$ и n -мерный вектор, $f(t) = \|\mathbf{Ax} - t\mathbf{x}\|_2 \equiv \sqrt{(\mathbf{Ax} - t\mathbf{x}, \mathbf{Ax} - t\mathbf{x})}$. Доказать, что $f(t)$ достигает минимума при $t = (\mathbf{Ax}, \mathbf{x})/(\mathbf{x}, \mathbf{x})$.
- 6.30.** Найти наилучшее приближение в $L_2(a, b)$ для функции $f(x)$ многочленами $Q_n(x)$ степени n :
- 1) $a = -1, b = 1; f(x) = |x|; n = 1$;
 - 2) $a = -1, b = 1; f(x) = x^2; n = 1$;
 - 3) $a = -1, b = 1; f(x) = x^3; n = 1$;
 - 4) $a = -1, b = 1; f(x) = x^3; n = 2$;

5) $a = 0, b = \pi; f(x) = \sin x; n = 2;$

6) $a = 0, b = 2; f(x) = x^3; n = 3.$

6.31. Для заданной функции $f(x)$ найти многочлен $Q_n(x)$ степени n , минимизирующий функционал $\int_{-1}^1 \frac{(f(x) - Q_n(x))^2}{\sqrt{1-x^2}} dx$, в следующих случаях:

1) $f(x) = \frac{1}{4}(x^2 + 2x + 1), n = 1;$ 2) $f(x) = x^2, n = 1;$

3) $f(x) = x^3, n = 2.$

6.32. Показать, что построение многочлена наилучшего приближения степени $(n - 1)$ для функции $f(x)$ в пространстве $L_2(0, 1)$ приводит к системе уравнений с матрицей Гильберта: $h_{ij} = 1/(i + j - 1), 1 \leq i, j \leq n.$

§ 7. Приближение сплайнами

Пусть на отрезке $[a, b]$ заданы: узлы $a = x_0 < x_1 < \dots < x_n = b$, $P_m(x)$ — множество многочленов степени не выше m ($m \geq 1$), $C^{(r)}[a, b]$ — множество функций, имеющих на $[a, b]$ непрерывные производные до r -го порядка включительно ($r \geq 0$).

Функцию $S_m(x) = S_{m,k}(x)$ называют полиномиальным сплайном степени m дефекта k ($1 \leq k \leq m$) с узлами $x_i, i = 0, 1, \dots, n$, для функции $f(x) \in C[a, b]$, если выполнены следующие условия:

- 1) на каждом из отрезков $[x_i, x_{i+1}], i = 0, 1, \dots, n - 1$, она является многочленом, т. е. $S_m(x) \in P_m(x)$;
- 2) на всем отрезке $[a, b]$ обладает непрерывностью производных, т. е. $S_m(x) \in C^{(m-k)}[a, b]$.

В дальнейшем термин «дефекта k » будем опускать, так как в параграфе рассматривается только случай $k = 1$.

Сплайн называется интерполяционным, если в узлах x_i справедливо равенство $S_m(x_i) = f(x_i), i = 0, 1, \dots, n$.

Далее равномерной сеткой с шагом h будем называть набор узлов $x_i = a + ih, 0 \leq i \leq n, nh = b - a$.

7.1. Построить линейный интерполяционный сплайн по значениям $f(0), f(1)$.

7.2. Получить оценки погрешности приближения функции $f(x)$ линейным интерполяционным сплайном на равномерной сетке с шагом h , если $f(x) \in C^{(2)}[0, 1]$.

7.3. Обозначим через M_i значения второй производной $S_3''(x)$ кубического интерполяционного сплайна в узлах x_i , $i = 0, 1, \dots, n$. Показать, что они удовлетворяют системе линейных уравнений $AM = d$, где

$$a_{ij} = \begin{cases} \frac{h_i}{6} & \text{при } j = i - 1, \\ \frac{h_i + h_{i+1}}{3} & \text{при } j = i, \\ \frac{h_{i+1}}{6} & \text{при } j = i + 1, \\ 0 & \text{при } |j - i| > 1; \end{cases} \quad d_i = \frac{f_{i+1} - f_i}{h_{i+1}} - \frac{f_i - f_{i-1}}{h_i},$$

$$i = 1, 2, \dots, n - 1, \quad h_i = x_i - x_{i-1}.$$

7.4. Построить кубический интерполяционный сплайн по значениям $f(0)$, $f(1)$, $f(2)$.

7.5. Пусть дополнительно к условию **7.3** задано $M_0 = M_n = 0$. Показать, что в этом случае решение системы $AM = d$ удовлетворяет неравенству

$$\max_{1 \leq i \leq n-1} |M_i| \leq 3 \max_{1 \leq i \leq n-1} |d_i| / \min_{1 \leq i \leq n-1} h_i.$$

7.6. Пусть $f(x) \in C^{(4)}[a, b]$, $\max_{[a, b]} |f^{(4)}(x)| \leq A_4$, задана равномерная сетка с шагом h и дополнительные условия для определения кубического интерполяционного сплайна имеют следующий вид:

$$S_3'(x_0 + 0) = f'(x_0), \quad S_3'(x_n - 0) = f'(x_n).$$

Показать, что справедливы оценки погрешностей

$$|S_3^{(l)}(x) - f^{(l)}(x)| \leq K_l A_4 h^{4-l}, \quad l = 0, 1, 2, 3,$$

с постоянными K_l , не зависящими от h .

7.7. На равномерной сетке с шагом h построены естественные сплайны $S_3(x)$ и $S_3^*(x)$ при использовании точных f_i и приближенных f_i^* значений функции, так что $|f_i - f_i^*| \leq \varepsilon$. Показать

справедливость оценки

$$\max_x |S_3(x) - S_3^*(x)| \leq 10\varepsilon.$$

Используют также *локальные (аппроксимационные) сплайны*, значения которых в узлах, как правило, не совпадают со значениями $f(x)$. Это обстоятельство не принципиально, так как значения $f(x)$ обычно известны приближенно. Рассмотрим построение локального сплайна третьей степени на равномерной сетке с шагом h для отрезка $[0, 1]$. Возьмем *стандартный* сплайн $B(x)$, определяемый соотношениями

$$B(x) = \begin{cases} \frac{2}{3} - x^2 + \frac{1}{2}|x|^3 & \text{при } |x| \leq 1, \\ \frac{1}{6}(2 - |x|)^3 & \text{при } 1 \leq |x| \leq 2, \\ 0 & \text{при } 2 \leq |x|. \end{cases}$$

Локальные сплайны третьей степени $B_2^{(1)}(x)$ и $B_2^{(2)}(x)$ записываются в одинаковом виде

$$B_2^{(k)}(x) = \sum_{i=-1}^{n+1} \alpha_i^{(k)} B\left(\frac{x - ih}{h}\right), \quad k = 1, 2,$$

но отличаются выбором коэффициентов.

При $k = 1$ доопределяют значения f_{-1} и f_{n+1} линейной интерполяцией по значениям f_0, f_1 и f_n, f_{n-1} , соответственно, и полагают $\alpha_i = f_i$ (где $f_i = f(x_i)$) при $-1 \leq i \leq n+1$.

При $k = 2$ доопределяют значения f_{-2}, f_{-1} и f_{n+1}, f_{n+2} кубической интерполяцией по значениям f_0, f_1, f_2, f_3 и $f_n, f_{n-1}, f_{n-2}, f_{n-3}$, соответственно, и полагают $\alpha_i = (8f_i - f_{i+1} - f_{i-1})/6$.

В задачах **7.8–7.14** сетка предполагается равномерной с шагом h .

- 7.8.** Показать, что при любых $\alpha_i^{(k)}$, $k = 1, 2$, функции $B_2^{(k)}(x)$ являются сплайнами третьей степени, причем они тождественно равны нулю вне отрезка $[-3h, 1 + 3h]$.
- 7.9.** Вывести формулы для значений f_{-1} и f_{n+1} , необходимых для определения локального сплайна $B_2^{(1)}(x)$.
- 7.10.** Вывести формулы для значений f_{-2}, f_{-1} и f_{n+1}, f_{n+2} , необходимых для определения локального сплайна $B_2^{(2)}(x)$.

7.11. Показать, что величина $B_2^{(1)}(x)$ зависит только от значений f_i в четырех ближайших к x точках x_i , а величина $B_2^{(2)}(x)$ — в шести точках.

7.12. Показать справедливость следующих равенств:

$$\begin{aligned} B_2^{(i)}(x_0) &= f_0, & B_2^{(i)}(x_n) &= f_n, & i &= 1, 2; \\ B_2^{(2)}(x_1) &= f_1, & B_2^{(2)}(x_{n-1}) &= f_{n-1}. \end{aligned}$$

7.13. Пусть $f(x) \in C^{(4)}[a, b]$, $\max_{[a, b]} |f^{(4)}(x)| \leq A_4$ и задана равномерная сетка с шагом h . Показать справедливость следующих оценок погрешностей с постоянными K_l , не зависящими от h :


$$\left| \left(B_2^{(2)}(x) \right)^{(l)} - f^{(l)}(x) \right| \leq K_l A_4 h^{4-l}, \quad l = 0, 1, 2, 3.$$

7.14. Пусть $f(x) \in C^{(2)}[a, b]$, $\max_{[a, b]} |f''(x)| \leq A_2$ и задана равномерная сетка с шагом h . Показать справедливость следующих оценок погрешностей с постоянными K_l , не зависящими от h :

$$\left| \left(B_2^{(1)}(x) \right)^{(l)} - f^{(l)}(x) \right| \leq K_l A_2 h^{2-l}, \quad l = 0, 1.$$

Глава III

Численное интегрирование



Рассмотрим интеграл вида

$$I(f) = \int_{\Omega} p(x) f(x) dx,$$

где Ω — конечный или бесконечный промежуток числовой оси и $f(x)$ — произвольная функция из некоторого класса F . В первую очередь под Ω будем иметь в виду отрезок $[a, b]$ и, если не оговорено противное, то считать, что $f(x)$ непрерывна на нем. Заданная функция $p(x)$ называется *весовой*. Будем предполагать, что на $[a, b]$ она измерима, тождественно не равна нулю (как правило, почти всюду положительна) и ее произведение на любую $f(x) \in F$ суммируемо.

Для приближенного вычисления интеграла $I(f)$ строят линейные квадратурные формулы (*квадратуры*) следующего вида:

$$S_n(f) = \sum_{i=1}^n c_i f(x_i).$$

Постоянные c_i называются *коэффициентами (весами) квадратуры*, x_i — ее *узлами*.

Для каждой функции $f(x) \in F$ погрешность квадратурной формулы $S_n(f)$ определяется как $R_n(f) = I(f) - S_n(f)$. При этом оценкой погрешности на классе F называют величину

$$R_n(F) = \sup_{f \in F} |R_n(f)|, \quad \|R_n(F)\| = \sup_{f \in F, \|f\|_F \neq 0} |R_n(f)| / \|f\|_F.$$

На практике часто используют оценки сверху для $|R_n(f)|$, которые будем обозначать через R_n .

§ 8. Квадратурные формулы интерполяционного типа

Имеется большая группа квадратурных формул, построенных на основе замены $f(x)$ алгебраическим интерполяционным многочленом. Пусть на отрезке $[a, b]$ по заданному набору различных узлов

$x_i, 1 \leq i \leq n$, функция $f(x)$ приближается интерполяционным многочленом Лагранжа $L_n(x)$ степени $n - 1$

$$L_n(x) = \sum_{i=1}^n f(x_i) \prod_{\substack{j=1 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}.$$

Положим

$$S_n(f) = \int_a^b p(x) L_n(x) dx.$$

Отсюда получаем явные формулы для коэффициентов $c_i, 1 \leq i \leq n$, и оценку погрешности R_n , такую, что $|I(f) - S_n(f)| \leq R_n$:

$$c_i = \int_a^b p(x) \prod_{\substack{j=1 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j} dx, \quad R_n = \frac{\|f^{(n)}(x)\|}{n!} \int_a^b |p(x)| |\omega_n(x)| dx,$$

где $\|f^{(n)}(x)\| = \max_{x \in [a, b]} |f^{(n)}(x)|, \omega_n(x) = \prod_{i=1}^n (x - x_i)$. В оценках, приводимых ниже, также используется равномерная норма.

Квадратурные формулы интерполяционного типа, построенные в случае весовой функции $p(x) \equiv 1$ для n равноотстоящих узлов: $x_1 = (a + b)/2$ при $n = 1$ и $x_i = a + (i - 1)(b - a)/(n - 1), 1 \leq i \leq n$ при $n > 1$, называются *формулами Ньютона—Котеса*.

- 8.1.** Получить формулы Ньютона—Котеса и соответствующие оценки погрешностей при числе узлов $n = 1, 2, 3$.
- 8.2.** Рассмотреть формулы Ньютона—Котеса при $n = 1$ (прямоугольников) и $n = 2$ (трапеций) и сравнить оценки их погрешностей в случае гладких подынтегральных функций.
- 8.3.** Пусть весовая функция $p(x)$ четная, узлы x_i расположены симметрично относительно нуля, т. е. $x_{n+1-i} = -x_i, i = 1, \dots, n$. Доказать, что в интерполяционной квадратурной формуле для вычисления интеграла $I(f) = \int_{-a}^a p(x) f(x) dx$ коэффициенты, соответствующие симметричным узлам, равны, т. е. $c_{n+1-i} = c_i, i = 1, \dots, n$.

8.4. Доказать, что для погрешности квадратурной формулы трапеций справедливо представление

$$R_2(f) = \int_a^b f(x) dx - \frac{b-a}{2} (f(a) + f(b)) = \frac{1}{2} \int_a^b (a-\xi)(b-\xi) f''(\xi) d\xi.$$

8.5. Вычислить интеграл $\int_0^1 \exp(x^2) dx$ по формуле Ньютона–Котеса с узлами $0, 1/4, 1/2, 3/4, 1$ и оценить погрешность.

Составные квадратурные формулы. Рассмотрим задачи на построение составных квадратурных формул и вывод оценок их погрешностей. Пусть $h = (b-a)/N$ и $x_k = a + kh$, $k = 0, 1, \dots, N$.

Введем следующие обозначения: $I^{(k)}(f) = \int_{x_k}^{x_{k+1}} p(x) f(x) dx$,

$S_n^{(k)}(f) = S_n(f)$ для отрезка $[x_k, x_{k+1}]$, $k = 0, \dots, N-1$.

Исходный интеграл равен сумме: $I(f) = \sum_{k=0}^{N-1} I^{(k)}(f)$, поэтому со-

ставную квадратурную формулу строим в виде $S_n^N(f) = \sum_{k=0}^{N-1} S_n^{(k)}(f)$;

тогда для ее погрешности будет справедливо неравенство

$|R_n^N(f)| \leq \sum_{k=0}^{N-1} |R_n^{(k)}(f)|$. Например, в случае составной формулы прямоугольников

$$S_1^N(f) = \frac{b-a}{N} \sum_{k=0}^{N-1} f\left(x_k + \frac{h}{2}\right)$$

для погрешности на отрезке $[x_k, x_{k+1}]$ имеем неравенство

$$\left| R_1^{(k)}(f) \right| \leq \|f''(x)\|_{[x_k, x_{k+1}]} \frac{(x_{k+1} - x_k)^3}{24} = \|f''(x)\|_{[x_k, x_{k+1}]} \frac{(b-a)^3}{24 N^3}.$$

Следовательно, для всего отрезка $[a, b]$ оценка погрешности получается суммированием по всем $k = 0, 1, \dots, N-1$:

$$R_1^N = \|f''(x)\|_{[a,b]} \frac{(b-a)^3}{24 N^2}.$$

8.6. Пусть для вычисления $\int_0^1 f(x) dx$ применяется составная формула трапеций. Оценить минимальное число разбиений N отрезка $[0, 1]$, обеспечивающее погрешность не более $0,5 \cdot 10^{-3}$ на следующих классах функций:

$$1) \|f''(x)\| \leq 1; \quad 2) \int_0^1 |f''(x)| dx \leq 1.$$

8.7. Для составной квадратурной формулы трапеций с шагом $h = (b - a)/N$, погрешность которой имеет вид

$$R_2^N(f) = \int_a^b f(x) dx - \frac{b-a}{N} \left(\frac{1}{2} f(x_0) + \frac{1}{2} f(x_N) + \sum_{k=1}^{N-1} f(x_k) \right),$$

получить оценки погрешности

$$1) R_2^N = \frac{h^2}{8} \int_a^b |f''(x)| dx; \quad 2) R_2^N = \frac{h^2}{2} \sqrt{\frac{b-a}{30}} \left(\int_a^b |f''(x)|^2 dx \right)^{1/2}.$$

Найти оценку погрешности вычисления интеграла $I(f)$ для $f(x) = 1/(1+x^2)$, $p(x) \equiv 1$ по составной квадратурной формуле

$$\mathbf{8.8.} \quad S(f) = [f(0) + 4f(0,1) + 2f(0,2) + 4f(0,3) + \dots + 2f(0,8) + 4f(0,9) + f(1,0)]/30.$$

$$\mathbf{8.9.} \quad S(f) = [f(0) + 2f(0,1) + 2f(0,2) + \dots + 2f(0,9) + f(1,0)]/20.$$

8.10. Оценить минимальное число разбиений N отрезка $[0, 1]$ для вычисления интеграла $I(f)$ при $f(x) = \sin(x^2)$, $p(x) \equiv 1$ по составной квадратурной формуле трапеций, обеспечивающее погрешность не более 10^{-4} .

8.11. Оценить минимальное число разбиений N отрезка $[0, 1]$ для вычисления интеграла $I(f)$ при $f(x) = \exp(x^2)$, $p(x) \equiv 1$ по составной квадратурной формуле прямоугольников, обеспечивающее погрешность не более 10^{-4} .

8.12. Оценить минимальное число разбиений N отрезка $[0, 1]$ для вычисления интеграла $I(f)$ при $f(x) = \exp(x^2)$, $p(x) \equiv 1$ по составной квадратурной формуле трапеций, обеспечивающее погрешность не более 10^{-4} .

8.13. Оценить минимальное число разбиений N отрезка $[0, 2]$ для вычисления интеграла $I(f)$ при $p(x) \equiv 1$ по составной квадратурной формуле Симпсона, обеспечивающее погрешность не более $0,5 \cdot 10^{-4}$ на классе функций, удовлетворяющих условию $\|f^{(4)}(x)\| \leq 1$.

8.14. Предложить составную квадратурную формулу для вычисления интегралов: $I(f) = \int_0^{\infty} e^{-x} f(x) dx$, $I(f) = \int_0^{\infty} x e^{-x} f(x) dx$, с погрешностью не более 10^{-4} , если для некоторого фиксированного $k \geq 1$ выполнено $\|f^{(k)}(x)\| \leq 1$.

8.15. Доказать справедливость следующих представлений погрешностей квадратурных формул:

$$1) \int_a^b f(x) dx - \frac{b-a}{8} \left(f(a) + 3f\left(\frac{2a+b}{3}\right) + 3f\left(\frac{a+2b}{3}\right) + f(b) \right) = - \left(\frac{b-a}{3}\right)^5 \frac{3}{80} f^{(4)}(\xi), \quad a < \xi < b;$$

$$2) \int_a^b f(x) dx - \frac{b-a}{90} \left(7f(a) + 32f\left(\frac{3a+b}{4}\right) + 12f\left(\frac{a+b}{2}\right) + 32f\left(\frac{a+3b}{4}\right) + 7f(b) \right) = - \left(\frac{b-a}{4}\right)^7 \frac{8}{945} f^{(6)}(\xi), \quad a < \xi < b;$$

$$3) \int_a^b f(x) dx - \frac{b-a}{2} (f(a) + f(b)) + \frac{(b-a)^2}{12} (f'(a) - f'(b)) = \frac{(b-a)^5}{720} f^{(4)}(\xi), \quad a < \xi < b.$$

8.16. Доказать, что ни для какой системы узлов и коэффициентов погрешность квадратурной формулы $R_n(f)$ не стремится сильно к нулю на пространстве непрерывных функций (т. е. для $f(x) \in C[a, b]$). Более того, всегда справедливо равенство

$$\|R_n(C)\| = \int_a^b p(x) dx + \sum_{i=1}^n |c_i|.$$

8.17. Пусть $A_l = \int_a^b |f^{(l)}(x)| dx < \infty$, $l = 1, 2$. Получить оценку погрешности составной формулы трапеций $|R_2^N(f)| \leq K_l A_l h^l$,

где K_l — абсолютная постоянная, $h = (b - a)/N$ — шаг интегрирования.

- 8.18.** Пусть $A_l = \int_a^b |f^{(l)}(x)| dx < \infty$, $l = 1, 2, 3, 4$. Получить оценку погрешности составной формулы Симпсона $|R_3^N(f)| \leq K_l A_l h^l$, где K_l — абсолютная постоянная, $h = (b - a)/N$ — шаг интегрирования.

В следующих упражнениях рассматривается приближенное вычисление интеграла $I(f_b) = \int_0^1 f_b(x) dx$ от функции с параметром $|b| < 1$:

$$f_b(x) = \begin{cases} 0 & \text{при } x = 0, \\ x^b & \text{при } x \in (0, 1]. \end{cases}$$

- 8.19.** Интеграл $I(f_b)$ вычисляется по составной квадратурной формуле трапеций с постоянным шагом $1/N$. Доказать, что суммарная погрешность удовлетворяет соотношению $|R_2^N(f)| \leq D_1(b)/N^{1+b}$, $D_1(b) \neq 0$.
- 8.20.** Интеграл $I(f_b)$ вычисляется по составной квадратурной формуле трапеций с распределением узлов по закону $x_k = \varphi(k/N)$, где $\varphi(t) = t^{3/(1+b)}$. Доказать, что суммарная погрешность удовлетворяет соотношению $|R_2^N(f)| \leq D_2(b)/N^2$, $D_2(b) \neq 0$.
- 8.21.** Интеграл $I(f_b)$ вычисляется по составной квадратурной формуле трапеций с распределением узлов $x_k = \varphi(k/N)$, $\varphi(t) = t^a$. Доказать, что при $a > 2/(b + 1)$ суммарная погрешность удовлетворяет соотношению $|R_2^N(f)| \leq D(a, b)/N^2$, $D(a, b) \neq 0$. Проверить, что $D(a, b) > D_2(b)$, где $D_2(b)$ определяется в **8.20**.

§ 9. Метод неопределенных коэффициентов

Если интегралы $I(x^k) = \int_a^b p(x)x^k dx$ вычисляются просто, то при заданном наборе различных узлов можно найти коэффициенты c_i из условия точности квадратурной формулы $S_n(f) = \sum_{i=1}^n c_i f(x_i)$ для произвольного многочлена наиболее высокой степени, т. е. из

равенств $I(x^k) = S_n(x^k)$, $k = 0, 1, \dots, (n - 1)$. Такая система линейных уравнений относительно c_i , $1 \leq i \leq n$, имеет единственное решение.

Если квадратура точна для многочлена степени m (говорят, что она имеет *алгебраический порядок точности*, равный m), то справедливо равенство $R_n(f) = R_n(f - P_m)$. Взяв в качестве $P_m(x)$ интерполяционный многочлен для $f(x)$, построенный по нулям многочлена Чебышёва, можно получить оценку

$$|R_n(f)| \leq \frac{\|f^{(m+1)}(x)\|}{(m+1)!} \frac{(b-a)^{m+1}}{2^{2m+1}} \left(\int_a^b |p(x)| dx + \sum_{i=1}^n |c_i| \right).$$

Из условия точности квадратурной формулы для функций заданного вида можно выписать уравнения (в общем случае нелинейные) не только для определения коэффициентов, но и для узлов квадратуры.

Квадратурными формулами Чебышёва называются квадратуры с одинаковыми коэффициентами, т. е.

$$S_n(f) = c \sum_{i=1}^n f(x_i), \quad c = \frac{1}{n} \int_a^b p(x) dx.$$

Их построение заключается в нахождении узлов x_i , $1 \leq i \leq n$, из условия точности квадратуры для многочлена максимально высокой степени. Квадратурные формулы Чебышёва (их удастся построить при $n = 1, 2, 3, 4, 7, 10$) обычно применяют, если значения $f(x_i)$ известны с независимыми случайными погрешностями, имеющими одинаковую дисперсию и нулевое математическое ожидание. В этом случае выбор одинаковых коэффициентов c обеспечивает минимальную дисперсию приближенного значения $S_n(f)$.

- 9.1.** Получить формулу Симпсона методом неопределенных коэффициентов.
- 9.2.** Для формул трапеций и Симпсона найти оценки погрешности, следующие из метода неопределенных коэффициентов.
- 9.3.** Для вычисления интегралов $I(f)$:

$$1) \int_0^2 (x+1)f(x) dx; \quad 2) \int_{-1}^0 x^2 f(x) dx; \quad 3) \int_{-1}^1 x^2 f(x) dx$$

построить формулу вида $S(f) = c_1 f(\tilde{x}) + c_2 f(x_2)$ с одним фиксированным узлом $\tilde{x} = 0$, точную для многочленов максимально высокой степени.

9.4. Для многочлена

$$P_n(x) = (x - x_1) \dots (x - x_n) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \dots + a_n$$

доказать, что величины $B_j = \sum_{k=1}^n x_k^j$, $1 \leq j \leq n$, удовлетворяют равенствам

$$\begin{aligned} B_1 &= -a_1, \\ a_1 B_1 + B_2 &= -2a_2, \\ a_2 B_1 + a_1 B_2 + B_3 &= -3a_3, \\ &\dots \dots \dots \\ a_{n-1} B_1 + a_{n-2} B_2 + \dots + a_1 B_{n-1} + B_n &= -na_n. \end{aligned}$$

9.5. Построить квадратурные формулы Чебышёва на отрезке $[-1, 1]$ для интеграла $I(f)$ с весом $p(x) \equiv 1$ при $n = 2, 3, 4$.

9.6. Квадратурная формула $S_n(f) = \frac{\pi}{n} \sum_{j=1}^n f\left(\cos \frac{2j-1}{2n} \pi\right)$ при-

меняется для вычисления интеграла $I(f) = \int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx$.

Показать, что она точна для всех многочленов степени $2n - 1$.

9.7. Квадратурная формула

$$S_n(f) = \frac{\pi}{n+1} \sum_{j=1}^n \sin^2\left(\frac{\pi j}{n+1}\right) f\left(\cos \frac{\pi j}{n+1}\right)$$

применяется для вычисления интеграла

$$I(f) = \int_{-1}^1 f(x) \sqrt{1-x^2} dx.$$

Показать, что она точна для всех многочленов степени $2n - 1$.

9.8. Квадратурная формула $S_n(f) = \frac{\omega}{n} \sum_{j=0}^{n-1} f\left(\frac{j\omega}{n}\right)$ применяется

для вычисления интеграла $I(f) = \int_0^\omega f(x) dx$. Показать, что она

точна для всех тригонометрических многочленов с периодом ω степени $n - 1$.

- 9.9.** Пусть T — треугольник на плоскости, $s(T)$ — его площадь, A , B , C — середины сторон. Показать, что кубатурная формула

$$S(f) = \frac{1}{3} s(T) (f(A) + f(B) + f(C)) \approx \iint_T f(x) dx,$$

где $x = (x_1, x_2)$, $dx = dx_1 dx_2$, точна для всех многочленов второй степени вида $a_0 + a_1 x_1 + a_2 x_2 + a_{11} x_1^2 + a_{12} x_1 x_2 + a_{22} x_2^2$.

- 9.10.** Пусть P — прямоугольник на плоскости, $s(P)$ — его площадь, A , B , C , D — середины сторон, E — точка пересечения диагоналей. Показать, что кубатурная формула

$$S(f) = \frac{1}{6} s(P) (f(A) + f(B) + f(C) + f(D) + 2f(E)) \approx \iint_P f(x) dx,$$

где $x = (x_1, x_2)$, $dx = dx_1 dx_2$, точна для всех многочленов от двух переменных третьей степени.

- 9.11.** Для вычисления интегралов $I(f)$

$$1) \int_0^2 f(x) dx; \quad 2) \int_0^1 f(x) dx; \quad 3) \int_{-1}^0 f(x) dx; \quad 4) \int_{-2}^0 f(x) dx$$

построить квадратурную формулу Чебышёва с тремя узлами.

- 9.12.** Для вычисления интегралов $I(f)$

$$1) \int_{-2}^0 x^2 f(x) dx; \quad 2) \int_0^1 x f(x) dx;$$

$$3) \int_0^{\pi/2} \cos(x) f(x) dx; \quad 4) \int_0^2 (x+2) f(x) dx$$

построить квадратурную формулу вида $S(f) = c_1 f(0) + c_2 f(x_2)$, точную для многочленов максимально высокой степени.

- 9.13.** Определить параметры c_1 , c_2 , x_2 так, чтобы квадратурная формула $S(f) = c_1 f(a) + c_2 f(x_2)$ для вычисления интеграла $\int_a^b f(x) dx$ была точной на многочленах максимально высокой степени.

9.14. Определить параметры c_1, c_2, c_3, x_2 так, чтобы квадратурная формула $S(f) = c_1 f(-1) + c_2 f(x_2) + c_3 f(1)$ для вычисления интеграла $I(f) = \int_{-1}^1 x^2 f(x) dx$ была точной на многочленах максимально высокой степени.

9.15. Для вычисления интеграла $I(f) = \int_0^1 f(x) dx$ построить квадратурную формулу $S_2(f) = c_1 f(0) + c_2 f(2/3)$, точную для многочленов максимально высокой степени.

9.16. Для вычисления интеграла $I(f) = \int_0^1 f(x) dx$ построить квадратурную формулу $S_2(f) = c_1 f(1/2) + c_2 f(2/3)$, точную для многочленов максимально высокой степени.

9.17. Для вычисления интеграла $I(f) = \int_0^2 f(x) dx$ построить квадратурную формулу $S_3(f) = c_1 f(0) + c_2 f(1/2) + c_3 f(2)$, точную для многочленов максимально высокой степени.

9.18. Для вычисления интеграла $I(f) = \int_a^b e^{\alpha x} f(x) dx$ построить квадратурную формулу $S_2(f) = c_1 f(a) + c_2 f(b)$, точную для многочленов максимально высокой степени.

9.19. Для вычисления интеграла $I(f) = \int_a^b f(x) dx$ построить квадратурную формулу

$$S_4(f) = c_1 f(a) + c_2 f\left(a + \frac{b-a}{3}\right) + c_3 f\left(a + 2\frac{b-a}{3}\right) + c_4 f(b),$$

точную для многочленов максимально высокой степени.

9.20. Пусть $f \in C^{(1)}[-1, 1]$ и $P_5(x)$ — многочлен пятой степени, удовлетворяющий условиям $P(x_k) = f(x_k)$, $P'(x_k) = f'(x_k)$, $k = 1, 2, 3$, где $x_1 = -1, x_2 = 0, x_3 = 1$. Рассмотрим квадратурную формулу

$$S_5(f) = (7f(-1) + 16f(0) + 7f(1) + f'(-1) - f'(1))/15.$$

Проверить, что $S_5(f)$ точна на многочленах пятой степени $\int_{-1}^1 P_5(x) dx = S_5(P_5)$, но найдется многочлен степени шесть, на котором она не точна.

§ 10. Квадратурные формулы Гаусса

Рассмотрим следующую задачу: при заданном числе узлов n построить для вычисления интегралов вида $I(f) = \int_a^b p(x)f(x) dx$ квадратурную формулу Гаусса $S_n(f) = \sum_{i=1}^n c_i f(x_i)$, точную для многочленов максимально высокой степени. Весовая функция $p(x)$ предполагается почти всюду положительной.

В этой постановке имеется $2n$ свободных параметров (узлы и коэффициенты неизвестны), поэтому можно попытаться построить квадратуру, точную для многочленов степени $2n - 1$. Несложно убедиться в том, что не существует квадратуры с n узлами, точной для всех многочленов степени $2n$. Действительно, возьмем $P_{2n}(x) = (x - x_1)^2 \dots (x - x_n)^2$. Тогда $0 = S_n(P_{2n}) \neq I(P_{2n}) > 0$.

Важную роль при построении квадратурных формул Гаусса играют многочлены, ортогональные на отрезке $[a, b]$ с весом $p(x) > 0$. Эти многочлены могут быть получены, например, в результате стандартной процедуры ортогонализации, примененной к системе функций $\{1, x, \dots, x^k, \dots\}$, при скалярном произведении

$$(f, g) = \int_a^b p(x)f(x)g(x) dx.$$

Пусть на отрезке $[a, b]$ имеется система ортогональных с весом $p(x)$ многочленов: $1, \psi_1(x), \psi_2(x), \dots, \psi_k(x), \dots$. Тогда многочлен k -й степени $\psi_k(x)$ ортогонален с весом $p(x)$ произвольному многочлену $P_l(x)$ при $l = 0, \dots, k - 1$. Действительно, многочлен $P_l(x)$

представим в виде $P_l(x) = \sum_{j=0}^l c_j \psi_j(x)$, и при $k \neq l$ имеют место равенства $\int_a^b p(x)\psi_k(x)\psi_l(x) dx = 0$. Аналогичным свойством обладают и многочлены, ортогональные на бесконечном промежутке Ω .

На практике наиболее употребительны следующие *ортгоналные многочлены*:

- Лежандра ($\Omega = [-1, 1]$, $p(x) \equiv 1$),
- Чебышёва первого рода ($\Omega = [-1, 1]$, $p(x) = 1/\sqrt{1-x^2}$),
- Лагерра ($\Omega = [0, \infty)$, $p(x) = e^{-x}$),
- Эрмита ($\Omega = (-\infty, \infty)$, $p(x) = e^{-x^2}$).

При построении квадратурных формул Гаусса базовой является следующая

Теорема. Пусть x_1, \dots, x_n — нули ортогонального на $[a, b]$ с весом $p(x) > 0$ многочлена $\psi_n(x)$ степени n и $S_n(f)$ — интерполяционная квадратура, построенная по этим узлам. Тогда $S_n(f)$ будет точна для многочленов степени $2n - 1$.

На основании этого утверждения процесс построения квадратурных формул Гаусса может быть разбит на два последовательных этапа:

- 1) нахождение нулей ортогонального многочлена;
- 2) нахождение весов методом неопределенных коэффициентов.

Приведем оценку погрешности квадратурных формул Гаусса

$$R_n = \left\| f^{(2n)}(x) \right\| \int_a^b p(x) \frac{\psi_n^2(x)}{(2n)!} dx,$$

которая в случае отрезка $[-1, 1]$ и веса $p(x) \equiv 1$ принимает вид

$$R_n = \left\| f^{(2n)}(x) \right\| \frac{2^{2n+1}(n!)^4}{((2n)!)^3(2n+1)}.$$

10.1. Методом ортогонализации построить многочлены Лежандра со старшим коэффициентом 1, ортогональные на отрезке $[-1, 1]$ с весом $p(x) \equiv 1$.

10.2. Доказать, что ортогональный с весом $p(x) > 0$ на отрезке $[a, b]$ многочлен степени n имеет ровно n различных корней на этом отрезке.

10.3. Построить квадратуру Гаусса с одним узлом для вычисления интегралов:

$$1) I(f) = \int_0^1 xf(x) dx; \quad 2) I(f) = \int_0^1 e^x f(x) dx.$$

10.4. Построить квадратуру Гаусса с двумя узлами для вычисления интегралов:

$$1) I(f) = \int_{-1}^1 x^2 f(x) dx, \quad 2) I(f) = \int_{-\pi/2}^{\pi/2} \cos(x) f(x) dx.$$

10.5. Построить квадратуру Гаусса с тремя узлами для вычисления интеграла $I(f) = \int_{-1}^1 f(x) dx$.

10.6. Доказать, что все коэффициенты квадратуры Гаусса положительны.

10.7. Пусть весовая функция $p(x)$ — четная относительно середины отрезка интегрирования, т. е. точки $(a + b)/2$. Доказать, что узлы квадратуры Гаусса для вычисления интеграла $I(f) = \int_a^b p(x) f(x) dx$ расположены симметрично относительно $(a + b)/2$, а соответствующие симметричным узлам коэффициенты квадратуры равны.

10.8. На отрезке $[-1, 1]$ при весовой функции $p(x) = 1/\sqrt{1-x^2}$ построена квадратурная формула Гаусса с n узлами. Вычислить ее погрешность $R_n(f)$ для функции $f(x) = x^{2n}$ и показать, что $\lim_{n \rightarrow \infty} 2^{2n-1} |R_n(f)| = \pi$.

10.9. На отрезке $[-1, 1]$ при весовой функции $p(x) = \sqrt{1-x^2}$ построена квадратурная формула Гаусса с n узлами. Вычислить ее погрешность $R_n(f)$ для функции $f(x) = x^{2n}$ и показать, что $\lim_{n \rightarrow \infty} 2^{2n} |R_n(f)| = \pi$.

10.10. Пусть $f(x)$ — непрерывная на отрезке $[a, b]$ функция. Доказать, что для квадратурных формул Гаусса $|R_n(f)| \rightarrow 0$ при $n \rightarrow \infty$.

10.11. Построить составную квадратурную формулу Гаусса с двумя узлами на каждом отрезке разбиения для вычисления интеграла $I(f) = \int_a^b e^{\alpha x} f(x) dx$. Оценить погрешность построенной формулы.

10.12. Доказать, что не существует квадратурных формул $S_n(f)$ с n узлами, точных для всех тригонометрических многочленов степени n при весовой функции $p(x) \equiv 1$.

10.13. Построить квадратурную формулу Гаусса с одним узлом для вычисления интеграла $I(f)$: 1) $\int_0^1 x^2 f(x) dx$; 2) $\int_{-1}^1 |x| f(x) dx$.

10.14. Построить квадратурную формулу Гаусса с двумя узлами для вычисления интеграла $I(f)$:

$$1) \int_{-1}^1 |x| f(x) dx; \quad 2) \int_{-1}^1 x^4 f(x) dx.$$

10.15. Построить квадратурную формулу Гаусса с двумя узлами для вычисления интеграла $I(f) = \int_0^1 p(x) f(x) dx$:

$$1) p(x) = x; \quad 2) p(x) = \sin(\pi x); \quad 3) p(x) = e^x; \\ 4) p(x) = \cos\left(x - \frac{1}{2}\right); \quad 5) p(x) = 1 - x; \quad 6) p(x) = e^{-x}.$$

10.16. Квадратурная формула

$$S_3(f) = \sqrt{\pi}/6 \left(f\left(-\sqrt{3/2}\right) + 4f(0) + f\left(\sqrt{3/2}\right) \right)$$

применяется для вычисления интеграла

$$I(f) = \int_{-\infty}^{+\infty} \exp(-x^2) f(x) dx.$$

Показать, что она точна для всех многочленов пятой степени.

10.17. Квадратурная формула

$$S_3(f) = \frac{\pi}{3} \left(f\left(-\frac{\sqrt{3}}{2}\right) + f(0) + f\left(\frac{\sqrt{3}}{2}\right) \right)$$

применяется для вычисления интеграла $I(f) = \int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx$.

Показать, что она точна для всех многочленов пятой степени.

10.18. Построить квадратуру Гаусса с четырьмя узлами для вычисления интеграла $I(f) = \int_{-1}^1 f(x) dx$.

Построить ортогональный многочлен $\psi_3(x) = x^3 + \dots + a_0$ на промежутке Ω при заданной весовой функции $p(x)$.

10.19. $\Omega = (-\infty, \infty)$, $p(x) = \exp(-x^2)$.

10.20. $\Omega = [-1, 1]$, $p(x) = 1/\sqrt{1-x^2}$.

10.21. $\Omega = [-1, 1]$, $p(x) = \sqrt{1-x^2}$.

10.22. $\Omega = [0, \infty)$, $p(x) = \exp(-x)$.

Построить квадратурную формулу Гаусса с двумя узлами для вычисления интегралов

10.23. $I(f) = \int_0^{\pi} \sin(x)f(x) dx$. **10.24.** $I(f) = \int_0^{\infty} \exp(-x)f(x) dx$.

10.25. $I(f) = \int_0^1 \left(x - \frac{1}{2}\right)^2 f(x) dx$.

10.26. Доказать, что ни с каким весом $p(x) > 0$ многочлены $\{x^m\}_{m=0}^{\infty}$ не могут быть ортогональны на отрезке $[0, 1]$.

10.27. Пусть на отрезке $[a, b]$ имеется система ортогональных многочленов $\{\psi_m(x)\}_{m=0}^{\infty}$ с весом $p(x)$ и старшим коэффициентом, равным единице. Доказать, что среди всех многочленов степени n вида $P_n(x) = x^n + a_{n-1}x^{n-1} + \dots + a_0$ минимальную норму $\|P_n\|^2 = \int_a^b p(x)P_n^2(x) dx$ имеет ортогональный многочлен $\psi_n(x)$.

10.28. Пусть на отрезке $[a, b]$ имеется система ортогональных многочленов $\{\psi_m(x)\}_{m=0}^{\infty}$ с весом $p(x)$ и старшим коэффициентом, равным единице. Показать при $n \geq 2$ справедливость рекуррентного соотношения $\psi_n(x) = (x + b_n)\psi_{n-1}(x) - c_n\psi_{n-2}(x)$ с коэффициентом $c_n > 0$.

10.29. Доказать, что ортогональные многочлены на симметричном относительно нуля отрезке с четным весом $p(x)$ обладают свойством $\psi_n(-x) = (-1)^n\psi_n(x)$.

10.30. Пусть задан отрезок $[a, b]$. Доказать, что при $b > a \geq 0$ все коэффициенты ортогонального многочлена отличны от нуля.

10.31. Доказать, что нули ортогональных многочленов $\psi_n(x)$ с фиксированным на отрезке $[a, b]$ весом $p(x) > 0$ перемежаются, т. е. справедливо

$$a < x_1^{(n)} < x_1^{(n-1)} < \dots < x_{n-1}^{(n-1)} < x_n^{(n)} < b,$$

где $x_i^{(n)}$, $1 \leq i \leq n$ — нули $\psi_n(x)$.

10.32. Пусть x_1, x_2, \dots, x_n — нули многочлена Лежандра и постоянные $\gamma_i = \int_{-1}^1 \prod_{\substack{j=1 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j} dx$. Доказать, что если $f(x), g(x)$ —

многочлены степени $n-1$, то $\int_{-1}^1 f(x)g(x) dx = \sum_{i=1}^n \gamma_i f(x_i)g(x_i)$.

10.33. Для вычисления интеграла $\int_{-1}^1 f(x) dx$ построить *квадратурную формулу Маркова—Радона*

$$S_n(f) = c_1 f(-1) + \sum_{i=2}^n c_i f(x_i),$$

точную для произвольного многочлена степени $2n-2$.

10.34. Для вычисления интеграла $\int_{-1}^1 f(x) dx$ построить *квадратурную формулу Маркова—Лобатто*

$$S_n(f) = c_1 f(-1) + \sum_{i=2}^{n-1} c_i f(x_i) + c_n f(1),$$

точную для произвольного многочлена степени $2n-3$.

§ 11. Главный член погрешности

Будем считать промежуток интегрирования Ω конечным (обозначим его $[a, b]$) и предположим, что $f(x)$ имеет на отрезке $[a, b]$ непрерывные производные до порядка $m+s$. Для квадратурной формулы

$S_n(f) = \sum_{i=1}^n c_i f(x_i)$, имеющей алгебраический порядок точности $m-1$, справедливо равенство

$$I(f) = \int_a^b p(x) f(x) dx = S_n(f) + R_m(f).$$

Используя формулу Тейлора для $f(a + (x - a))$ с остаточным членом в интегральной форме $\int_a^x \frac{(x-t)^m}{m!} f^{(m+1)}(t) dt$, можно получить следующее представление погрешности:

$$R_m(f) = \int_a^b f^{(m)}(t)K(t) dt.$$

Здесь ядро $K(t)$ имеет вид

$$K(t) = \int_t^a p(x) \frac{(x-t)^{m-1}}{(m-1)!} dx - \sum_{i=1}^n c_i E(x_i - t) \frac{(x_i - t)^{m-1}}{(m-1)!},$$

где «гасящая» функция $E(x)$ определяется формулой

$$E(x) = \begin{cases} 1 & \text{при } x > 0, \\ \frac{1}{2} & \text{при } x = 0, \\ 0 & \text{при } x < 0. \end{cases}$$

Справедливо представление Эйлера для погрешности:

$$R_m(f) = A_0 \left[f^{(m-1)}(b) - f^{(m-1)}(a) \right] + \dots \\ \dots + A_{s-1} \left[f^{(m+s-2)}(b) - f^{(m+s-2)}(a) \right] + R_{m+s}(f),$$

$$A_j = \frac{1}{b-a} \int_a^b L_j(t) dt, \quad L_{j+1}(t) = \int_a^t [A_j - L_j(x)] dx, \quad L_0(t) = K(t),$$

$$R_{m+s}(f) = \int_a^b f^{(m+s)}(t) L_s(t) dt.$$

Главным членом погрешности обычно называют первое слагаемое в этом представлении.

Правило Рунге. Пусть на отрезке длины h для вычисления интеграла $I(f)$ используется некоторая квадратурная формула $S_h(f)$, имеющая алгебраический порядок точности $m - 1$. Разлагая $f(x)$ в ряд Тейлора в середине отрезка (точке c), получим

$$I(f) - S_h(f) = \alpha f^{(m)}(c) h^{m+1} + O(h^{m+2}).$$

Обозначим через $S_{h/2}(f)$ составную формулу, полученную с помощью формулы $S_h(f)$ для двух половинок исходного отрезка. Тогда при том же α находим

$$I(f) - S_{h/2}(f) = \alpha f^{(m)}(c) \frac{h^{m+1}}{2^m} + O(h^{m+2}).$$

Следовательно, с точностью до членов $O(h^{m+2})$ справедливо следующее *правило Рунге*:

$$I(f) - S_{h/2}(f) \approx \frac{S_{h/2}(f) - S_h(f)}{2^m - 1}.$$

11.1. Пусть интеграл $I(f) = \int_a^b f(x) dx$, где $f(x)$ — гладкая функция, вычисляется по составной формуле трапеций $S_2^N(f)$ с постоянным шагом $h = (b - a)/N$ и $R_2^N(f) = I(f) - S_2^N(f)$ — ее суммарная погрешность.

1) Показать, что справедливо разложение по степеням h

$$R_2^N(f) = a_1 h^2 + a_2 h^4 + a_3 h^6 + \dots$$

2) Показать, что имеет место равенство

$$R_2^N(f) = -\frac{h^2}{12} \int_a^b f''(x) dx + Z(f), \quad Z(f) = o(h^2).$$

3) Пусть $|f^{(3)}(x)| \leq M_3$ на отрезке $[a, b]$. Показать, что $|Z(f)| \leq c_3 M_3 (b - a) h^3$.

4) Пусть $|f^{(4)}(x)| \leq M_4$ на отрезке $[a, b]$. Показать, что $|Z(f)| \leq c_4 M_4 (b - a) h^4$.

11.2. Пусть $I(f) = \int_0^1 f(x) dx$ вычисляется по составной формуле трапеций с переменным шагом интегрирования: $x_k = \varphi(kh)$, $\varphi(t)$ — гладкая функция, $\varphi(0) = 0$, $\varphi(1) = 1$. Доказать, что главный член погрешности есть

$$-\frac{h^2}{12} \int_0^1 f''(\varphi(t)) (\varphi'(t))^3 dt.$$

11.3. Пусть интеграл $I(f) = \int_a^b f(x) dx$, где $f(x)$ — гладкая функция, вычисляется по составной формуле Симпсона $S_3^N(f)$ с посто-

янным шагом $h = (b - a)/N$ и $R_3^N(f) = I(f) - S_3^N(f)$ — ее суммарная погрешность. Показать, что справедливо разложение по степеням h

$$R_3^N(f) = b_1 h^4 + b_2 h^6 + \dots$$

- 11.4. Пусть интеграл $I(f) = \int_0^1 x^\lambda f(x) dx$, где $f(x)$ — гладкая функция и $f(0) \neq 0$, вычисляется по составной формуле трапеций с постоянным шагом $h = 1/N$. Показать, что при $-1 < \lambda < 1$ для суммарной погрешности справедливо разложение по степеням h

$$R_2^N(f) = a_1 h^{1+\lambda} + a_2 h^{2+\lambda} + \dots$$

- 11.5. Используя значения S_h и $S_{h/2}$ квадратуры с главным членом погрешности ch^m , т. е. $I = S_h + ch^m$, построить квадратурную формулу более высокого порядка точности.
- 11.6. Показать, что при применении правила Рунге к формуле трапеций получится формула Симпсона. Насколько при этом увеличится порядок главного члена погрешности?
- 11.7. Показать, что операция построения формулы

$$S_{h,h/2} = S_{h/2} + \frac{S_{h/2} - S_h}{2^m - 1}$$

является экстраполяционной, т. е. при $S_h \neq S_{h/2}$ величина $S_{h,h/2}$ всегда лежит вне отрезка с концами S_h и $S_{h/2}$.

- 11.8. Пусть для вычисления интеграла I от некоторой функции используется квадратурная формула S_h , фактический порядок p главного члена погрешности которой неизвестен для данной функции. Предложить способ численной оценки значения p .

Следующие упражнения иллюстрируют возможные обобщения правила Рунге.

- 11.9. Пусть задана некоторая составная квадратурная формула S_h для приближенного вычисления интеграла I , имеющая представление погрешности вида $I - S_h = ch^m + O(h^{m+1})$. С помощью S_h посчитаны значения при шагах h_1 и $h_2 = h_1/\lambda$.

Доказать, что

$$I - S_{h_2} \approx \frac{S_{h_2} - S_{h_1}}{\lambda^m - 1}.$$

Имеется в виду предельный переход при $h_2 \rightarrow 0$, $\lambda = \text{const} > 1$.

- 11.10.** Пусть задана некоторая составная квадратурная формула S_h для приближенного вычисления интеграла I , имеющая представление погрешности вида $I - S_h = ch^m + O(h^{m+1})$. С помощью S_h посчитаны значения при шагах h_1 и $h_2 = h_1/\lambda$. Доказать, что

$$I - S_{h_2} \approx \frac{S_{h_2} - S_{h_1}}{(h_1/h_2)^m - 1}$$

при следующих условиях: $h_1 \rightarrow 0$, $\lambda \rightarrow \infty$, $(\lambda - 1)/h_1 \rightarrow \infty$.

- 11.11.** Пусть задана некоторая составная квадратурная формула S_h для приближенного вычисления интеграла I , имеющая представление погрешности вида $I - S_h = ch^m + O(h^{m+2})$. С помощью S_h посчитаны значения при шагах h_1 и $h_2 = h_1/\lambda$. Доказать, что

$$I - S_{h_2} \approx \frac{S_{h_2} - S_{h_1}}{(h_1/h_2)^m - 1}$$

при следующих условиях: $h_1 \rightarrow 0$, $h_1 > h_2$.

§ 12. Численное интегрирование функций с особенностями

Быстро осциллирующие функции. Пусть требуется вычислить интеграл $\int_a^b \exp(i\omega x) f(x) dx$, где $\omega(b-a) \gg 1$, $f(x)$ — гладкая функция. Функции $\text{Re}(\exp(i\omega x) f(x))$, $\text{Im}(\exp(i\omega x) f(x))$ имеют на рассматриваемом отрезке примерно $\omega(b-a)/\pi$ нулей. Многочлен степени n имеет не более n нулей на этом отрезке, поэтому такие функции могут быть хорошо приближены многочленами степени n лишь при $n \gg \omega(b-a)/\pi$. Следовательно, для непосредственного вычисления интегралов от таких функций потребуется применение квадратур, точных для многочленов очень высокой степени.

Более выгодным может оказаться использование $\exp(i\omega x)$ в качестве весовой функции. Зададимся узлами интерполирования

$$x_j = \frac{b+a}{2} + \frac{b-a}{2} d_j, \quad j = 1, 2, \dots, n,$$

построим многочлен Лагранжа $L_n(x)$ и рассмотрим квадратурную формулу

$$S_n^\omega(f) = \int_a^b \exp(i\omega x) L_n(x) dx = \frac{b-a}{2} \exp\left(i\omega \frac{b+a}{2}\right) \sum_{j=1}^n D_j(p) f(x_j),$$

где использованы обозначения

$$p = \omega \frac{b-a}{2}, \quad D_j(p) = \int_{-1}^1 \left(\prod_{k \neq j} \frac{\xi - d_k}{d_j - d_k} \right) \exp(ip\xi) d\xi.$$

При этом оценка погрешности

$$R_n = D(d_1, \dots, d_n) \max_{x \in [a, b]} \left| f^{(n)}(x) \right| \left(\frac{b-a}{2} \right)^{n+1}$$

не зависит от ω .

12.1. Для приближенного вычисления интеграла $I(f) =$

$$= \int_0^1 \sin(100\pi x) f(x) dx$$

построить методом неопределенных коэффициентов квадратурную формулу с заданными узлами $S(f) = c_1 f(0) + c_2 f(1)$, точную для многочленов максимальной степени.

12.2. Для приближенного вычисления интеграла $I(f) =$

$$= \int_0^1 \cos(10^4 \pi x) f(x) dx$$

построить методом неопределенных коэффициентов квадратурную формулу с заданными узлами $S(f) = c_1 f(0) + c_2 f(1)$, точную для многочленов максимальной степени.

12.3. Построить квадратурную формулу $S_2^\omega(f)$ с узлами $d_1 = -1$, $d_2 = 1$.

12.4. Показать, что при малых ω формула, полученная в **12.3**, может иметь большую вычислительную погрешность.

12.5. Построить формулу $S_3^\omega(f)$ с узлами $d_1 = -1$, $d_2 = 0$, $d_3 = 1$ (ее часто называют *формулой Филона*).

12.6. Построить формулу $S_5^\omega(f)$ с узлами $d_1 = -1$, $d_2 = -0,5$, $d_3 = 0$, $d_4 = 0,5$, $d_5 = 1$.

Вычисление интегралов от функций с особенностями. Значительная часть реально встречающихся подынтегральных функций — это функции с особенностями, причем особенность может содержаться либо в самой функции, либо в ее производных. Если нерегулярность функции не вызвана колебательным характером ее поведения, то для вычисления больших серий интегралов такого типа используют специальные приемы: выделение особенности в весовую функцию, разбиение интеграла на части, аддитивное представление подынтегральной функции, замену переменных и т. д.

12.7. Пусть требуется вычислить интеграл $I(f) = \int_0^1 f(x)x^\alpha dx$, где $\alpha \in (0, 1)$, $f(x)$ — гладкая функция, $f(0) \neq 0$. Построить составную квадратурную формулу $S(f) = \sum_{k=0}^N D_k f(k/N)$ с оценкой погрешности вида $\text{const} \cdot \max_{x \in [0,1]} |f''(x)| \cdot N^{-2}$.

12.8. Пусть требуется вычислить интеграл $\int_0^1 \frac{f(x)\lambda}{\lambda^2 + x^2} dx$, $f \in C^{(2)}[0, 1]$, $|\lambda| \ll 1$. Показать, что при использовании составной формулы трапеций с постоянным шагом $h = 1/N$ суммарная погрешность оценивается величиной $\text{const} \cdot \min\left(\frac{h}{\lambda}, \frac{h^2}{\lambda^2}\right)$.

12.9. Для вычисления интеграла $\int_0^1 \frac{f(x)\lambda}{\lambda^2 + x^2} dx$, $f \in C^{(1)}[0, 1]$, $|\lambda| \ll 1$, используется следующая составная квадратурная формула с постоянным шагом $h = 1/N$:

$$S(f) = \sum_{k=1}^N f(\xi_k) \left[\text{arctg} \left(\frac{kh}{\lambda} \right) - \text{arctg} \left(\frac{(k-1)h}{\lambda} \right) \right],$$

где $(k-1)h \leq \xi_k \leq kh$. Получить оценку суммарной погрешности вида $|R^N| \leq \text{const} \cdot \max_{x \in [0,1]} |f'(x)| \cdot N^{-1}$.

- 12.10.** Предложить способ вычисления интеграла $\int_0^1 \frac{\ln x}{1+x^2} dx$ по составной квадратурной формуле с постоянным шагом h , чтобы суммарная погрешность имела порядок $O(h^2)$.
- 12.11.** Предложить способ вычисления интеграла $\int_0^1 \frac{\ln x}{1+x^2} dx$ по составной квадратурной формуле с постоянным шагом h , чтобы суммарная погрешность имела порядок $O(h^4)$.
- 12.12.** Пусть $f(x)$ — достаточно гладкая функция и $f(0) \neq 0$. Предложить квадратурную формулу для вычисления интеграла $\int_0^1 f(x)x^{-\alpha} \sin(\omega x) dx$ при $\alpha > 1$, $\omega \gg 1$.
- 12.13.** Построить квадратурную формулу для вычисления с суммарной погрешностью не более 10^{-4} интеграла $\int_1^\infty \frac{f(x)}{1+x^2} dx$, если для некоторого фиксированного $l \geq 1$ справедливо $|f^{(l)}(x)| \leq A_l$.
- 12.14.** Построить квадратурную формулу для вычисления с суммарной погрешностью не более 10^{-4} интеграла $\int_0^\infty f(x)e^{-x} dx$, если для некоторого фиксированного $l \geq 1$ справедливо $|f^{(l)}(x)| \leq A_l$.
- 12.15.** Построить квадратурную формулу (не проводя замену переменных) для вычисления с суммарной погрешностью не более 10^{-3} интеграла $\int_0^1 \frac{f(x)}{\sqrt{x}} dx$, если для некоторого фиксированного $l \geq 1$ справедливо $|f^{(l)}(x)| \leq A_l$.
- 12.16.** Построить квадратурную формулу для вычисления с суммарной погрешностью не более 10^{-4} интеграла $\int_0^1 f(x) \frac{\sqrt{\sin x}}{x} dx$, если для некоторого фиксированного $l \geq 1$ справедливо $|f^{(l)}(x)| \leq A_l$.

Глава IV

Матричные вычисления



§ 13. Векторные и матричные нормы

Нормой вектора $\mathbf{x} = (x_1, \dots, x_n)^T$ называется функционал, обозначаемый $\|\mathbf{x}\|$ и удовлетворяющий условиям:

$$\begin{aligned}\|\mathbf{x}\| &> 0, \quad \mathbf{x} \neq 0, \quad \|\mathbf{0}\| = 0, \\ \|\alpha\mathbf{x}\| &= |\alpha| \|\mathbf{x}\|, \\ \|\mathbf{x} + \mathbf{y}\| &\leq \|\mathbf{x}\| + \|\mathbf{y}\|.\end{aligned}$$

Наиболее употребительны следующие нормы:

$$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|, \quad \|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|, \quad \|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{(\mathbf{x}, \mathbf{x})}.$$

Нормы $\|\cdot\|_I$ и $\|\cdot\|_{II}$ называются *эквивалентными*, если для всех $\mathbf{x} \in \mathbb{R}^n$ справедливы неравенства

$$c_1 \|\mathbf{x}\|_{II} \leq \|\mathbf{x}\|_I \leq c_2 \|\mathbf{x}\|_{II}$$

с одними и теми же положительными постоянными c_1 и c_2 .

Нормой матрицы A называется функционал, обозначаемый $\|A\|$ и удовлетворяющий условиям:

$$\begin{aligned}\|A\| &> 0, \quad A \neq 0, \quad \|0\| = 0, \\ \|\alpha A\| &= |\alpha| \|A\|, \\ \|A + B\| &\leq \|A\| + \|B\|, \\ \|AC\| &\leq \|A\| \|C\|.\end{aligned}$$

Пусть задана некоторая векторная норма $\|\cdot\|_v$. Тогда матричную норму можно определить как операторную

$$\|A\|_v = \sup_{\|\mathbf{x}\|_v \neq 0} \frac{\|A\mathbf{x}\|_v}{\|\mathbf{x}\|_v} = \sup_{\|\mathbf{x}\|_v = 1} \|A\mathbf{x}\|_v.$$

В этом случае матричная норма называется *подчиненной* соответствующей векторной норме $\|\cdot\|_v$.

- 13.1.** Является ли функционал $\min(|x_1| + 2|x_2|, 2|x_1| + |x_2|)$ нормой в \mathbb{R}^2 ?
- 13.2.** Является ли функционал $\max_{t \in [0,1]} \left| \sum_{k=1}^n x_k t^{k-1} \right|$ нормой в \mathbb{R}^n ?
- 13.3.** Найти константы эквивалентности норм $\|\mathbf{x}\|_\infty$, $\|\mathbf{x}\|_1$, $\|\mathbf{x}\|_2$, а также векторы, на которых они достигаются.
- 13.4.** Пусть A — симметричная и положительно определенная матрица: $A = A^T > 0$. 1) Доказать, что $\sqrt{(A\mathbf{x}, \mathbf{x})}$ является нормой вектора \mathbf{x} и 2) найти константы ее эквивалентности норме $\|\mathbf{x}\|_2$.
- 13.5.** Найти матричные нормы, подчиненные векторным нормам $\|\cdot\|_1$, $\|\cdot\|_2$ и $\|\cdot\|_\infty$.
- 13.6.** Доказать, что модуль любого собственного значения матрицы не больше любой ее нормы.
- 13.7.** Пусть A — вещественная матрица размерности $n \times m$. Доказать следующие свойства спектральной нормы $\|A\|_2$:
 1) $\|A\|_2 = \sup_{\substack{\|\mathbf{x}\|_2=1 \\ \|\mathbf{y}\|_2=1}} |\mathbf{y}^T A \mathbf{x}|$; 2) $\|A^T\|_2 = \|A\|_2$;
 3) $\|A^T A\|_2 = \|A A^T\|_2 = \|A\|_2^2$.
- 13.8.** Пусть A — вещественная прямоугольная матрица. Показать, что умножение ее справа или слева на ортогональную матрицу Q соответствующих размеров не меняет ее спектральную норму.
- 13.9.** Используя выражения для матричных норм из **13.5**, показать справедливость неравенства $\|A\|_2^2 \leq \|A\|_1 \|A\|_\infty$.
- 13.10.** Показать, что функция от элементов матрицы $\eta(A) = \max_{i,j} |a_{i,j}|$ не является нормой в пространстве матриц (хотя и является нормой вектора с компонентами a_{ij} в \mathbb{R}^{n^2}).
- 13.11.** Доказать, что выражение $M(A) = n \eta(A)$ (см. **13.10**) является матричной нормой.
- 13.12.** Доказать, что для вектора $\mathbf{x} = (x_1, x_2)^T$ и числа $h > 0$ выражение $\|\mathbf{x}\|_h = \max\left(|x_1|, \frac{|x_2 - x_1|}{h}\right)$ является нормой. Найти матричную норму, подчиненную этой векторной норме.

- 13.13.** Доказать, что выражение $N(A) = \left(\sum_{i,j=1}^n a_{ij}^2 \right)^{1/2}$ является матричной нормой. Найти наилучшие константы эквивалентности, связывающие $N(A)$ и нормы $\|\cdot\|_1$, $\|\cdot\|_2$ и $\|\cdot\|_\infty$.
- 13.14.** Доказать, что выражение $\max_k (d_k |x_k|)$, $d_k > 0$, — норма вектора \mathbf{x} , и найти подчиненную ей норму матрицы.
- 13.15.** Доказать, что выражение $\sum_{k=1}^n d_k |x_k|$, $d_k > 0$, — норма вектора \mathbf{x} , и найти подчиненную ей норму матрицы.
- 13.16.** Доказать, что выражение $\sqrt{\sum_{k=1}^n d_k x_k^2}$, $d_k > 0$, — норма вектора \mathbf{x} , и найти подчиненную ей норму матрицы.
- 13.17.** Доказать, что выражение $\max_{1 \leq i \leq n} \left(\sum_{k=1}^i x_k \right)$ — норма вектора \mathbf{x} , и найти подчиненную ей норму матрицы.
- 13.18.** Пусть $M(A) = n \max_{1 \leq i, j \leq n} |a_{ij}|$. Найти наилучшие константы c_1, c_2 в неравенстве $c_1 M(A) \leq \|A\|_2 \leq c_2 M(A)$.
- 13.19.** Пусть $M(A) = n \max_{1 \leq i, j \leq n} |a_{ij}|$. Найти наилучшие константы c_1, c_2 в неравенстве $c_1 M(A) \leq \|A\|_1 \leq c_2 M(A)$.
- 13.20.** Пусть $\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$, $p \geq 1$. Доказать неравенство Йенсена $\|\mathbf{x}\|_p \leq \|\mathbf{x}\|_q$, $1 \leq q \leq p \leq \infty$.
- 13.21.** Доказать равенство $\lim_{p \rightarrow \infty} \|\mathbf{x}\|_p = \|\mathbf{x}\|_\infty \quad \forall \mathbf{x} \in \mathbb{R}^n$.
- 13.22.** Доказать, что сходимость в любой норме в пространстве \mathbb{R}^n эквивалентна покоординатной сходимости.
- 13.23.** Пусть $\|\cdot\|$ — векторная норма в \mathbb{R}^m и $A \in \mathbb{R}^{m \times n}$ — прямоугольная матрица размерности $m \times n$. Показать, что если ранг матрицы $\text{rang}(A) = n$, то $\|A\mathbf{x}\|$ — векторная норма в \mathbb{R}^n .
- 13.24.** Проверить, что $\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$, $p \geq 1$, является нормой в пространстве \mathbb{C}^n векторов с комплексными координатами. Показать, что при $\mathbf{x} \in \mathbb{C}^n$ справедливо неравенство

$\|\mathbf{x}\|_p \leq \|\operatorname{Re}(\mathbf{x})\|_p + \|\operatorname{Im}(\mathbf{x})\|_p$. Найти такую максимальную постоянную $c_0 > 0$, что $c_0(\|\operatorname{Re}(\mathbf{x})\|_2 + \|\operatorname{Im}(\mathbf{x})\|_2) \leq \|\mathbf{x}\|_2$ для всех $\mathbf{x} \in \mathbb{C}^n$.

13.25. Пусть $\|\cdot\|$ — некоторая норма в \mathbb{R}^n . Доказать, что функционал

$$\|\mathbf{x}\|_* = \sup_{\mathbf{y} \neq 0} \frac{(\mathbf{x}, \mathbf{y})}{\|\mathbf{y}\|}$$

также задает норму в \mathbb{R}^n , называемую двойственной к $\|\cdot\|$. Найти норму, двойственную к $\|\cdot\|_\infty$.

13.26. Пусть $1 \leq p \leq \infty$ и B — любая подматрица квадратной матрицы A . Доказать, что $\|B\|_p \leq \|A\|_p$.

13.27. Доказать, что если $D = \operatorname{diag}(d_1, d_2, \dots, d_k) \in \mathbb{R}^{m \times n}$, где $k = \min\{m, n\}$, то $\|D\|_p = \max_i |d_i|$.

13.28. Пусть B — невырожденная матрица, $\|\cdot\|$ — векторная норма. Доказать, что $\|\mathbf{x}\|_* = \|B\mathbf{x}\|$ также является векторной нормой. Какая норма в пространстве матриц порождается нормой $\|\mathbf{x}\|_*$?

13.29. Показать, что если A — невырожденная матрица, то для нормы матрицы, подчиненной векторной норме, справедливо равенство

$$\|A^{-1}\|^{-1} = \inf_{\mathbf{x} \neq 0} \|A\mathbf{x}\| / \|\mathbf{x}\|.$$

13.30. Доказать неравенство $\|A\|_2 \leq \|A\|^{1/2} \|A^T\|^{1/2}$ для любой нормы A , подчиненной какой-либо векторной норме.

13.31. Пусть $A = A^T > 0$. Показать, что существует единственная матрица $X = X^T > 0$ такая, что $A = X^2$.

13.32. Пусть $A = A^T > 0$ и $\|\mathbf{x}\|_A = (\mathbf{A}\mathbf{x}, \mathbf{x})^{1/2}$. Доказать, что для произвольного многочлена $p_m(t)$ степени $m \geq 0$ верно равенство

$$\|p_m(A)\|_A = \|p_m(A)\|_2.$$

13.33. Доказать, что если матрица A — вещественная и $(\mathbf{A}\mathbf{x}, \mathbf{x}) > 0$ для всех вещественных $\mathbf{x} \neq 0$, то существует такая постоянная $\delta > 0$, не зависящая от \mathbf{x} , что $(\mathbf{A}\mathbf{x}, \mathbf{x}) \geq \delta (\mathbf{x}, \mathbf{x})$.

- 13.34.** Привести пример положительно определенной вещественной матрицы, спектр которой не является вещественным.
- 13.35.** Доказать, что матричные нормы $M(A) = n \max_{1 \leq i, j \leq n} |a_{ij}|$ и $N(A) = \left(\sum_{i, j=1}^n a_{ij}^2 \right)^{1/2}$ не подчинены никаким векторным нормам.
- 13.36.** Показать, что для любого собственного значения $\lambda(A)$ невырожденной матрицы A справедлива оценка $1/\|A^{-1}\| \leq |\lambda(A)|$.
- 13.37.** Доказать, что для любого собственного значения $\lambda(A)$ матрицы A справедливо неравенство $|\lambda(A)| \leq \inf_k \|A^k\|^{1/k}$, k — натуральное число.
- 13.38.** Доказать, что если A — нормальная матрица ($AA^T = A^T A$), то $\|A\|_2 = \rho(A)$, где $\rho(A) = \max_i |\lambda_i(A)|$ — спектральный радиус матрицы A .
- 13.39.** Убедиться, что матрица A размерности $n \times n$ при $n \geq 2$ не определяется однозначно значениями квадратичной формы (Ax, x) на произвольном векторе x , т. е. найдутся две различные матрицы A и B , для которых $(Ax, x) \equiv (Bx, x)$ для любых вещественных x .
- 13.40.** Пусть $A = A^T > 0$ и $F(x) = (Ax, x) - 2(b, x)$ — квадратичная функция. Доказать, что:
- 1) $F(x) = \|x^* - x\|_A^2 - \|x^*\|_A^2$, где x^* — решение системы $Ax = b$;
 - 2) равенство $F(x^*) = \min_x F(x)$ выполняется тогда и только тогда, когда x^* — решение системы $Ax = b$;
 - 3) для градиента функции $F(x)$ справедлива формула

$$\text{grad } F(x) = 2(Ax - b).$$
- 13.41.** Доказать, что всякая норма $\|\cdot\|_m$ матрицы согласована с какой-либо векторной нормой $\|\cdot\|_v$, т. е. верна оценка $\|Ax\|_v \leq \|A\|_m \|x\|_v$.
- 13.42.** Пусть A — квадратная матрица, $\rho(A)$ — ее спектральный радиус. Доказать, что для любого $\varepsilon > 0$ существует по крайней мере одна матричная норма такая, что $\rho(A) \leq \|A\| \leq \rho(A) + \varepsilon$.

§ 14. Элементы теории возмущений

Рассмотрим систему линейных алгебраических уравнений

$$A\mathbf{x} = \mathbf{b}$$

с квадратной невырожденной матрицей A . При ее решении в результате вычислений с конечной разрядностью вместо \mathbf{x} получается *приближенное* решение $\tilde{\mathbf{x}}$, которое можно рассматривать как *точное* решение *возмущенной* системы $(A + \delta A)\tilde{\mathbf{x}} = \mathbf{b}$, где матрица возмущений δA мала в каком-либо смысле. Другой источник ошибок в $\tilde{\mathbf{x}}$ определяется возмущениями δA и $\delta \mathbf{b}$ в элементах матрицы A и в компонентах вектора правой части \mathbf{b} (например, вследствие ошибок округлений, возникающих в процессе ввода вещественных чисел в память компьютера).

Чтобы оценить, насколько приближенное решение $\tilde{\mathbf{x}}$ отличается от точного решения \mathbf{x} , используют нормы векторов и подчиненные нормы матриц. Пусть в системе $A\mathbf{x} = \mathbf{b}$ возмущается только вектор \mathbf{b} , т. е. вместо исходной системы решается возмущенная система $A\tilde{\mathbf{x}} = \tilde{\mathbf{b}} \equiv \mathbf{b} + \delta \mathbf{b}$, и пусть $\tilde{\mathbf{x}}$ — точное решение возмущенной системы. Тогда для относительной ошибки в $\tilde{\mathbf{x}}$ верна оценка

$$\frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} \leq \|A\| \|A^{-1}\| \frac{\|\mathbf{b} - \tilde{\mathbf{b}}\|}{\|\mathbf{b}\|}.$$

Величину $\|A\| \|A^{-1}\|$ называют *числом обусловленности* матрицы A и часто обозначают $\text{cond}(A)$. Для вырожденных матриц $\text{cond}(A) = \infty$. Конкретное значение $\text{cond}(A)$ зависит от выбора матричной нормы, однако, в силу их эквивалентности, при практических оценках этим различием можно пренебречь.

Из приведенного выше неравенства следует, что даже если *вектор невязки* $\mathbf{r} = \mathbf{b} - A\tilde{\mathbf{x}}$ мал, относительные возмущения в решении могут быть большими, если $\text{cond}(A)$ велико (такие матрицы называют *плохо обусловленными*).

14.1. Доказать неравенство

$$\frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} \leq \text{cond}(A) \frac{\|\mathbf{b} - \tilde{\mathbf{b}}\|}{\|\mathbf{b}\|} = \text{cond}(A) \frac{\|\mathbf{b} - A\tilde{\mathbf{x}}\|}{\|\mathbf{b}\|}.$$

14.2. Найти решения двух систем с близкими коэффициентами:

$$\begin{cases} x + 3y = 4, \\ x + 3,00001y = 4,00001, \end{cases} \quad \begin{cases} x + 3y = 4, \\ x + 2,99999y = 4,00001 \end{cases}$$

и объяснить полученный результат.

14.3. Показать, что $\text{cond}(A) \geq 1$ для любой матрицы A и $\text{cond}_2(Q) = \|Q\|_2 \|Q^{-1}\|_2 = 1$ для ортогональной матрицы Q .

14.4. Можно ли утверждать, что если определитель матрицы мал, то матрица плохо обусловлена?

14.5. Пусть A — двухдиагональная матрица размерности $n \times n$ с элементами $a_{ij} = \{1 \text{ для } i = j; a \text{ для } i = j - 1; 0 \text{ для остальных индексов}\}$, $|a| \neq 1$. Вычислить $\text{cond}_\infty(A)$ и оценить возмущение в компоненте x_1 решения системы $A\mathbf{x} = \mathbf{b}$, если возмущена на ε только последняя компонента b_n вектора \mathbf{b} .

14.6. Пусть A — двухдиагональная матрица размерности $n \times n$ с элементами $a_{ij} = \{p \text{ для } i = j; q \text{ для } i = j - 1; 0 \text{ для остальных индексов}\}$. Вычислить матрицу A^{-1} и показать, что при $|q| < |p|$ матрица A хорошо обусловлена, а при $|q| > |p|$ и больших значениях n — плохо обусловлена.

14.7. Пусть матрица A определена как в 14.6. Выразить явно решение системы $A\mathbf{x} = \mathbf{b}$ через правую часть.

14.8. Решается система $A\mathbf{x} = \mathbf{b}$ с матрицей

$$A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & \varepsilon & \varepsilon \\ 1 & \varepsilon & \varepsilon \end{pmatrix}, \quad 0 < |\varepsilon| \ll 1.$$

В результате замены $x'_1 = x_1$, $x'_2 = \varepsilon x_2$, $x'_3 = \varepsilon x_3$ для нахождения новых неизвестных \mathbf{x}' имеем систему $A'\mathbf{x}' = \mathbf{b}'$ с матрицей

$$A' = \begin{pmatrix} \varepsilon & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

В каком случае число обусловленности меньше?

- 14.9.** Пусть $A = A^T > 0$, $\lambda(A) \in [m, M]$ и $A \neq \beta I$, где I — единичная матрица. Доказать, что $\text{cond}_2(A + \alpha I)$ монотонно убывает по α при $\alpha > 0$.
- 14.10.** Существуют ли несимметричные матрицы, для которых справедливо: $\text{cond}^2(A) = \text{cond}(A^2) > 1$?
- 14.11.** Доказать неравенство $n^{-1} \leq \text{cond}_1(A)/\text{cond}_2(A) \leq n$.
- 14.12.** Оценить снизу и сверху $\text{cond}_\infty(A)$ невырожденной матрицы A размерности $n \times n$, используя границы собственных значений матрицы $A^T A$, т. е. $\lambda(A^T A) \in [\alpha, \beta]$.
- 14.13.** Получить неравенство $\text{cond}(A) \geq |\lambda_{\max}(A)/\lambda_{\min}(A)|$ для произвольной невырожденной матрицы A и любой матричной нормы, используемой при определении числа обусловленности.
- 14.14.** Доказать, что $\text{cond}(AB) \leq \text{cond}(A)\text{cond}(B)$ для любой нормы в определении числа обусловленности и любых матриц A и B .
- 14.15.** Оценить $\text{cond}_2(A)$ трехдиагональной $n \times n$ матрицы A с элементами $a_{ij} = \{2 \text{ для } i = j; -1 \text{ для } |i - j| = 1; 0 \text{ для остальных индексов}\}$.
- 14.16.** Оценить $\text{cond}_2(A)$ трехдиагональной $n \times n$ матрицы A с элементами $a_{ij} = \{2/3 \text{ для } i = j; 1/6 \text{ для } |i - j| = 1; 0 \text{ для остальных индексов}\}$.
- 14.17.** Пусть I — единичная матрица и $\|\delta I\| < 1$. Показать, что матрица $I - \delta I$ невырожденная и выполнена оценка

$$\|(I - \delta I)^{-1}\| \leq \frac{1}{1 - \|\delta I\|}.$$

- 14.18.** Пусть I — единичная матрица и $\|\delta I\| < 1$. Получить оценку отклонения матрицы I от матрицы $(I - \delta I)^{-1}$.
- 14.19.** Пусть A — невырожденная матрица и $\|A^{-1}\delta A\| < 1$. Показать, что матрица $A + \delta A$ невырожденная и выполнена оценка

$$\|(A + \delta A)^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\delta A\|}.$$

- 14.20.** Пусть A — невырожденная матрица и $\|A^{-1}\delta A\| < 1$. Получить оценку отклонения матрицы $(A + \delta A)^{-1}$ от A^{-1} .

14.21. Оценить снизу число обусловленности $\text{cond}_2(A)$ матрицы:

$$1) A = \begin{pmatrix} 10 & 10 & 30 \\ 0,1 & 0,5 & 0,1 \\ 0,03 & 0,01 & 0,01 \end{pmatrix}; \quad 2) A = \begin{pmatrix} 1 & 20 & -400 \\ 0,2 & -2 & -20 \\ -0,04 & -0,2 & 1 \end{pmatrix}.$$

14.22. Система $Ax = b$, где

$$A = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 10^{-10} & 10^{-10} \\ 1 & 10^{-10} & 10^{-10} \end{pmatrix}, \quad b = \begin{pmatrix} 2(1 + 10^{-10}) \\ -10^{-10} \\ 10^{-10} \end{pmatrix}$$

имеет решение $x = (10^{-10}, -1, 1)^T$. Доказать, что если $(A + \delta A)y = b$, $\|\delta A\| \leq 10^{-8}\|A\|$, то $\|x - y\| \leq 10^{-7}$. Таким образом, в данном случае относительно малые изменения в элементах матрицы A не приводят к большим изменениям в решении, хотя $\text{cond}_\infty(A)$ имеет порядок 10^{10} .

14.23. Доказать, что число обусловленности $\text{cond}_2(A)$ матрицы

$$A = \begin{pmatrix} 100 & 99 \\ 99 & 98 \end{pmatrix}$$

максимально среди всех невырожденных матриц второго порядка, элементы которых — положительные целые числа, не превосходящие 100.

14.24. Пусть матрица A имеет строгое диагональное преобладание: при некотором $0 < q < 1$ для элементов каждой строки i выполнено неравенство $\sum_{j \neq i} |a_{ij}| \leq q|a_{ii}|$. Оценить снизу и сверху

$\text{cond}_\infty(A)$, используя только диагональные элементы матрицы и параметр q .

14.25. Пусть R — верхняя треугольная матрица размерности $n \times n$, у которой $|r_{ij}| \leq 1$ для всех i, j и $r_{ii} = 1$ для всех i . Найти максимально возможное значение числа обусловленности $\text{cond}_\infty(R)$.

14.26. Показать, что определитель D_n матрицы Коши K_n с элементами $k_{ij} = 1/(a_i + b_j)$, $1 \leq i, j \leq n$ равен

$$D_n = \prod_{1 \leq i < j \leq n} (a_i - a_j)(b_i - b_j) \left(\prod_{1 \leq i, j \leq n} (a_i + b_j) \right)^{-1}.$$

14.27. Пусть задана матрица Гильберта H_n с элементами $h_{ij} = 1/(i+j-1)$, $1 \leq i, j \leq n$. Показать, что элементами матрицы H_n^{-1} являются целые числа, которые можно вычислить по формуле

$$\tilde{h}_{ij} = (-1)^{i+j} \frac{(i+n-1)!(j+n-1)!}{[(i-1)!]^2 [(j-1)!]^2 (n-i)!(n-j)!(i+j-1)}.$$

14.28. Оценить рост числа обусловленности $\text{cond}_\infty(H_n)$ матрицы Гильберта относительно параметра размерности n .

14.29. Пусть система $A\mathbf{x} = \mathbf{b}$ с матрицей

$$A = \begin{pmatrix} \varepsilon & 1 \\ 1 & 1 \end{pmatrix}$$

решается методом LR -разложения: $A = LR$, $L\mathbf{y} = \mathbf{b}$, $R\mathbf{x} = \mathbf{y}$. Вычислить $\text{cond}_\infty(L)$ и $\text{cond}_\infty(R)$, если LR -разложение строится методом Гаусса: 1) без выбора ведущего элемента; 2) с выбором ведущего элемента.

§ 15. Метод простой итерации

Преобразуем систему линейных алгебраических уравнений

$$A\mathbf{x} = \mathbf{b} \quad (1)$$

с невырожденной матрицей A к виду

$$\mathbf{x} = B\mathbf{x} + \mathbf{c}. \quad (2)$$

Если решение системы (2) находят как предел последовательности

$$\mathbf{x}^{k+1} = B\mathbf{x}^k + \mathbf{c}, \quad (3)$$

то такой процесс называют *методом простой итерации*, а матрицу B — *матрицей или оператором перехода*. Справедливы следующие теоремы о сходимости метода.

Теорема 1. Если $\|B\| < 1$, то система уравнений (2) имеет единственное решение и итерационный процесс (3) сходится к решению со скоростью геометрической прогрессии.

Теорема 2. Пусть система (2) имеет единственное решение. Итерационный процесс (3) сходится к решению системы (2) при

любом начальном приближении тогда и только тогда, когда все собственные значения матрицы B по модулю меньше 1.

Асимптотической скоростью сходимости $R_\infty(B)$ итерационного метода называют величину $R_\infty(B) = -\ln \rho(B)$, где $\rho(B)$ — спектральный радиус (максимальное по модулю собственное значение) оператора перехода B .

Рассмотрим общий способ перехода от системы (1) к системе (2). Всякая система

$$\mathbf{x} = \mathbf{x} - D(A\mathbf{x} - \mathbf{b}) \quad (4)$$

имеет вид (2) и при $\det(D) \neq 0$ равносильна системе (1). В то же время, всякая система (2), равносильная (1), записывается в виде (4) с матрицей $D = (I - B)A^{-1}$.

Оптимальный линейный одношаговый метод. Для систем со знакоопределенными матрицами метод (3) обычно строят в виде

$$\frac{\mathbf{x}^{k+1} - \mathbf{x}^k}{\tau} + A\mathbf{x}^k = \mathbf{b}, \quad \text{т. е. } B = I - \tau A, \quad \mathbf{c} = \tau \mathbf{b}, \quad (5)$$

где τ — итерационный параметр. Так как точное решение \mathbf{x} удовлетворяет уравнению (5), то для вектора ошибки $\mathbf{z}^k = \mathbf{x} - \mathbf{x}^k$ справедливо

$$\mathbf{z}^{k+1} = (I - \tau A)\mathbf{z}^k, \quad \|\mathbf{z}^{k+1}\| \leq \|I - \tau A\| \|\mathbf{z}^k\|, \quad k = 0, 1, \dots$$

Итерационный параметр τ ищется из условия минимума нормы оператора перехода. Если $A = A^T > 0$ и выбрана евклидова векторная норма, то минимизационная задача $\min_{\tau} \|I - \tau A\|_2 = \min_{\tau} \left(\max_{\lambda(A)} |1 - \tau \lambda(A)| \right) = q$ решается явно. Пусть известны точные границы спектра матрицы A , т. е. $\lambda(A) \in [m, M]$, тогда оптимальные значения таковы:

$$\tau = \frac{2}{m + M}, \quad q = \frac{M - m}{M + m} < 1,$$

и справедлива оценка $\|\mathbf{x} - \mathbf{x}^k\|_2 \leq q^k \|\mathbf{x} - \mathbf{x}^0\|_2$.

Оптимальный линейный N -шаговый метод. Будем считать, что в итерационном алгоритме

$$\frac{\mathbf{x}^{k+1} - \mathbf{x}^k}{\tau_{k+1}} + A\mathbf{x}^k = \mathbf{b} \quad (6)$$

допускается циклическое изменение (с периодом N) параметра τ в зависимости от номера итерации, т. е. $\tau_1, \tau_2, \dots, \tau_N, \tau_1, \tau_2, \dots$. В этом случае после N итераций для вектора ошибки имеем:

$$\mathbf{z}^{k+N} = \prod_{j=1}^N (I - \tau_j A) \mathbf{z}^k, \quad \|\mathbf{z}^{k+N}\|_2 \leq \left\| \prod_{j=1}^N (I - \tau_j A) \right\|_2 \|\mathbf{z}^k\|_2, \quad k=0, 1, 2, \dots$$

Будем искать набор $\tau_j, j = 1, \dots, N$, из условия минимума нормы оператора перехода. Если $A = A^T > 0$, то

$$\min_{\tau_j} \left\| \prod_{j=1}^N (I - \tau_j A) \right\|_2 = \min_{\tau_j} \left(\max_{\lambda(A)} \left| \prod_{j=1}^N (1 - \tau_j \lambda(A)) \right| \right).$$

Пусть известны точные границы спектра матрицы A , т. е. $\lambda(A) \in [m, M]$, тогда оптимальные значения параметров равны обратным величинам корней многочлена Чебышёва степени N на отрезке $[m, M]$:

$$\tau_j^{-1} = \frac{M+m}{2} + \frac{M-m}{2} \cos \frac{\pi(2j-1)}{2N}, \quad 1 \leq j \leq N,$$

и справедлива оценка погрешности после N итераций:

$$\|\mathbf{x} - \mathbf{x}^N\|_2 \leq \frac{2q_1^N}{1+q_1^{2N}} \|\mathbf{x} - \mathbf{x}^0\|_2 \leq 2q_1^N \|\mathbf{x} - \mathbf{x}^0\|_2, \quad q_1 = \frac{\sqrt{M} - \sqrt{m}}{\sqrt{M} + \sqrt{m}}.$$

Для устойчивой численной реализации N -шагового метода требуется специальным образом упорядочить итерационные параметры τ_j .

Недостатком рассмотренных методов является требование информации о границах спектра матрицы A .

15.1. Пусть элементы матрицы B имеют вид $b_{ij} = \frac{1}{2} \cdot 3^{-|i-j|}$.

Доказать, что система $\mathbf{x} = B\mathbf{x} + \mathbf{c}$ имеет единственное решение и метод простой итерации сходится при любом начальном приближении.

15.2. Пусть матрица B имеет вид

$$B = \begin{pmatrix} \alpha & \beta & 0 \\ \beta & \alpha & \beta \\ 0 & \beta & \alpha \end{pmatrix}.$$

Найти все α, β , при которых метод простой итерации $\mathbf{x}^{k+1} = B\mathbf{x}^k + \mathbf{c}$ сходится с произвольного начального приближения.

- 15.3.** Привести пример задачи $\mathbf{x} = B\mathbf{x} + \mathbf{c}$ такой, что у матрицы B есть собственное значение λ вне единичного круга, но метод $\mathbf{x}^{k+1} = B\mathbf{x}^k + \mathbf{c}$ сходится при некотором начальном приближении.
- 15.4.** Пусть матрица B в методе $\mathbf{x}^{k+1} = B\mathbf{x}^k + \mathbf{c}$ имеет вид

$$B = \begin{pmatrix} \alpha & 4 \\ 0 & \beta \end{pmatrix}, \quad 0 < \alpha, \beta < 1.$$

Показать, что величина ошибки $\mathbf{z}^k = \mathbf{x} - \mathbf{x}^k$ в норме $\|\cdot\|_\infty$ начинает монотонно убывать лишь с некоторого номера итерации N . Оценить N при $\alpha = \beta \approx 1$.

- 15.5.** Пусть все собственные значения матрицы A вещественные и положительные. Доказать сходимость метода $\mathbf{x}^{k+1} = (I - \tau A)\mathbf{x}^k + \tau \mathbf{b}$ при $\tau = \|A\|^{-1}$ с любой матричной нормой.
- 15.6.** Доказать, что все собственные значения матрицы A размерности $n \times n$ принадлежат области комплексной плоскости $G(A)$, представляющей собой объединение кругов

$$G_i(A) = \{z : |z - a_{ii}| \leq R_i(A) = \sum_{j \neq i} |a_{ij}|\}, \quad i = 1, \dots, n.$$

- 15.7.** Доказать, что все собственные значения матрицы A размерности $n \times n$ принадлежат области $G(A) \cap G(A^T)$.
- 15.8.** Доказать, что у матрицы

$$\begin{pmatrix} 2 & 0,4 & 0,4 \\ 0,3 & 4 & 0,4 \\ 0,1 & 0,1 & 5 \end{pmatrix}$$

все собственные значения вещественные и найти интервалы, которым они принадлежат.

- 15.9.** Привести пример, демонстрирующий ложность утверждения: все собственные значения матрицы A размерности $n \times n$ принадлежат объединению кругов

$$|z - a_{ii}| \leq \min\{R_i(A), R_i(A^T)\}, \quad i = 1, \dots, n.$$

- 15.10.** Доказать, что все собственные значения матрицы A размерности $n \times n$ принадлежат объединению кругов

$$|z - a_{ii}| \leq R_i^\alpha(A) R_i^{1-\alpha}(A^T), \quad i = 1, \dots, n,$$

где α — произвольное число из отрезка $[0, 1]$ (*теорема Островского*).

- 15.11.** Доказать, что все собственные значения матрицы A размерности $n \times n$ принадлежат объединению $n(n-1)/2$ овалов Кассини:

$$|z - a_{ii}| |z - a_{jj}| \leq R_i(A) R_j(A), \quad i, j = 1, \dots, n, \quad i \neq j.$$

- 15.12.** Доказать, что если для некоторого i и при всех j выполняются неравенства $|a_{ii} - a_{jj}| > R_i(A) + R_j(A)$, то в круге $|z - a_{ii}| \leq R_i(A)$ лежит точно одно собственное значение матрицы A .

- 15.13.** Пусть A — матрица *простой структуры*, т. е. подобна диагональной ($A = QDQ^{-1}$, где столбцы \mathbf{q}_i матрицы Q — собственные векторы матрицы A , а элементы диагональной матрицы D — соответствующие собственные значения, т. е. $d_{ii} = \lambda_i$), и все $\lambda(A) \in [m, M]$, $m > 0$. Доказать, что метод $\mathbf{x}^{k+1} = (I - \tau A)\mathbf{x}^k + \tau \mathbf{b}$ сходится с произвольного начального приближения при $0 < \tau < 2/M$.

- 15.14.** Пусть матрица системы $A\mathbf{x} = \mathbf{b}$ имеет вид

$$A = \begin{pmatrix} 2 & 0,3 & 0,5 \\ 0,1 & 3 & 0,4 \\ 0,1 & 0,1 & 4,8 \end{pmatrix}.$$

Доказать, что метод простой итерации $\mathbf{x}^{k+1} = (I - \tau A)\mathbf{x}^k + \tau \mathbf{b}$ при $0 < \tau < 2/5$ сходится с произвольного начального приближения.

- 15.15.** Пусть λ и \mathbf{q} — собственное значение и соответствующий собственный вектор невырожденной матрицы простой структуры A , \mathbf{x}^0 — начальное приближение в методе простой итерации $\mathbf{x}^{k+1} = (I - \tau A)\mathbf{x}^k + \tau \mathbf{b}$ для решения системы $A\mathbf{x} = \mathbf{b}$. Найти такое значение параметра τ , чтобы в разложении ошибки по собственным векторам коэффициент при векторе \mathbf{q} на первой итерации был равен нулю.

- 15.16.** Пусть известны все собственные значения $\lambda_1, \dots, \lambda_n$ невырожденной матрицы простой структуры A . Построить итерационный метод $\mathbf{x}^{k+1} = (I - \tau_{k+1}A)\mathbf{x}^k + \tau_{k+1}\mathbf{b}$ с переменными τ_k , который не более чем за n шагов приводил бы при отсутствии округлений к решению системы $A\mathbf{x} = \mathbf{b}$.
- 15.17.** Пусть матрица простой структуры A имеет одно отрицательное собственное значение $\lambda_1 \in [-2 - \varepsilon, -2 + \varepsilon]$, $\varepsilon = 0,01$, а остальные значения — положительные: $\lambda_i \in [1, 3]$, $i = 2, \dots, n$. Предложить итерационный метод $\mathbf{x}^{k+1} = (I - \tau_{k+1}A)\mathbf{x}^k + \tau_{k+1}\mathbf{b}$ с переменными τ_k для решения системы $A\mathbf{x} = \mathbf{b}$.

- 15.18.** Для решения системы $\mathbf{x} = B\mathbf{x} + \mathbf{c}$ рассмотрим алгоритм

$$\mathbf{y}^{k+1} = B\mathbf{x}^k + \mathbf{c}, \quad \mathbf{x}^{k+1} = \alpha\mathbf{x}^k + (1 - \alpha)\mathbf{y}^{k+1}.$$

Пусть $B = B^T$, $\lambda(B) \in [m, M]$, $m > 1$. Найти оптимальное значение итерационного параметра α .

- 15.19.** Построить матрицу A размерности 31×31 с элементами $|a_{ij}| \leq 1$ и собственными значениями $|\lambda(A)| \leq 1$ такую, что $\|A^{30}\|_\infty \geq 10^9$.
- 15.20.** Пусть A — невырожденная матрица размерности $n \times n$ и X_0 — произвольная матрица размерности $n \times n$. Рассмотрим итерационный процесс $X_{k+1} = X_k + X_k(I - AX_k)$, $k = 0, 1, \dots$. Доказать, что $\lim_{k \rightarrow \infty} X_k = A^{-1}$ тогда и только тогда, когда спектральный радиус матрицы $I - AX_0$ меньше 1. При этом $I - AX_k = (I - AX_0)^{2^k}$, $k = 0, 1, \dots$. Доказать также, что если $AX_0 = X_0A$, то $AX_k = X_kA$ для всех k .
- 15.21.** При каких значениях параметра τ метод $\mathbf{x}^{k+1} = (I - \tau A)\mathbf{x}^k + \tau\mathbf{b}$ для системы уравнений $A\mathbf{x} = \mathbf{b}$ с матрицей:

$$1) A = \begin{pmatrix} 5 & 0,8 & 4 \\ 2,5 & 2 & 0 \\ 2 & 0,8 & 4 \end{pmatrix}; \quad 2) A = \begin{pmatrix} 2 & 1 & 0,5 \\ 3 & 5 & 1 \\ 1 & 3 & 3 \end{pmatrix};$$

$$3) A = \begin{pmatrix} 1 & 0,5 & 0,3 \\ 1 & 3 & 0 \\ 1 & 1 & 2 \end{pmatrix}; \quad 4) A = \begin{pmatrix} 3 & 1,2 & 0,8 \\ 1,4 & 2 & 0,1 \\ 0,6 & 0,4 & 1 \end{pmatrix}$$

сходится с произвольного начального приближения?

15.22. Пусть $A = A^T > 0$. Записать наилучший по скорости сходимости итерационный процесс вида

$$\mathbf{x}^{k+1} = \mathbf{x}^k - P_1(A)(A\mathbf{x}^k - \mathbf{b}), \quad P_1(t) = \alpha t + \beta.$$

15.23. Пусть приближения метода $\mathbf{x}^{k+1} = B\mathbf{x}^k + \mathbf{c}$, $\|B\| < 1$, сходятся к решению \mathbf{x} . Доказать, что $\|\mathbf{x} - \mathbf{x}^k\| \leq \|(I-B)^{-1}\| \|\mathbf{x}^{k+1} - \mathbf{x}^k\|$.

15.24. Для приближений метода $\mathbf{x}^{k+1} = B\mathbf{x}^k + \mathbf{c}$, $\|B\| < 1$, доказать оценку $\|\mathbf{x}^k\| \leq \|B\|^k \|\mathbf{x}^0\| + \|\mathbf{c}\|/(1 - \|B\|)$.

15.25. Пусть $A = I - B$, $b_{ij} \geq 0$. Доказать, что если все компоненты векторов \mathbf{x} и \mathbf{c} из задачи $A\mathbf{x} = \mathbf{c}$ неотрицательны, то приближения метода $\mathbf{x}^{k+1} = B\mathbf{x}^k + \mathbf{c}$, $\mathbf{x}^0 = 0$, сходятся к \mathbf{x} .

15.26. Пусть спектр матрицы A удовлетворяет условиям: $0 < \delta \leq \operatorname{Re}(\lambda(A)) \leq 1$, $|\operatorname{Im}(\lambda(A))| \leq 1$. Найти область значений вещественного параметра τ , при которых итерационный метод $\mathbf{x}^{k+1} = (I - \tau A)\mathbf{x}^k + \tau\mathbf{b}$ для системы $A\mathbf{x} = \mathbf{b}$ сходится с произвольного начального приближения.

15.27. Исследовать сходимость метода $\mathbf{x}^{k+1} = B\mathbf{x}^k + \mathbf{c}$ для решения системы уравнений $\mathbf{x} = B\mathbf{x} + \mathbf{c}$ с матрицей

$$B = \begin{pmatrix} 0 & 1/4 & 1/8 & 1/16 & \dots & 1/2^n & 1/2^{n+1} \\ 1/4 & 0 & 1/4 & 1/8 & \dots & 1/2^{n-1} & 1/2^n \\ 1/8 & 1/4 & 0 & 1/4 & \dots & 1/2^{n-2} & 1/2^{n-1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1/2^{n+1} & 1/2^n & 1/2^{n-1} & \dots & \dots & 1/4 & 0 \end{pmatrix}.$$

15.28. Построить сходящийся с произвольного начального приближения метод простой итерации $\mathbf{x}^{k+1} = (I - \tau A)\mathbf{x}^k + \tau\mathbf{b}$ для системы $A\mathbf{x} = \mathbf{b}$, матрица которой имеет четный порядок и состоит из элементов $a_{ij} = \{1 \text{ для } i = j = 1, 3, 5, \dots; 2 \text{ для } i = j = 2, 4, 6, \dots; 0,5 \text{ для } i = j - 1; 0 \text{ для остальных индексов}\}$.

15.29. При каких условиях на спектр матрицы B итерационный метод $\mathbf{x}^{k+1} = (2B^2 - I)\mathbf{x}^k + 2(B + I)\mathbf{c}$ сходится быстрее метода простой итерации $\mathbf{x}^{k+1} = B\mathbf{x}^k + \mathbf{c}$?

15.30. Пусть решение системы $A\mathbf{x} = \mathbf{b}$ с матрицей $A = A^T > 0$ ищется как точка минимума функционала $F(\mathbf{x}) = (A\mathbf{x}, \mathbf{x}) - 2(\mathbf{b}, \mathbf{x})$

(см. 13.40) по следующему алгоритму:

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \delta_k \text{grad } F(\mathbf{x}^k),$$

а δ_k выбирается из условия минимума величины $F(\mathbf{x}^k - \delta_k \text{grad } F(\mathbf{x}^k))$. Доказать, что $2\delta_k = (\mathbf{r}^k, \mathbf{r}^k) / (A\mathbf{r}^k, \mathbf{r}^k)$, где $\mathbf{r}^k = \mathbf{b} - A\mathbf{x}^k$ — вектор невязки (этот алгоритм называют методом *наискорейшего градиентного спуска*).

15.31. Пусть $A = A^T > 0$ и $\lambda(A) \in [m, M]$. Доказать, что метод наискорейшего градиентного спуска для решения системы $A\mathbf{x} = \mathbf{b}$ сходится с произвольного начального приближения и верна оценка:

$$\|\mathbf{z}^k\|_A \leq \left(\frac{M-m}{M+m} \right)^k \|\mathbf{z}^0\|_A, \quad \text{где } \mathbf{z}^k = \mathbf{x} - \mathbf{x}^k, \quad \|\mathbf{z}\|_A^2 = (A\mathbf{z}, \mathbf{z}).$$

15.32. Пусть $A = A^T > 0$ и $\lambda(A) \in [m, M]$. Доказать следующую оценку скорости сходимости метода наискорейшего градиентного спуска:

$$\|\mathbf{z}^k\|_2 \leq \left(1 - \frac{m}{M} \right)^k \|\mathbf{z}^0\|_2, \quad \text{где } \mathbf{z}^k = \mathbf{x} - \mathbf{x}^k.$$

§ 16. Методы релаксации

Скорость сходимости рассмотренных итерационных процессов зависела от отношения M/m границ спектра матрицы $A = A^T > 0$, т. е. от числа обусловленности $\text{cond}_2(A)$. Для «улучшения» исходной задачи можно перейти к некоторой эквивалентной системе $B^{-1}A\mathbf{x} = B^{-1}\mathbf{b}$ при условии невырожденности матрицы B . Применим метод

$$\frac{\mathbf{x}^{k+1} - \mathbf{x}^k}{\tau} + B^{-1}A\mathbf{x}^k = B^{-1}\mathbf{b}. \quad (1)$$

Неявные методы. Пусть $A = A^T > 0$. Перепишем итерационный алгоритм (1) в виде

$$B \frac{\mathbf{x}^{k+1} - \mathbf{x}^k}{\tau} + A\mathbf{x}^k = \mathbf{b}, \quad (2)$$

который часто называют *обобщенным методом простой итерации или методом с предобуславливателем B* .

Неявный двухслойный итерационный алгоритм (2) требует на каждом шаге решения задач вида $B\mathbf{y} = \mathbf{f}$ и совпадает с рассмотренными выше методами при $B = I$. Известно, что алгоритм (2) сходится при $B > \tau A/2$. Если дополнительно $B = B^T > 0$ и $m_1 B \leq A \leq M_1 B$, то при $\tau = 2/(m_1 + M_1)$ метод сходится со скоростью геометрической прогрессии с показателем $q = (M_1 - m_1)/(M_1 + m_1)$.

При удачном выборе оператора B можно принципиально улучшить скорость сходимости, однако необходимо учитывать трудоемкость нахождения $\mathbf{y} = B^{-1}\mathbf{f}$. Например, при $B = A$, $\tau = 1$ метод (2) сойдется за одну итерацию, но потребует решения исходной задачи $A\mathbf{x} = \mathbf{b}$.

Методы релаксации. Рассмотрим неявные методы с диагональной либо треугольной матрицей B . Представим матрицу системы $A\mathbf{x} = \mathbf{b}$ в виде $A = L + D + R$, где D — диагональная матрица, L и R — соответственно левая нижняя и правая верхняя треугольные матрицы с нулевыми диагоналями (строго нижняя и строго верхняя треугольные матрицы). Будем предполагать, что все диагональные элементы исходной матрицы a_{ii} отличны от нуля, следовательно, любая матрица вида $D + \omega L$ с произвольным параметром ω обратима.

Методы релаксации описывают формулой (2) с матрицей $B = D + \omega L$. Здесь итерационный параметр ω называется *параметром релаксации*. Методы *Якоби* ($\omega = 0, \tau = 1$), *Гаусса–Зейделя* ($\omega = \tau = 1$) и *верхней релаксации* (в англоязычной литературе — SOR) ($\omega = \tau$) удобно представить, соответственно, в виде

$$\begin{aligned} D(\mathbf{x}^{k+1} - \mathbf{x}^k) + A\mathbf{x}^k &= \mathbf{b}, \\ (D + L)(\mathbf{x}^{k+1} - \mathbf{x}^k) + A\mathbf{x}^k &= \mathbf{b}, \\ (D + \omega L) \frac{\mathbf{x}^{k+1} - \mathbf{x}^k}{\omega} + A\mathbf{x}^k &= \mathbf{b}. \end{aligned}$$

В случае $A = A^T > 0$ ($R = L^T$) используют также *симметричный метод релаксации* (в англоязычной литературе — SSOR):

$$\begin{aligned} (D + \omega L) \frac{\mathbf{x}^{k+1/2} - \mathbf{x}^k}{\omega} + A\mathbf{x}^k &= \mathbf{b}, \\ (D + \omega R) \frac{\mathbf{x}^{k+1} - \mathbf{x}^{k+1/2}}{\omega} + A\mathbf{x}^{k+1/2} &= \mathbf{b}. \end{aligned}$$

16.1. Для решения системы $A\mathbf{x} = \mathbf{b}$ с матрицей

$$A = \begin{pmatrix} \alpha & \beta & 0 \\ \beta & \alpha & \beta \\ 0 & \beta & \alpha \end{pmatrix}$$

применяются методы Якоби и Гаусса—Зейделя. Для каждого алгоритма найти все значения параметров α , β , обеспечивающие сходимость с произвольного начального приближения.

16.2. Доказать, что для систем линейных уравнений второго порядка методы Якоби и Гаусса—Зейделя сходятся и расходятся одновременно.

16.3. Пусть невырожденная матрица A обладает свойством строгого диагонального преобладания, т. е. для всех i справедливо неравенство

$$\sum_{j \neq i} |a_{ij}| \leq q |a_{ii}|, \quad 0 \leq q < 1.$$

Доказать, что для вектора ошибки в методе Гаусса—Зейделя имеет место оценка $\|\mathbf{x} - \mathbf{x}^k\|_\infty \leq q^k \|\mathbf{x} - \mathbf{x}^0\|_\infty$.

16.4. Исследовать сходимость метода Гаусса—Зейделя для матриц размерности $n \times n$ с элементами:

$$1) a_{ij} = 3^{-|i-j|}, \quad 2) a_{ij} = \begin{cases} 2 & \text{при } i = j, \\ -1 & \text{при } |i - j| = 1, \\ 0 & \text{при } |i - j| > 1. \end{cases}$$

16.5. Показать, что выполнение неравенства $0 < \tau < 2$ является необходимым для сходимости метода верхней релаксации.

16.6. Пусть матрица A простой структуры имеет собственные значения $\lambda(A) \in [m, M]$, $m > 0$. Доказать, что при любом положительном значении итерационного параметра τ сходится метод

$$\frac{\mathbf{x}^{k+1} - \mathbf{x}^k}{\tau} + A \left(\frac{\mathbf{x}^{k+1} + \mathbf{x}^k}{2} \right) = \mathbf{b}.$$

Определить оптимальное значение τ_{opt} .

16.7. При каких $\alpha \in [0, 1]$ для матрицы A из **16.6** метод

$$\frac{\mathbf{x}^{k+1} - \mathbf{x}^k}{\tau} + A(\alpha \mathbf{x}^{k+1} + (1 - \alpha)\mathbf{x}^k) = \mathbf{b}$$

сходится при любом $\tau > 0$?

16.8. Невырожденную систему $A\mathbf{x} = \mathbf{b}$ с матрицей $A = \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}$

решают методом Гаусса—Зейделя. Доказать, что:

- 1) если $|a| > 1$, то для некоторого начального приближения итерационный процесс не сходится;
- 2) если $|a| < 1$, то итерации сходятся при любом начальном приближении.

16.9. Построить пример системы уравнений третьего порядка, для которой метод Якоби сходится, а метод Гаусса—Зейделя расходится.

16.10. Построить пример системы уравнений третьего порядка, для которой метод Гаусса—Зейделя сходится, а метод Якоби расходится.

16.11. Доказать, что обобщенный метод простой итерации

$$B \frac{\mathbf{x}^{k+1} - \mathbf{x}^k}{\tau} + A\mathbf{x}^k = \mathbf{b}, \quad A = A^T > 0, \quad \tau > 0,$$

сходится при условии $B - \frac{\tau}{2}A > 0$ (т. е. $(B\mathbf{x}, \mathbf{x}) > > \frac{\tau}{2}(A\mathbf{x}, \mathbf{x}) \forall \mathbf{x} \neq 0$).

16.12. Пусть $A = A^T > 0$. Доказать, что метод верхней релаксации сходится с любого начального приближения при $\tau \in (0, 2)$.

16.13. Пусть $B = L + R$, где L — нижняя треугольная матрица с нулями на диагонали, R — верхняя треугольная матрица. Пусть, далее, $\|B\|_\infty < 1$, так что итерационный процесс $\mathbf{x}^{k+1} = B\mathbf{x}^k + \mathbf{c}$ сходится. Доказать, что метод $\mathbf{x}^{k+1} = L\mathbf{x}^{k+1} + R\mathbf{x}^k + \mathbf{c}$ также сходится.

16.14. Для системы уравнений $A\mathbf{x} = \mathbf{b}$ вида

$$4x_{i,j} - x_{i+1,j} - x_{i-1,j} - x_{i,j+1} - x_{i,j-1} = h^2 b_{ij}, \quad i, j = 1, 2, \dots, n-1; \\ x_{0,i} = x_{i,0} = x_{n,i} = x_{i,n} = 0, \quad i = 0, 1, \dots, n; \quad nh = 1,$$

записать расчетные формулы и найти асимптотическую скорость сходимости следующих итерационных методов: 1) метода Якоби; 2) метода Гаусса—Зейделя; 3) метода верхней релаксации с оптимальным параметром релаксации; 4) метода симметричной верхней релаксации с оптимальным параметром релаксации.

- 16.15.** Исследовать сходимость метода Якоби для решения системы уравнений $A\mathbf{x} = \mathbf{b}$ с матрицей

$$A = \begin{pmatrix} 2 & -0,2 & 0,3 & 0,4 \\ 0,3 & -3 & 1 & -1,4 \\ 0,4 & 0,8 & 4 & 2,4 \\ -0,5 & 1,2 & -2,5 & -5 \end{pmatrix}.$$

- 16.16.** Найти все α, β , при которых метод Гаусса—Зейделя является сходящимся для системы уравнений $A\mathbf{x} = \mathbf{b}$ с матрицей:

$$1) A = \begin{pmatrix} \alpha & 0 & \beta \\ 0 & \alpha & 0 \\ \beta & 0 & \alpha \end{pmatrix}; \quad 2) A = \begin{pmatrix} \alpha & \beta & 0 \\ \beta & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix}; \quad 3) A = \begin{pmatrix} \alpha & \alpha & 0 \\ \alpha & \beta & \beta \\ 0 & \beta & \alpha \end{pmatrix}.$$

- 16.17.** Пусть матрицы $A_i, i = 1, 2$, простой структуры имеют собственные значения $\lambda(A_i) \in [m, M], m > 0$ и $A_1A_2 = A_2A_1, A = A_1 + A_2$. Доказать, что при любом положительном значении параметра τ сходится следующий итерационный метод решения системы уравнений $A\mathbf{x} = \mathbf{b}$:

$$\frac{\mathbf{x}^{k+1/2} - \mathbf{x}^k}{\tau} + A_1\mathbf{x}^{k+1/2} + A_2\mathbf{x}^k = \mathbf{b},$$

$$\frac{\mathbf{x}^{k+1} - \mathbf{x}^{k+1/2}}{\tau} + A_1\mathbf{x}^{k+1/2} + A_2\mathbf{x}^{k+1} = \mathbf{b}.$$

Определить оптимальное значение τ_{opt} .

- 16.18.** Доказать сходимость итерационного метода из **16.17**, если матрицы простой структуры A_1, A_2 положительно определены, но не обязательно выполнено $A_1A_2 = A_2A_1$.

- 16.19.** Пусть матрицы $A_i, i = 1, 2$, простой структуры имеют собственные значения $\lambda(A_i) \in [m, M], m > 0$ и $A_1A_2 = A_2A_1, A = A_1 + A_2$. Доказать, что при любом $\tau > 0$ сходится

следующий итерационный метод решения системы уравнений $A\mathbf{x} = \mathbf{b}$:

$$\frac{\mathbf{x}^{k+1/2} - \mathbf{x}^k}{\tau} + A_1\mathbf{x}^{k+1/2} + A_2\mathbf{x}^k = \mathbf{b},$$

$$\frac{\mathbf{x}^{k+1} - \mathbf{x}^{k+1/2}}{\tau} + A_2(\mathbf{x}^{k+1} - \mathbf{x}^k) = 0.$$

Определить оптимальное значение τ_{opt} .

- 16.20.** Доказать сходимость итерационного процесса из **16.19**, если матрицы простой структуры A_1, A_2 положительно определены, но не обязательно выполнено $A_1A_2 = A_2A_1$.
- 16.21.** Показать, что если матрица $A = M - N$ вырожденная, то нельзя получить оценку $\rho(M^{-1}N) < 1$ ни для какой невырожденной матрицы M .
- 16.22.** Пусть $A = M - N$ и итерации $M\mathbf{x}^{k+1} = N\mathbf{x}^k + \mathbf{b}$ сходятся при произвольном начальном приближении. Доказать, что $\rho(M^{-1}N) < 1$.
- 16.23.** Пусть решаются задачи $A_i\mathbf{x} = \mathbf{b}_i$, $i = 1, 2$, где

$$A_1 = \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & -3/4 \\ -1/12 & 1 \end{pmatrix},$$

B_1 и B_2 — соответствующие этим матрицам операторы перехода в итерационном методе Якоби. Показать, что $\rho(B_1) > \rho(B_2)$, т. е. опровергнуть мнение о том, что относительное усиление диагонального преобладания обязательно влечет за собой более быструю сходимость метода Якоби.

§ 17. Задачи на собственные значения

Число $\lambda \in \mathbb{C}^1$ и ненулевой вектор $\mathbf{x} \in \mathbb{C}^n$ называют *собственным значением* и *собственным вектором* матрицы A размерности $n \times n$, если справедливо равенство $A\mathbf{x} = \lambda\mathbf{x}$.

Пусть S — произвольная невырожденная матрица. Говорят, что матрицы одинаковой размерности A и $B = S^{-1}AS$ *подобны*, а матрица S осуществляет подобие.

- 17.1.** Пусть $B = S^{-1}AS$. Показать, что матрицы A и B имеют одни и те же собственные значения, а \mathbf{x}_A является собственным

вектором A тогда и только тогда, когда $\mathbf{x}_B = S^{-1}\mathbf{x}_A$ — собственный вектор B .

17.2. Пусть A — симметричная матрица размерности $n \times n$, $\lambda \in \mathbb{R}^1$, $\mathbf{x} \in \mathbb{R}^n$ — соответственно произвольное число и произвольный вектор, причем $\|\mathbf{x}\|_2 = 1$. Доказать, что существует собственное значение λ_k матрицы A , для которого $|\lambda_k - \lambda| \leq \|A\mathbf{x} - \lambda\mathbf{x}\|_2$.

17.3. Показать, что для максимального и минимального собственных значений симметричной матрицы A справедливы оценки:

$$\lambda_{\min}(A) \leq \min_{1 \leq i \leq n} a_{ii}; \quad \lambda_{\max}(A) \geq \max_{1 \leq i \leq n} a_{ii}.$$

17.4. Доказать, что у вещественной трехдиагональной матрицы A с элементами $a_{ij} = \{b_i \text{ для } i = j; a_i \text{ для } i = j + 1; c_i \text{ для } i = j - 1; 0 \text{ для остальных индексов}\}$ все собственные значения вещественные, если $a_{i+1}c_i > 0$, $i = 1, 2, \dots, n - 1$.

17.5. Доказать, что для трехдиагональной матрицы из **17.4** верно неравенство $|\lambda_k(A)| < 1 \forall k$, если $|a_i| + |b_i| + |c_i| \leq 1 \forall i$, $a_1 = c_n = 0$, и хотя бы для одного значения индекса i неравенство строгое, а $a_{i+1}c_i \neq 0$, $i = 1, 2, \dots, n - 1$.

17.6. Пусть A и B — матрицы размерности $m \times n$ и $n \times m$ соответственно, $m \geq n$; $P_C(\lambda) = \det(\lambda I - C)$ — характеристический многочлен квадратной матрицы C . Доказать справедливость равенства

$$P_{AB}(\lambda) = \lambda^{m-n} P_{BA}(\lambda).$$

17.7. Доказать, что для квадратных матриц A, B одинаковой размерности спектры матриц AB и BA совпадают.

17.8. Доказать, что если A, B — симметричные матрицы размерности $n \times n$, то необходимым и достаточным условием равенства $AB = BA$ является существование базиса в пространстве \mathbb{R}^n , составленного из общих собственных векторов матриц A и B .

17.9. Доказать, что если матрицы A и B коммутируют, то существует собственное значение $\lambda(AB)$, равное произведению собственных значений $\lambda(A)\lambda(B)$.

- 17.10. Доказать, что если A — симметричная и положительно определенная матрица, а B — симметричная матрица, то все собственные значения $\lambda(AB)$ матрицы AB вещественные.
- 17.11. Доказать, что если A, B — симметричные и положительно определенные матрицы, то все собственные значения $\lambda(AB)$ матрицы AB положительные.
- 17.12. Пусть A — симметризуемая матрица, т. е. существует невырожденная матрица S такая, что SAS^{-1} — симметричная матрица. Доказать, что система собственных векторов матрицы A образует базис.
- 17.13. Доказать, что если A — симметричная и положительно определенная матрица, а B — симметричная матрица, то система собственных векторов матрицы AB образует базис.
- 17.14. Доказать, что если A, B — симметричные и положительно определенные, коммутирующие матрицы, то матрица AB положительно определена.
- 17.15. Доказать положительную определенность матрицы

$$1) A = \begin{pmatrix} 2 & -1 & 1/2 & -1/3 \\ -1 & 3 & -1 & -1/2 \\ 1/2 & -1 & 4 & 2 \\ -1/3 & -1/2 & 2 & 5 \end{pmatrix},$$

$$2) A = \begin{pmatrix} 12 & -6 & 3 & -2 \\ -6 & 18 & -6 & 6 \\ 3 & -6 & 24 & 15 \\ -2 & 6 & 15 & 20 \end{pmatrix}.$$

- 17.16. Пусть матрицы $A, A^T \in \mathbb{R}^{n \times n}$ имеют строгое диагональное преобладание и положительные диагональные элементы. Доказать, что матрица A положительно определена, т. е. $(Ax, x) > 0 \forall x \neq 0$.
- 17.17. Построить пример симметричной положительно определенной матрицы размерности 3×3 , трехдиагональная часть которой не является положительно определенной.
- 17.18. Пусть $A = A^T > 0$. Доказать, что если $\lambda_{\max}(A) = a_{kk}$ при некотором k , где $1 \leq k \leq n$, то $a_{ik} = a_{kj} = 0$ при всех $i \neq k$, $j \neq k$.

17.19. Пусть $A_n(a, b)$ — вещественная трехдиагональная $n \times n$ матрица с элементами $a_{ij} = \{a \text{ для } i = j; b \text{ для } |i - j| = 1; 0 \text{ для остальных индексов}\}$. Доказать следующие равенства:

1) $\det A_{n+1}(a, b) = a \det A_n(a, b) - b^2 \det A_{n-1}(a, b), n \geq 2;$

2) $\det A_n(a, b) = \left(\left(a/2 + \sqrt{a^2/4 - b^2} \right)^{n+1} - \left(a/2 - \sqrt{a^2/4 - b^2} \right)^{n+1} \right) / 2\sqrt{a^2/4 - b^2}, n \geq 1;$

3) $\det A_n(a, b) = \sum_{k=0}^{[n/2]} C_{n+1}^{2k+1} (a^2/4 - b^2)^k (a/2)^{n-2k}, k \geq 1.$

17.20. Найти все собственные значения и собственные векторы матрицы $A_n(a, b)$ из **17.19**.

17.21. Доказать, что матрица $A_n(a, b)$ из **17.19** положительно определена тогда и только тогда, когда $a - 2|b| \cos \frac{\pi}{n+1} > 0$.

17.22. Матрица Уилкинсона

$$A = \begin{pmatrix} 20 & 20 & 0 & 0 & \dots & 0 & 0 \\ 0 & 19 & 20 & 0 & \dots & 0 & 0 \\ 0 & 0 & 18 & 20 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 2 & 20 \\ \varepsilon & 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

при $\varepsilon = 0$ имеет наименьшее по модулю собственное значение, равное 1. Как оно изменится при $\varepsilon = 20^{-19} \cdot 20! \approx 5 \cdot 10^{-7}$?

17.23. Доказать, что характеристическое уравнение $n \times n$ матрицы

$$A_n(\alpha) = \begin{pmatrix} 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ \alpha & 0 & \dots & 0 \end{pmatrix}$$

имеет вид $\lambda^n = \alpha$. Найти собственные числа близких матриц:

- 1) $A_{20}(2^{-20}),$ 2) $A_{20}(0).$

Степенной метод. Алгоритм вычисления максимального по модулю собственного значения λ матрицы A имеет вид

$$\mathbf{x}^{k+1} = A\mathbf{x}^k, \quad \lambda^{k+1} = \frac{(\mathbf{x}^{k+1}, \mathbf{x}^k)}{(\mathbf{x}^k, \mathbf{x}^k)}, \quad \mathbf{x}^k \neq 0; \quad k = 0, 1, 2, \dots$$

При его практической реализации на каждом шаге текущее приближение нормируют, т. е. заменяют \mathbf{x}^k на $\mathbf{x}^k / \|\mathbf{x}^k\|$.

Метод обратной итерации. Этот метод, по сути соответствующий степенному методу для матрицы A^{-1} , можно применять для вычисления наименьшего по модулю собственного значения λ :

$$A\mathbf{x}^{k+1} = \mathbf{x}^k, \quad \lambda^{k+1} = \frac{(\mathbf{x}^k, \mathbf{x}^{k+1})}{(\mathbf{x}^{k+1}, \mathbf{x}^{k+1})}.$$

При этом на каждом шаге алгоритма требуется решать систему $A\mathbf{x}^{k+1} = \mathbf{x}^k$ с нормированным вектором \mathbf{x}^k .

Степенной метод и метод обратной итерации можно также применять к матрице со *сдвигом* $A - cI$, что позволяет влиять на сходимость. При этом допустимо изменение сдвига от шага к шагу. В частности, скорость сходимости метода обратной итерации можно значительно повысить, если выбирать в качестве переменного сдвига c_k отношение *Рэля* для текущего приближения, т. е. величину

$$R_A(\mathbf{x}^k) = \frac{(A\mathbf{x}^k, \mathbf{x}^k)}{(\mathbf{x}^k, \mathbf{x}^k)}.$$

17.24. Пусть A — матрица простой структуры (собственные векторы $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$ матрицы образуют базис в \mathbb{C}^n). Пусть далее $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$ и L — линейная оболочка векторов $\mathbf{q}_2, \mathbf{q}_3, \dots, \mathbf{q}_n$. Доказать, что для степенного метода при условии $\mathbf{x}^0 \notin L$ справедлива оценка $\lambda^k = \lambda_1 + O(|\lambda_2/\lambda_1|^k)$.

17.25. Доказать, что если в условии **17.24** матрица A является симметричной, то для степенного метода справедлива оценка $\lambda^k = \lambda_1 + O(|\lambda_2/\lambda_1|^{2k})$.

17.26. Пусть матрица A размерности $n \times n$ имеет n различных собственных значений. Предположим, что \mathbf{x}^0 принадлежит линейной оболочке некоторых собственных векторов $\mathbf{q}_{i_1}, \mathbf{q}_{i_2}, \dots, \mathbf{q}_{i_t}$, но не принадлежит никакой их линейной оболочке. К какому собственному значению матрицы сходятся

ся итерации степенного метода при отсутствии округлений и с какой скоростью?

- 17.27.** Пусть собственные значения симметричной матрицы A удовлетворяют цепочке неравенств $\lambda_1 < \lambda_2 < \dots < \lambda_n$. Выяснить, к какому собственному значению λ_s в зависимости от параметра c и с какой скоростью сходится итерационный процесс

$$\mathbf{x}^{k+1} = (A - cI)\mathbf{x}^k, \quad \lambda^{k+1} = c + \frac{(\mathbf{x}^{k+1}, \mathbf{x}^k)}{(\mathbf{x}^k, \mathbf{x}^k)}.$$

- 17.28.** В условии **17.27** выбрать постоянную c так, чтобы итерационный процесс с наилучшей скоростью сходил к λ_1 (или к λ_n).

- 17.29.** Пусть собственные значения симметричной матрицы A удовлетворяют цепочке неравенств $\lambda_1 < \lambda_2 < \dots < \lambda_n$. Выяснить, к какому собственному значению λ_t сходится метод обратной итерации со сдвигом

$$(A - cI)\mathbf{x}^{k+1} = \mathbf{x}^k, \quad \lambda^{k+1} = c + \frac{(\mathbf{x}^k, \mathbf{x}^{k+1})}{(\mathbf{x}^{k+1}, \mathbf{x}^{k+1})}$$

в зависимости от параметра c . Найти скорость сходимости.

- 17.30.** Пусть $A = A^T > 0$. Доказать, что

$$\lambda_{\max}(A) = \max_{\mathbf{x} \neq 0} R_A(\mathbf{x}), \quad \lambda_{\min}(A) = \min_{\mathbf{x} \neq 0} R_A(\mathbf{x}), \quad \text{где } R_A(\mathbf{x}) = \frac{(A\mathbf{x}, \mathbf{x})}{(\mathbf{x}, \mathbf{x})}.$$

- 17.31.** Пусть $A = A^T > 0$. Доказать, что $\forall \mathbf{x} \in \mathbb{R}^n$ и $\forall \mu \in \mathbb{R}^1$ имеет место свойство минимальности невязки $\|(A - R_A(\mathbf{x})I)\mathbf{x}\|_2 \leq \|(A - \mu I)\mathbf{x}\|_2$.

- 17.32.** Пусть метод обратной итерации с постоянным сдвигом c сходится к собственному значению λ_c матрицы $A = A^T$. Показать, что начиная с некоторого k выполняется оценка $|\lambda_c - \lambda^{k+1}| \leq 1/\|\mathbf{x}^{k+1}\|_2$, т.е. величина $\|\mathbf{x}^{k+1}\|_2^{-1}$ характеризует скорость сходимости итерационного процесса.

Глава V

Решение нелинейных уравнений



Итерационные методы вычисления изолированного (отделенного от других) корня z уравнения $f(x) = 0$, как правило, требуют указания какой-либо области D , содержащей этот единственный корень, и алгоритма нахождения очередного приближения x_{n+1} по уже имеющимся x_n, \dots, x_{n-k} .

Широко используемые способы отделения корней — графический и табличный — базируются на свойствах гладкости функции; в случае, когда $f(x)$ является многочленом степени n , имеются аналитические подходы.

Если $f(x)$ — непрерывная, то вещественный корень z принадлежит любому отрезку, на концах которого функция имеет значения разных знаков. Деля отрезок пополам, получаем универсальный метод вычисления корня (метод бисекции). Этот подход не требует знания хорошего начального приближения. Если оно имеется, то для гладких функций используются более эффективные методы.

Пусть z — единственный на отрезке $[a, b]$ корень уравнения $f(x) = 0$ с непрерывной функцией $f(x)$. Если в его окрестности функция представима в виде $f(x) = (x - z)^p g(x)$, где p — натуральное число, а $g(x)$ — ограниченная функция такая, что $g(z) \neq 0$, то p называют *кратностью* корня. Если $p = 1$, то корень называют *простым*. При нечетном p функция $f(x)$ меняет знак на $[a, b]$, т. е. $f(a)f(b) < 0$, а при четном p — нет.

Итерационный метод решения порождает последовательность приближений x_n , которая сходится к корню: $\lim_{n \rightarrow \infty} |x_n - z| = 0$.

Величину $e_n = |x_n - z|$ называют *абсолютной ошибкой* на n -й итерации. Итерационный метод имеет *порядок t* (или *скорость сходимости t*), если t — наибольшее положительное число, для которого существует такая постоянная $q > 0$, что

$$\limsup_{n \rightarrow \infty} \frac{e_{n+1}}{e_n^t} \leq q < \infty.$$

Постоянную q называют *константой асимптотической ошибки*, ее обычно оценивают через производные функции $f(x)$ в точке

$x = z$. При $m = 1$ ($q \in (0, 1)$) сходимость называется *линейной*, при $1 < m < 2$ — *сверхлинейной*, при $m = 2$ — *квадратичной* и т. д. Из сходимости с порядком $m > 1$ следует оценка $e_{n+1} \leq q_n e_n$, $q_n \rightarrow 0$ при $n \rightarrow \infty$. При этом $e_{n+1} \leq e_0 \prod_{i=0}^n q_i$. Иногда скорость сходимости может замедляться при приближении к искомому решению, что соответствует $q_n \rightarrow 1$, но $e_n \rightarrow 0$ при $n \rightarrow \infty$. Таким свойством обладают методы с *полиномиальной* скоростью сходимости $e_n \leq e_0(1 + \alpha n e_0^l)^{-1/l}$. Данная оценка верна, например, если $e_{n+1} \leq (1 - \alpha e_n^l)e_n$ с некоторыми $l \geq 1$ и $0 < \alpha \leq e_0^{-l}$. Для методов с полиномиальной скоростью сходимости число итераций n , необходимое для достижения ошибки порядка ε , имеет асимптотику $n = O(\varepsilon^{-l})$, что существенно ограничивает их применение для расчетов с высокой точностью.

Особое внимание в теории решения нелинейных уравнений уделяется методам со сверхлинейной скоростью сходимости. При практических расчетах традиционно применяют методы с квадратичной сходимостью, так как итерационные процессы более высокого порядка ($m > 2$) обычно требуют серьезного увеличения вычислительных затрат.

§ 18. Метод простой итерации и смежные вопросы

Исходное уравнение $f(x) = 0$ часто заменяют эквивалентным ему уравнением $x = \varphi(x)$. Эту замену можно сделать, положив, например, $\varphi(x) = x + \psi(x)f(x)$, где $\psi(x)$ — произвольная непрерывная и знакопостоянная функция.

Метод простой итерации. Выберем некоторое начальное приближение $x_0 \in [a, b]$ к корню z , дальнейшие приближения будем вычислять по формуле

$$x_{n+1} = \varphi(x_n), \quad n = 0, 1, 2, \dots$$

Последовательность x_n стремится к z , например, если отображение $y = \varphi(x)$ является *сжимающим*, т. е. при некотором $0 < q < 1$ выполнено условие $\rho(\varphi(x_1), \varphi(x_2)) \leq q \rho(x_1, x_2)$, либо *слабо сжимающим*, т. е. при некоторых $\alpha > 0, l \geq 1$ выполнено условие

$$\rho(\varphi(x_1), \varphi(x_2)) \leq \frac{\rho(x_1, x_2)}{(1 + \alpha \rho^l(x_1, x_2))^{1/l}}.$$

Здесь $\rho(x_1, x_2)$ — расстояние между точками x_1 и x_2 .

Отметим, что гарантировать сходимость метода можно либо при условии, что соответствующие оценки выполняются для всех точек $x_{1,2} \in \mathbb{R}^1$, либо для точек $x_{1,2} \in [a, b]$, но дополнительно требуется, чтобы $\varphi(x) \in [a, b] \forall x \in [a, b]$. Из второго условия следует, что все приближения x_n принадлежат отрезку $[a, b]$.

Метод секущих. Пусть x_{n-1} и x_n — последовательные приближения к корню. Заменим кривую $y = f(x)$ прямой, проходящей через точки $(x_{n-1}, f(x_{n-1}))$ и $(x_n, f(x_n))$. В качестве следующего приближения к корню возьмем точку пересечения этой прямой с осью абсцисс. Расчетная формула принимает вид

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n).$$

Метод хорд. Пусть $f(a)f(b) < 0$. Идея метода (его еще называют методом *ложного положения*) состоит в замене кривой $y = f(x)$ хордами, проходящими через концы отрезков, в которых $f(x)$ имеет противоположные знаки. Метод хорд требует, чтобы один конец отрезка, на котором ищется корень, был неподвижен. В качестве неподвижного конца x_0 выбирают конец отрезка, для которого знак $f(x)$ совпадает со знаком второй производной $f''(x)$. Расчетная формула имеет вид

$$x_{n+1} = x_n - \frac{x_n - x_0}{f(x_n) - f(x_0)} f(x_n).$$

Метод парабол. Пусть x_{n-2}, x_{n-1} и x_n — три последовательных приближения к корню. Заменим кривую $y = f(x)$ параболой, проходящей через точки $(x_{n-2}, f(x_{n-2}))$, $(x_{n-1}, f(x_{n-1}))$ и $(x_n, f(x_n))$. В качестве следующего приближения к корню возьмем ближайшую к x_n точку пересечения этой параболы с осью абсцисс. Этот подход исключительно эффективен для нахождения корней многочлена как с действительными, так и с комплексными коэффициентами.

При нахождении кратных корней ($p > 1$) для большинства алгоритмов характерно замедление скорости сходимости.

- 18.1.** Пусть уравнение $f(x) \equiv x - \varphi(x) = 0$ имеет корень z , для его вычисления применяется метод простой итерации $x_{n+1} = \varphi(x_n)$, функция $\varphi(x)$ имеет непрерывную производную в открытой окрестности z и $|\varphi'(z)| \leq q < 1$. Доказать, что метод сходится к z при выборе начального приближения x_0 из некоторой окрестности корня.

18.2. Дополнительно к условиям **18.1** потребуем $\varphi'(z) = 0$, непрерывность и ограниченность $\varphi''(x)$. Доказать, что для метода $x_{n+1} = \varphi(x_n)$ в окрестности корня z верна квадратичная оценка сходимости

$$|z - x_{n+1}| \leq C(z - x_n)^2.$$

18.3. Пусть на некотором отрезке $Q_\delta = [a - \delta, a + \delta]$ функция $\varphi(x)$ удовлетворяет условию Липшица $|\varphi(x') - \varphi(x'')| \leq q|x' - x''|$ с константой $q < 1$ и в точке a выполняется неравенство

$$|a - \varphi(a)| \leq (1 - q)\delta.$$

Показать, что на отрезке Q_δ уравнение $f(x) \equiv x - \varphi(x) = 0$ имеет единственный корень z и последовательность $x_{n+1} = \varphi(x_n)$ сходится к корню z для произвольного $x_0 \in Q_\delta$.

18.4. Пусть уравнение $f(x) = 0$ имеет корень на отрезке $[a, b]$, причем $f(x)$ дифференцируема, а $f'(x)$ знакопостоянна на этом отрезке. Построить равносильное уравнение вида $x = \varphi(x)$, для которого на $[a, b]$ выполнено условие $|\varphi'(x)| \leq q < 1$.

18.5. Построить итерационный процесс вычисления всех корней уравнения $f(x) \equiv x^3 + 3x^2 - 1 = 0$ методом простой итерации.

18.6. Определить область начальных приближений x_0 , для которых итерационный процесс $x_{n+1} = (x_n^3 + 1)/20$ сходится.

18.7. Оценить скорость сходимости итерационного процесса $x_{n+1} = x_n - x_n^3 + x_n^4$ к корню $z = 0$ при малых x_0 .

18.8. Пусть z — простой корень уравнения $f(x) = 0$. Оценить скорость сходимости метода хорд в его окрестности.

18.9. Пусть z — простой корень уравнения $f(x) = 0$. Оценить скорость сходимости метода секущих.

18.10. Доказать, что все корни уравнения

$$f(x) \equiv a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = 0$$

расположены в кольце $\frac{|a_0|}{b + |a_0|} \leq |z| \leq 1 + \frac{c}{|a_n|}$, где

$$c = \max\{|a_0|, |a_1|, \dots, |a_{n-1}|\}, \quad b = \max\{|a_1|, |a_2|, \dots, |a_n|\}.$$

18.11. Доказать, что если при $x = a$ имеют место неравенства

$$f(a) > 0, f'(a) > 0, \dots, f^{(n)}(a) > 0,$$

то уравнение $f(x) \equiv a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = 0$ не имеет действительных корней, больших a .

18.12. Найти границы действительных корней уравнения

$$x^4 - 35x^3 + 380x^2 - 1350x + 1000 = 0.$$

18.13. Пусть $x_{n+1} = \sqrt{x_n + 2}$. Доказать, что $\lim_{n \rightarrow \infty} x_n = 2$ для любого $x_0 \geq -2$.

18.14. Доказать, что итерационный процесс $x_{n+1} = \cos x_n$ сходится для любого начального приближения $x_0 \in \mathbb{R}^1$.

18.15. Исследовать сходимость метода простой итерации $x_{n+1} = x_n^2 - 2x_n + 2$ в зависимости от выбора начального приближения x_0 .

18.16. Уравнение $x = 2x^{-1}$, имеющее два корня $z_1 = 1$ и $z_2 = 2$, решают методом простой итерации. Исследовать его сходимость в зависимости от выбора начального приближения x_0 .

18.17. Доказать, что метод простой итерации для решения уравнения $x = \varphi(x)$ сходится при любом начальном приближении:

1) $\varphi(x) = \alpha \sin^2 x + \beta \cos^2 x + \gamma$, где $|\alpha - \beta| < 1$;

2) $\varphi(x) = a e^{-b x^2} + c$, где $b \geq 0$, $2a^2 b < e$.

18.18. Уравнение $x + \ln x = 0$, имеющее корень $z \approx 0,6$, предлагается решать одним из следующих методов простой итерации:

1) $x_{n+1} = -\ln x_n$; 2) $x_{n+1} = e^{-x_n}$;

3) $x_{n+1} = \frac{x_n + e^{-x_n}}{2}$; 4) $x_{n+1} = \frac{3x_n + 5e^{-x_n}}{8}$.

Исследовать сходимость этих методов и сравнить их скорости.

18.19. Пусть функция $\varphi'(x)$ непрерывна на отрезке $[z - \delta, z + \delta]$, где z — единственная неподвижная точка для $\varphi(x)$. Может ли метод простой итерации сходиться к z , если $|\varphi'(z)| = 1$? Может ли он расходиться в этом случае?

18.20. Показать, что для всякого a существует единственное решение $z(a, \varepsilon)$ уравнения $x + \varepsilon \sin x + a = 0$ при $|\varepsilon| \leq 1$.

18.21. Найти область сходимости метода простой итерации для следующих уравнений:

$$1) x = e^{2x} - 1, \quad 2) x = 1/2 - \ln x, \quad 3) x = \operatorname{tg} x.$$

18.22. Записать расчетные формулы метода парабол и найти корни уравнения $2x + \lg x = -0,5$ с точностью 10^{-2} .

18.23. Оценить скорость сходимости метода секущих в окрестности p -кратного корня z уравнения $f(x) = 0$.

§ 19. Метод Ньютона. Итерации высшего порядка

Метод Ньютона. В случае одного уравнения $f(x) = 0$ формула метода Ньютона имеет вид

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Геометрически метод состоит в замене дуги кривой $y = f(x)$ касательной к ней в процессе каждой итерации. Это видно из уравнения касательной, проведенной в точке $(x_n, f(x_n))$:

$$y - f(x_n) = f'(x_n)(x - x_n),$$

из которого следует формула итерационного процесса, если положить $y = 0$ и $x = x_{n+1}$.

Метод Ньютона соответствует методу простой итерации $x_{n+1} = x_n - \tau_n f(x_n)$ с оптимальным, в некотором смысле, переменным параметром τ_n . Действительно, пусть z — изолированный простой (т. е. $f'(z) \neq 0$) корень, пусть также z и все x_n принадлежат некоторому отрезку $[a, b]$. Тогда

$$z - x_{n+1} = z - x_n + \tau_n f(x_n) - \tau_n f(z) = (1 - \tau_n f'(\xi_n))(z - x_n),$$

следовательно, при $\tau_n = 1/f'(\xi_n)$ метод сходится за одну итерацию. Точка ξ_n неизвестна, поэтому на текущем шаге выбираем $\tau_n = 1/f'(x_n)$, при этом верна оценка $|z - x_{n+1}| \leq \max_{\xi \in [a, b]} |1 - \tau_n f'(\xi)| |z - x_n|$.

В случае системы m нелинейных уравнений

$$\mathbf{F}(\mathbf{x}) = 0, \quad \text{где } \mathbf{x} = (x_1, \dots, x_m)^T, \quad \mathbf{F} = (f_1, \dots, f_m)^T,$$

будем считать отображение $\mathbf{F}: \mathbb{R}^m \rightarrow \mathbb{R}^m$ непрерывно дифференцируемым в некоторой окрестности решения \mathbf{z} , так что

$$\mathbf{F}'(\mathbf{x}) = \left[\frac{\partial f_i(\mathbf{x})}{\partial x_j} \right], \quad 1 \leq i, j \leq m.$$

В предположении обратимости этого оператора метод Ньютона можно записать в виде

$$\mathbf{x}_{n+1} = \mathbf{x}_n - (\mathbf{F}'(\mathbf{x}_n))^{-1} \mathbf{F}(\mathbf{x}_n).$$

Введем обозначение: $\Omega_a = \{\mathbf{x} : \|\mathbf{x} - \mathbf{z}\| < a\}$, где $\|\cdot\|$ — норма в \mathbb{R}^m . Пусть при некоторых $a, a_1, a_2 : 0 < a, 0 \leq a_1, a_2 < \infty$, выполнены следующие условия:

- 1) $\|(\mathbf{F}'(\mathbf{x}))^{-1} \mathbf{y}\| \leq a_1 \|\mathbf{y}\|$ при $\mathbf{x} \in \Omega_a$ и $\forall \mathbf{y}$;
 - 2) $\|\mathbf{F}(\mathbf{u}_1) - \mathbf{F}(\mathbf{u}_2) - \mathbf{F}'(\mathbf{u}_2)(\mathbf{u}_1 - \mathbf{u}_2)\| \leq a_2 \|\mathbf{u}_1 - \mathbf{u}_2\|^2$ при $\mathbf{u}_1, \mathbf{u}_2 \in \Omega_a$.
- Обозначим также $c = a_1 a_2, b = \min(a, c^{-1})$.

Теорема. При условиях 1), 2) и $\mathbf{x}_0 \in \Omega_b$ метод Ньютона сходится с оценкой погрешности $\|\mathbf{x}_n - \mathbf{z}\| \leq c^{-1} (c \|\mathbf{x}_0 - \mathbf{z}\|)^{2^n}$, т. е. квадратично.

Условия теоремы гарантируют, что корень \mathbf{z} простой. В случае двукратного корня ($p = 2$) метод Ньютона сходится линейно; скорость сходимости замедляется при повышении кратности.

Метод Чебышёва. Пусть z — простой корень уравнения $f(x) = 0$ и $F(y)$ — обратная к $f(x)$ функция. Тогда $x \equiv F(f(x))$ и $z = F(0)$. Разложим $F(0)$ в ряд Тейлора в окрестности некоторой точки y :

$$F(0) = F(y) + \sum_{k=1}^m F^{(k)}(y) \frac{(-y)^k}{k!} + \dots$$

Приблизим значение $F(0)$ значением частичной суммы в точке $y = f(x)$:

$$z = F(0) \approx \varphi_m(x) = x + \sum_{k=1}^m (-1)^k F^{(k)}(f(x)) \frac{(f(x))^k}{k!},$$

что соответствует замене функции F многочленом φ_m , производные которого совпадают с соответствующими производными F в точке $y = f(x)$. Итерационный метод вида $x_{n+1} = \varphi_m(x_n)$ имеет порядок сходимости $m + 1$.

19.1. Построить итерационный метод Ньютона для вычисления $\sqrt[p]{a}$, $a > 0$, где $p \neq 0$ — вещественное число.

19.2. Пусть уравнение $f(x) = 0$ имеет на отрезке $[a, b]$ простой корень z , причем $f(x)$ — трижды непрерывно дифференцируемая функция. Доказать квадратичную скорость сходимости метода Ньютона.

- 19.3.** Пусть уравнение $f(x) = 0$ имеет на отрезке $[a, b]$ корень z кратности p , причем $f(x)$ — дважды непрерывно дифференцируемая функция. Показать, что при этих условиях метод Ньютона сходится как геометрическая прогрессия (т. е. линейно) со знаменателем $(p - 1)/p$.
- 19.4.** Пусть уравнение $f(x) = 0$ имеет на отрезке $[a, b]$ корень z кратности p , причем $f(x)$ — дважды непрерывно дифференцируемая функция. Построить модификацию метода Ньютона, имеющую квадратичную скорость сходимости.
- 19.5.** Построить метод Ньютона для вычисления значения a^{-1} так, чтобы расчетные формулы не содержали операций деления. Определить область сходимости метода при $a > 0$.
- 19.6.** Пусть уравнение $f(x) = 0$ имеет на отрезке $[a, b]$ корень z неизвестной кратности $p > 1$, причем $f(x)$ — трижды непрерывно дифференцируемая функция. Построить модификацию метода Ньютона с квадратичной скоростью сходимости и предложить способ численной оценки величины кратности корня.
- 19.7.** Доказать, что если на отрезке $[a, b]$ функция $f'(x)$ не обращается в нуль, функция $f''(x)$ непрерывна и не меняет знака, а также выполнены условия

$$f(a)f(b) < 0, \quad \max \left[\left| \frac{f(a)}{f'(a)} \right|, \left| \frac{f(b)}{f'(b)} \right| \right] \leq b - a,$$

то метод Ньютона для уравнения $f(x) = 0$ сходится при любом $x_0 \in [a, b]$.

- 19.8.** Указать область сходимости метода решения уравнения $x = 1/a$, не содержащего операций деления: $x_{n+1} = (1 + C)x_n - aCx_n^2$, в зависимости от параметра $C \neq 0$.
- 19.9.** При вычислении методом Ньютона \sqrt{a} при $1 \leq a \leq 4$ начальное приближение x_0 выбрано равным значению многочлена наилучшего равномерного приближения для \sqrt{a} на $[1, 4]$, т. е. $x_0 = Q_1^0(a) = \frac{17}{24} + \frac{a}{3}$. Доказать справедливость оценки $|x_4 - \sqrt{a}| \leq 0,5 \cdot 10^{-25}$.

19.10. Для нахождения $a^{1/3}$ используют итерационный процесс

$$x_{n+1} = Ax_n + B \frac{a}{x_n^2} + C \frac{a^2}{x_n^5}.$$

Найти значения параметров A, B, C , обеспечивающие максимальный порядок сходимости.

19.11. Записать метод Чебышёва для функции $f(x) = x^p - a$.

19.12. Показать, что метод вычисления $a^{1/p}$

$$x_{n+1} = \varphi(x_n), \quad \varphi(x) = x \frac{(p-1)x^p + (p+1)a}{(p+1)x^p + (p-1)a},$$

имеет третий порядок сходимости.

19.13. Определить порядок сходимости метода

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} - \frac{f''(x_n)(f(x_n))^2}{2(f'(x_n))^3}.$$

19.14. Определить порядок сходимости модифицированного метода

$$\text{Ньютона } x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)}.$$

19.15. Определить порядок сходимости метода

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} - \frac{f(x_n - (f'(x_n))^{-1}f(x_n))}{f'(x_n)}.$$

19.16. Для нахождения простого нуля z функции $f(x) \in C^{(4)}$ используют итерационный метод $x_{n+1} = (y_{n+1} + v_{n+1})/2$, где

$$y_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad v_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)}, \quad g(x) = \frac{f(x)}{f'(x)}.$$

Доказать, что если метод сходится, то скорость сходимости — кубическая.

19.17. Для нахождения корня z уравнения $f(x) = 0$ используют итерационный метод

$$x_{n+1} = g(x_n), \quad g(x) = x - \frac{(f(x))^2}{f(x + f(x)) - f(x)}.$$

Исследовать поведение функции $g(x)$ в окрестности z .

19.18. Записать расчетную формулу метода Ньютона для системы уравнений

$$1) \begin{cases} \sin(x+y) - 1, 3x = 0, 1, \\ x^2 + y^2 = 1; \end{cases} \quad 2) \begin{cases} x^{10} + y^{10} = 1024, \\ e^x - e^y = 1. \end{cases}$$

19.19. Указать начальное приближение и оценить число итераций в методе Ньютона, требующихся для достижения точности 10^{-3} при решении системы уравнений

$$\begin{cases} x^3 - y^2 = 1, \\ xy^3 - y = 4. \end{cases}$$

19.20. Проверить, что $\mathbf{z} = (1, 1, 1)^T$ — одно из решений системы уравнений $\mathbf{F}(\mathbf{x}) = 0$, где $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ имеет вид

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} x_1 x_2^3 + x_2 x_3 - x_1^4 - 1 \\ x_2 + x_2^2 + x_3 - 3 \\ x_2 x_3 - 1 \end{bmatrix}.$$

Сходится ли метод Ньютона к \mathbf{z} при достаточно близких начальных приближениях?

19.21. Для приближенного решения нелинейной краевой задачи

$$y'' = f(x, y), \quad x \in [0, X], \quad y(0) = a, \quad y(X) = b,$$

рассматривают систему нелинейных алгебраических уравнений с параметром $h = X/N$:

$$\frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} = f(x_k, y_k), \quad k = 1, 2, \dots, N-1, \quad y_0 = a, \quad y_N = b.$$

Здесь y_k — приближения к значениям $y(kh)$. Записать расчетные формулы метода Ньютона для решения приведенной системы. Указать способ их реализации, если

- 1) $f(x, y) = x^2 + y^3$;
- 2) $f(x, y) = y^2 \exp(x)$;
- 3) $f(x, y) = \sin(y) \cos(x)$.

19.22. Для решения уравнения $x^3 - x = 0$ применяют метод Ньютона. При каком начальном приближении он сходится и к какому корню?

Глава VI

Разностные уравнения



Пусть неизвестная функция y и заданная функция f являются функциями одного целочисленного аргумента. Тогда линейное уравнение

$$a_0y(k) + a_1y(k+1) + \dots + a_ny(k+n) = f(k), \quad k = 0, 1, 2, \dots,$$

где $a_i, i = 0, 1, \dots, n$ — постоянные коэффициенты и $a_0 \neq 0, a_n \neq 0$, называют *линейным разностным уравнением n -го порядка с постоянными коэффициентами*. Если в этом уравнении положить $y(k+i) = y_{k+i}$ и $f(k) = f_k$, то оно принимает вид

$$a_0y_k + a_1y_{k+1} + \dots + a_ny_{k+n} = f_k, \quad k = 0, 1, 2, \dots$$

Для однозначного определения решения требуется задать n условий, например,

$$y_i = b_i, \quad i = 0, 1, \dots, n-1.$$

Как в постановках задач, так и в методах решения, имеется глубокая аналогия между рассмотренным разностным уравнением и обыкновенным дифференциальным уравнением с постоянными коэффициентами

$$\tilde{a}_0y(x) + \tilde{a}_1y'(x) + \dots + \tilde{a}_ny^{(n)}(x) = \tilde{f}(x).$$

§ 20. Однородные разностные уравнения

Если в разностном уравнении правая часть f_k равна нулю, то уравнение называют *однородным*. Напомним, как ищется общее решение однородного дифференциального уравнения с постоянными коэффициентами. Положим $y(x) = \exp(\lambda x)$. Подставляя это выражение в дифференциальное уравнение и сокращая затем на $\exp(\lambda x)$, получим характеристическое уравнение $\tilde{p}(\lambda) \equiv \sum_{j=0}^n \tilde{a}_j \lambda^j = 0$. Если $\lambda_1, \dots, \lambda_r$ — различные корни этого уравнения кратности $\sigma_1, \dots, \sigma_r$ соответственно, то общее решение можно записать в виде $y(x) = c_{11}e^{\lambda_1 x} + c_{12}xe^{\lambda_1 x} + \dots + c_{1\sigma_1}x^{\sigma_1-1}e^{\lambda_1 x} + \dots + c_{r1}e^{\lambda_r x} + \dots + c_{r\sigma_r}x^{\sigma_r-1}e^{\lambda_r x}$, где c_{ij} — произвольные постоянные.

Аналогично ищется решение разностного уравнения. Положим $y_k = \mu^k$. Подставляя это выражение в разностное уравнение и сокращая затем на μ^k , получим характеристическое уравнение

$$p(\mu) \equiv \sum_{j=0}^n a_j \mu^j = 0.$$

Пусть μ_1, \dots, μ_r — его различные корни, $\sigma_1, \dots, \sigma_r$ — их кратности. Тогда общее решение однородного разностного уравнения имеет вид $y_k = c_{11}\mu_1^k + c_{12}k\mu_1^k + \dots + c_{1\sigma_1}k^{\sigma_1-1}\mu_1^k + \dots + c_{r1}\mu_r^k + c_{r2}k\mu_r^k + \dots + c_{r\sigma_r}k^{\sigma_r-1}\mu_r^k$, где c_{ij} — произвольные постоянные. Таким образом, каждому корню μ кратности σ соответствует набор частных решений вида $\mu^k, k\mu^k, \dots, k^{\sigma-1}\mu^k$.

20.1. Найти общее решение уравнения $by_{k+1} - cy_k + ay_{k-1} = 0$.

20.2. Найти общее действительное решение уравнения $y_{k+1} - y_k + 2y_{k-1} = 0$.

20.3. Верно ли, что любое решение разностного уравнения $y_{k+1} - 5y_k + 6y_{k-1} = 0$ удовлетворяет уравнению

$$y_{k+1} - 9y_k + 27y_{k-1} - 23y_{k-2} - 24y_{k-3} + 36y_{k-4} = 0?$$

20.4. Пусть φ_k и z_k — частные решения уравнения

$$a_1y_{k+1} + a_0y_k + a_{-1}y_{k-1} = 0, \quad a_1, a_{-1} \neq 0.$$

Доказать, что определитель матрицы $A_k = \begin{pmatrix} \varphi_k & \varphi_{k+1} \\ z_k & z_{k+1} \end{pmatrix}$ либо равен нулю, либо отличен от нуля для всех k одновременно.

20.5. Найти решение разностной задачи

$$y_{k+4} + 2y_{k+3} + 3y_{k+2} + 2y_{k+1} + y_k = 0, \quad y_0 = y_1 = y_3 = 0, \quad y_2 = -1.$$

20.6. Показать, что для чисел Фибоначчи

$$f_{k+1} = f_k + f_{k-1}, \quad f_0 = 0, \quad f_1 = 1$$

справедливо равенство $f_k f_{k+2} - f_{k+1}^2 = (-1)^{k+1}$, $k = 0, 1, 2, \dots$.

20.7. Вычислить определитель Δ_k трехдиагональной $k \times k$ матрицы A с элементами $a_{ij} = \{b \text{ для } i = j; a \text{ для } i = j + 1; c \text{ для } i = j - 1; 0 \text{ для остальных индексов}\}$.

20.8. Используя разностное уравнение, записать формулу для вычисления интеграла

$$I_k(\alpha) = \frac{1}{\pi} \int_0^{\pi} \frac{\cos(kx) - \cos(k\alpha)}{\cos x - \cos \alpha} dx,$$

где α — параметр из отрезка $[0, \pi]$.

20.9. Предложить удобную форму записи решения уравнения

$$y_{k+1} - 2p y_k + y_{k-1} = 0, \quad k = 1, 2, \dots, \quad p > 0.$$

20.10. Показать, что если $-1 < \lambda < 1$, то любое решение разностного уравнения $y_{k+1} - 2\lambda y_k + y_{k-1} = 0$ ограничено при $k \rightarrow \infty$. Если λ — любое комплексное число, не принадлежащее интервалу действительной оси $-1 < \lambda < 1$, то среди решений этого разностного уравнения имеются неограниченные при $k \rightarrow \infty$.

20.11. Доказать, что любое решение разностного уравнения $y_{k+1} - 12y_{k-1} + 2y_{k-2} + 27y_{k-3} - 18y_{k-4} = 0$ однозначно представимо в виде суммы решений уравнений

$$y_{k+1} - 3y_{k-1} + 2y_{k-2} = 0 \quad \text{и} \quad y_{k+1} - 9y_{k-1} = 0.$$

20.12. Найти решение краевой задачи

$$y_{k+1} - y_k + y_{k-1} = 0, \quad 1 \leq k \leq N-1, \quad y_0 = 1, \quad y_N = 0.$$

Найти общее действительное решение разностного уравнения

$$\mathbf{20.13.} \quad 20y_{k-1} - 8y_k + y_{k+1} = 0. \quad \mathbf{20.14.} \quad 2y_{k-1} - 2y_k + y_{k+1} = 0.$$

$$\mathbf{20.15.} \quad 26y_{k-1} + 10y_k + y_{k+1} = 0. \quad \mathbf{20.16.} \quad 13y_{k-1} + 4y_k + y_{k+1} = 0.$$

Найти решение разностной задачи

$$\mathbf{20.17.} \quad y_{k+2} + 4y_{k+1} + 4y_k = 0, \quad y_0 = 1, \quad y_1 = 4.$$

$$\mathbf{20.18.} \quad y_{k+2} + 3y_{k+1} + 2y_k = 0, \quad y_0 = 2, \quad y_1 = 1.$$

$$\mathbf{20.19.} \quad y_{k+2} + y_k = 0, \quad y_0 = 2, \quad y_1 = 1.$$

$$\mathbf{20.20.} \quad y_{k+1} - 4y_k + y_{k-1} + 6y_{k-2} = 0, \quad y_0 = 6, \quad y_1 = 12, \quad y_4 = 276.$$

20.21. Пусть $I_k = \int_0^{\infty} x^k e^{-x} \sin x dx$, $k \geq 0$. Показать, что для целых неотрицательных k справедливо равенство $I_{4k+3} = 0$.

20.22. Для целых положительных чисел $a_0 > a_1$ наибольший общий делитель находится последовательным делением a_0 на a_1 , затем a_1 — на первый остаток и т. д. Указать оценку сверху для числа делений (длину алгоритма Евклида).

§ 21. Неоднородные разностные уравнения

Пусть y_k^0 — общее решение однородного, y_k^1 — частное решение неоднородного уравнения. Тогда общее решение линейного неоднородного уравнения с постоянными коэффициентами можно представить в виде их суммы $y_k = y_k^0 + y_k^1$.

Если правая часть имеет специальный вид, то частное решение можно найти методом неопределенных коэффициентов. Пусть

$$f_k = \alpha^k (P_m(k) \cos \beta k + Q_n(k) \sin \beta k),$$

где $P_m(k)$, $Q_n(k)$ — многочлены степени m и n соответственно. Тогда частное решение ищут в виде

$$y_k^1 = k^s \alpha^k (R_l(k) \cos \beta k + T_l(k) \sin \beta k), \quad (1)$$

где $s = 0$, если $\alpha e^{\pm i\beta}$ не являются корнями характеристического уравнения, и s равно кратности корня в противном случае; $l = \max(m, n)$ — степень многочленов $R_l(k)$ и $T_l(k)$. Чтобы найти коэффициенты этих многочленов, надо подставить выражение (1) в неоднородное уравнение и приравнять коэффициенты при подобных членах.

Найти частное решение разностного уравнения

21.1. $2y_k - y_{k+1} = 1 + 2k - k^2$. **21.2.** $2y_k - y_{k+1} = k 2^k$.

21.3. $2y_k - y_{k+1} = \sin k$.

Найти решение разностной задачи

21.4. $y_{k+1} - b y_k = a^k$, $y_0 = 1$ ($a, b \neq 0$).

21.5. $y_{k+1} - y_{k-1} = \frac{1}{k^2 - 1}$, $y_1 = 0$, $y_2 = 0$.

21.6. $y_{k+1} - k y_k = 2^k k!$, $k \geq 0$.

21.7. $y_{k+1} = \frac{y_k}{1 + y_k}$, $y_0 = 1$.

21.8. $y_{k+1} = \frac{a y_k + b}{c y_k + d}$, $y_0 = 1$, при условии $(a - d)^2 + 4bc > 0$.

21.9. Найти частное решение уравнения

$$y_{k+2} - y_{k+1} - 6y_k = 4 \cdot 2^k.$$

Найти решение разностной задачи

21.10. $y_{k+2} - 4y_k = 5 \cdot 3^k$, $y_0 = 0$, $y_1 = 1$.

21.11. $y_{k+2} - 4y_k = 2^k$, $y_0 = 0$, $y_1 = 1$.

21.12. $y_{k+2} - 4y_{k+1} + 4y_k = 2^k$, $y_0 = 0$, $y_1 = 1$.

21.13. $y_{k+4} - \frac{5}{2}y_{k+3} + \frac{5}{2}y_{k+1} - y_k = 1$,
 $y_0 = 0$, $y_1 = 11$, $y_2 = -8$, $y_3 = 6$.

21.14. Вычислить сумму

$$S_k = \sum_{n=0}^k a_n, \quad a_n = (1 + n + n^2) \cos \beta n.$$

Найти решение нелинейной разностной задачи

21.15. $y_{k+1} = 1/(2 - y_k)$, $y_0 = 1/2$.

21.16. $y_{k+1} = 2 - 1/y_k$, $y_0 = 2$.

Найти решение нелинейного уравнения

21.17. $y_{k+1}^2 - y_k^2 = 1$, $k \geq 0$.

21.18. $y_{k+1}^2 = 2y_k$.

Найти решение нелинейной разностной задачи

21.19. $y_k y_{k+2}^3 = y_{k+1}^3 y_{k+3}$, $y_0 = 1$, $y_1 = e^{-1/2}$, $y_2 = e^{-2}$.

21.20. $y_{k+1} = (a y_k + b)/(c y_k + d)$, $y_0 = 1$, при условии $(a - d)^2 + 4bc = 0$.

Найти частное решение уравнения

21.21. $\frac{1}{8}y_{k-1} - \frac{3}{4}y_k + y_{k+1} = \left(\frac{1}{2}\right)^k$. **21.22.** $y_{k+1} - y_k - 12y_{k-1} = 4^k$.

21.23. $3y_{k+1} + 17y_k - 6y_{k-1} = \left(\frac{1}{3}\right)^k$. **21.24.** $y_{k+1} - 5y_k + 6y_{k-1} = 2^k$.

Найти общее решение уравнения

$$21.25. y_{k+1} - \frac{5}{2} y_k + y_{k-1} = \cos k. \quad 21.26. y_{k+2} - 2y_{k+1} - 3y_k + 4y_{k-1} = k.$$

$$21.27. y_{k+1} + y_k - 5y_{k-1} + 3y_{k-2} = 1. \quad 21.28. y_{k+1} - 2y_k - 8y_{k-1} = \sin k.$$

§ 22. Фундаментальное решение и задачи на собственные значения

Фундаментальным решением G_k называют решение разностного уравнения

$$a_0 y_k + a_1 y_{k+1} + \dots + a_n y_{k+n} = f_k$$

с правой частью специального вида $f_k = \delta_k^0$, где

$$\delta_k^n = \begin{cases} 0 & \text{при } k \neq n, \\ 1 & \text{при } k = n. \end{cases}$$

22.1. Найти ограниченное фундаментальное решение уравнения

$$a y_k + b y_{k+1} = \delta_k^0.$$

22.2. Пусть $|a/b| \neq 1$, $|f_k| \leq F$, а G_k — ограниченное фундаментальное решение уравнения $a y_k + b y_{k+1} = f_k$. Показать, что частным решением этого уравнения является сходящийся ряд

$$y_k^1 = \sum_{n=-\infty}^{\infty} G_{k-n} f_n.$$

Найти ограниченное фундаментальное решение уравнения

$$22.3. y_{k-1} - 2y_k + y_{k+1} = \delta_k^0. \quad 22.4. y_{k-1} - y_k + y_{k+1} = \delta_k^0.$$

$$22.5. y_{k-1} - \frac{5}{2} y_k + y_{k+1} = \delta_k^0. \quad 22.6. y_{k+1} - 5y_k + 6y_{k-1} = \delta_k^0.$$

$$22.7. 3y_{k-1} - \frac{13}{2} y_k + y_{k+1} = \delta_k^0. \quad 22.8. \frac{1}{8} y_{k-1} - \frac{3}{4} y_k + y_{k+1} = \delta_k^0.$$

$$22.9. y_{k+1} - y_k - 12y_{k-1} = \delta_k^0.$$

22.10. Доказать, что решение разностной задачи

$$y_{k-1} - 2y_k + y_{k+1} = f_k, \quad y_0 = \alpha, \quad y_N = \beta$$

удовлетворяет неравенству

$$\max_{0 \leq k \leq N} |y_k| \leq \max(|\alpha|, |\beta|) + \max_{1 \leq k \leq N-1} |f_k| \frac{N^2}{8}.$$

Найти все решения задачи на собственные значения

$$22.11. \quad \frac{y_{k+1} - y_{k-1}}{2h} = -\lambda y_k, \quad 1 \leq k \leq N-1, \quad h = \frac{1}{N}, \quad y_0 = y_N = 0.$$

$$22.12. \quad \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} = -\lambda y_k, \quad 1 \leq k \leq N-1, \quad h = \frac{1}{N}, \quad y_0 = y_N = 0.$$

$$22.13. \quad \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} = -\lambda y_k, \quad 1 \leq k \leq N-1, \quad h = \frac{1}{N}, \\ \frac{2}{h^2}(y_1 - y_0) = -\lambda y_0, \quad -\frac{2}{h^2}(y_N - y_{N-1}) = -\lambda y_N.$$

$$22.14. \quad \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} = -\lambda y_k, \quad 1 \leq k \leq N-1, \quad h = \frac{1}{N}, \\ y_0 = 0, \quad -\frac{2}{h^2}(y_N - y_{N-1}) = -\lambda y_N.$$

$$22.15. \quad \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} = -\lambda y_k, \quad 1 \leq k \leq N-1, \quad h = \frac{1}{N}, \\ \frac{2}{h^2}(y_1 - y_0) = -\lambda y_0, \quad y_N = 0.$$

$$22.16. \quad \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} = -\lambda y_k, \quad k = 0, \pm 1, \pm 2, \dots, \quad h = \frac{1}{N}, \\ y_k = y_{k+N}.$$

$$22.17. \quad \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} = -\lambda y_k, \quad 1 \leq k \leq N-1, \quad h = \frac{1}{N-1}, \\ y_0 = y_1, \quad y_N = y_{N-1}.$$

$$22.18. \quad \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} = -\lambda y_k, \quad 1 \leq k \leq N-1, \quad h = \frac{1}{N-1}, \\ y_0 = -y_1, \quad y_N = -y_{N-1}.$$

$$22.19. \quad \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} = -\lambda y_k, \quad 1 \leq k \leq N-1, \quad h = \frac{1}{N}, \\ y_0 = y_1, \quad y_N = 0.$$

$$22.20. \quad \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} = -\lambda y_k, \quad 1 \leq k \leq N-1, \quad h = \frac{1}{N}, \\ y_0 = 0, \quad y_N = y_{N-1}.$$

$$22.21. \quad \frac{y_{k+1} - y_{k-1}}{2h} = -\lambda y_k, \quad 1 \leq k \leq N-1, \quad h = \frac{1}{N}, \\ y_0 = y_1, \quad y_N = y_{N-1}.$$

Решение дифференциальных уравнений



Постановки задач. Пусть в области D с границей Γ задана дифференциальная задача

$$Lu = f \quad \text{в} \quad D \tag{1}$$

с граничным условием

$$lu = \varphi \quad \text{на} \quad \Gamma. \tag{2}$$

Здесь L и l — дифференциальные операторы; f и φ — заданные, а u — искомый элемент некоторого линейного нормированного функционального пространства F , Φ и U , соответственно.

Если одной из переменных является время t , то наиболее часто рассматривают области вида $D(t, \mathbf{x}) = d(\mathbf{x}) \times [t_0, T]$, где t — время, $\mathbf{x} = (x_1, \dots, x_n)$ — совокупность пространственных координат. Это означает, что решение ищется в пространственной области $d(\mathbf{x})$ на отрезке времени $[t_0, T]$. В этом случае условия, заданные при $t = t_0$, называют *начальными*, а условия, заданные на границе $\gamma(\mathbf{x})$ области $d(\mathbf{x})$, — *граничными*, или *краевыми*.

Задачу только с начальными условиями называют *задачей Коши*. Задачу с начальными и граничными условиями называют *смешанной краевой задачей*.

Метод сеток. Для приближенного решения дифференциальных задач часто используют метод сеток. Для его применения определяют некоторую *сетку* — конечное множество точек (узлов) $\bar{D}_h = D_h \cup \Gamma_h$, принадлежащее области $\bar{D} = D \cup \Gamma$. Как правило, $\Gamma_h \subset \Gamma$. Будем рассматривать только сетки, узлами которых являются все точки пересечения заданных наборов параллельных прямых (плоскостей), причем по каждой переменной выбирается свой, как правило, постоянный *шаг*, т. е. расстояние между соседними узлами. Сетки по времени и пространству обычно определяют независимо. Сетка \bar{D}_h характеризуется параметром h , в общем случае векторным, компоненты которого состоят из шагов сетки h_i по каждой переменной. Для изучения свойств сеточных задач вводится понятие

величины шага сетки, в качестве которой принимается какая-либо норма вектора h , например, $\|h\|_\infty = \max_{1 \leq i \leq n} h_i$, где n — число независимых переменных в дифференциальной задаче. В приводимых ниже оценках под h понимается величина шага сетки.

Если $X \subset Y$ и функция v определена на множестве Y , то ее следом на множестве X называется функция, определенная на X и совпадающая там с v . Если функция v определена на некотором множестве Y , содержащем Y_h , то ее след на Y_h будем обозначать $(v)_h$. Часто пространства F_h , Φ_h и U_h определяют как пространства следов функций из F , Φ и U (или следов гладких функций из всюду плотных в этих пространствах подмножеств) на D_h , Γ_h и \overline{D}_h соответственно. При этом задают согласованные нормы пространств. Так, норма пространства U_h согласована с нормой U , если для достаточно гладких функций $u \in U$ выполняется соотношение

$$\lim_{h \rightarrow 0} \|(u)_h\|_{U_h} = \|u\|_U.$$

Все производные, входящие в уравнение и краевые условия, заменяют конечными разностями, или другими словами, *разностными аппроксимациями*. При записи этих аппроксимаций в некотором внутреннем узле сетки берут одно и то же количество соседних узлов, образующих строго определенную конфигурацию, называемую *шаблоном*. В результате дифференциальные операторы L и l заменяют разностными L_h и l_h .

Определим *разностную схему* как семейство сеточных задач, зависящих от параметра h :

$$L_h u_h = f_h \quad \text{в } D_h, \quad (3)$$

$$l_h u_h = \varphi_h \quad \text{на } \Gamma_h. \quad (4)$$

Решение разностной схемы u_h , называемое *разностным*, принимается в качестве приближенного решения дифференциальной задачи (1), (2).

Аппроксимация. Оператор L_h из (3) локально аппроксимирует в точке x_i сетки D_h дифференциальный оператор L из (1), если для достаточно гладкой функции $u \in U$ существуют такие положительные постоянные h_0 , c и p , не зависящие от h , что при всех $h \leq h_0$ справедливо неравенство

$$|(L_h(u)_h - (Lu)_h)|_{x=x_i} \leq ch^p.$$

Число p при этом называют *порядком аппроксимации*. Аналогично определяют порядок локальной аппроксимации для оператора l_h .

Говорят, что разностная схема (3), (4) *аппроксимирует* дифференциальную задачу (1), (2) с порядком аппроксимации $p = \min(p_1, p_2)$, если для любых достаточно гладких функций $u \in U$, $f \in F$ и $\varphi \in \Phi$ существуют такие положительные постоянные h_0 , c_1 , p_1 , c_2 и p_2 , не зависящие от h , что при всех $h \leq h_0$ справедливы неравенства

$$\begin{aligned} \|L_h(u)_h - (Lu)_h\|_{F_h} + \|(f)_h - f_h\|_{F_h} &\leq c_1 h^{p_1}, \\ \|l_h(u)_h - (lu)_h\|_{\Phi_h} + \|(\varphi)_h - \varphi_h\|_{\Phi_h} &\leq c_2 h^{p_2}. \end{aligned}$$

Выражения, стоящие под знаками норм, называют *погрешностями аппроксимации*.

Также используется понятие аппроксимации на решении, позволяющее строить схемы более высокого порядка точности на фиксированном шаблоне. Говорят, что разностная схема (3), (4) *аппроксимирует на решении* u с порядком аппроксимации $p = \min(p_1, p_2)$ дифференциальную задачу (1), (2), если существуют такие положительные постоянные h_0 , c_1 , p_1 , c_2 и p_2 , не зависящие от h , что при всех $h \leq h_0$ справедливы неравенства

$$\|L_h(u)_h - f_h\|_{F_h} \leq c_1 h^{p_1}, \quad \|l_h(u)_h - \varphi_h\|_{\Phi_h} \leq c_2 h^{p_2}.$$

Предполагается, что при этом выполнены условия нормировки сеточных функций f_h и φ_h : $\lim_{h \rightarrow 0} \|f_h\|_{F_h} = \|f\|_F$, $\lim_{h \rightarrow 0} \|\varphi_h\|_{\Phi_h} = \|\varphi\|_\Phi$.

Порядки аппроксимаций обычно оценивают с помощью разложения в ряды Тейлора. Порядок аппроксимации разностной схемы может быть разным по разным переменным. Если погрешность аппроксимации стремится к нулю при любом законе стремления шагов по различным переменным к нулю, то такую аппроксимацию называют *безусловной*. Если же погрешность аппроксимации стремится к нулю при одних законах убывания шагов и не стремится к нулю при других, то аппроксимацию называют *условной*.

Устойчивость. Разностную схему (3), (4) называют *устойчивой*, если ее решение существует, единственно и непрерывно зависит от заданных f_h , φ_h , причем эта зависимость равномерна относительно величины шага сетки h . Уточним это определение, ограничившись

для простоты только линейным случаем. Рассмотрим решения $u_h^{(i)}$ разностной схемы (3), (4) при соответствующих правых частях $f_h^{(i)}, \varphi_h^{(i)}, i = 1, 2$. Линейная разностная схема устойчива, если существуют такие положительные постоянные h_0, c_1 и c_2 , не зависящие от h , что для всех $h \leq h_0$ справедливо неравенство

$$\|u_h^{(1)} - u_h^{(2)}\|_{U_h} \leq c_1 \|f_h^{(1)} - f_h^{(2)}\|_{F_h} + c_2 \|\varphi_h^{(1)} - \varphi_h^{(2)}\|_{\Phi_h}.$$

Это означает, что для решения (3), (4) имеет место *априорная оценка*

$$\|u_h\|_{U_h} \leq c_1 \|f_h\|_{F_h} + c_2 \|\varphi_h\|_{\Phi_h},$$

из которой необходимо следует существование и единственность решения при любой правой части.

Устойчивость называют *безусловной*, если указанные неравенства выполняются при произвольном соотношении шагов по различным переменным. Если же для выполнения неравенств шаги должны удовлетворять дополнительным соотношениям, то устойчивость называют *условной*.

Непрерывную зависимость от f_h (равномерную относительно h) называют устойчивостью *по правой части*, а непрерывную зависимость от φ_h — устойчивостью *по граничным условиям*. Если рассматривается смешанная краевая задача, то устойчивость по граничному условию при $t = t_0$ называют устойчивостью *по начальным данным*.

Сходимость. Решение u_h разностной схемы (3), (4) *сходится* к решению u дифференциальной задачи (1), (2), если

$$\|(u)_h - u_h\|_{U_h} \rightarrow 0 \text{ при } h \rightarrow 0.$$

Если существуют такие положительные постоянные h_0, c и p , не зависящие от h , что для всех $h \leq h_0$ справедливо неравенство

$$\|(u)_h - u_h\|_{U_h} \leq c h^p,$$

то разностная схема сходится *с порядком p* ; говорят, также, что разностное решение u_h имеет порядок точности p .

Теорема (о связи аппроксимации, устойчивости и сходимости). Пусть выполнены следующие условия:

- 1) операторы L, l и L_h, l_h — линейные;
- 2) решение u дифференциальной задачи (1), (2) существует и единственно;

3) разностная схема (3), (4) аппроксимирует дифференциальную задачу (1), (2) с порядком p ;

4) разностная схема (3), (4) устойчива.

Тогда решение разностной схемы u_h сходится к решению и дифференциальной задачи с порядком не ниже p .

Для многомерных задач порядок аппроксимации по разным переменным может быть неодинаковым, поэтому порядок сходимости по разным переменным может быть различным. Если аппроксимация и (или) устойчивость разностной схемы условные, то сходимость имеет место только при тех соотношениях между шагами сетки по разным переменным, при которых выполнены условия аппроксимации и (или) устойчивости. В классе задач с решениями конечной гладкости требование устойчивости является необходимым условием сходимости.

§ 23. Методы построения разностных схем

Рассмотрим наиболее распространенные методы построения разностных схем. При этом будем считать, что входные данные и решение дифференциальной задачи обладают требуемой гладкостью.

Метод неопределенных коэффициентов. Пусть задано несколько расположенных группой узлов сетки (шаблон) и требуется с использованием этого шаблона построить разностный оператор L_h , локально аппроксимирующий дифференциальный оператор L в одном из узлов сетки x_i . В этом случае в выражении

$$(L_h(u)_h - (Lu)_h)|_{x=x_i}$$

оператор L_h берут с неопределенными коэффициентами. Для нахождения искомых коэффициентов с помощью формулы Тейлора строят разложения в точке x_i для всех функций $u(x_j)$, входящих в выражение $L_h(u)_h$, и группируют множители при $u(x_i), u'(x_i), u''(x_i), \dots$. Приравнявая найденные множители к нулю, приходят к системе линейных алгебраических уравнений для нахождения коэффициентов разностной схемы. Порядок аппроксимации и главный член погрешности определяется после подстановки найденных коэффициентов в первый ненулевой множитель при соответствующей производной функции $u(x)$ в точке x_i .

Интегро-интерполяционный метод. В качестве примера опишем применение этого метода к построению разностной схемы на равномерной сетке $\bar{D}_h = \{x_i = ih, i = 0, \dots, N; Nh = 1\}$ для задачи

$$Lu \equiv -u'' + p(x)u = f(x), \quad x \in [0, 1], \quad 0 \leq p(x) \leq p_1, \\ u'(0) = \alpha_1 u(0) + \beta_1, \quad u(1) = 0.$$

Введя обозначение $\omega(x) = u'(x)$, перепишем исходное уравнение в виде $\omega'(x) = p(x)u(x) - f(x)$ и затем проинтегрируем в пределах от $x_{i-1/2}$ до $x_{i+1/2}$ ($x_{i\pm 1/2} = x_i \pm h/2$):

$$\omega(x_{i+1/2}) - \omega(x_{i-1/2}) = \int_{x_{i-1/2}}^{x_{i+1/2}} [p(x)u(x) - f(x)] dx.$$

Полученное равенство служит основой для построения разностных схем. Заменим интеграл в правой части, например, по квадратурной формуле прямоугольников

$$\int_a^b \varphi(x) dx = (b-a) \varphi\left(\frac{a+b}{2}\right) + O((b-a)^3)$$

и разделим обе части на h . В результате получим

$$\frac{\omega(x_{i+1/2}) - \omega(x_{i-1/2})}{h} = p(x_i)u(x_i) - f(x_i) + O(h^2),$$

и осталось аппроксимировать соотношение $u'(x) = \omega(x)$ на сетке D_h :

$$u(x_{i+1}) - u(x_i) = h\omega(x_{i+1/2}) + O(h^3), \\ u(x_i) - u(x_{i-1}) = h\omega(x_{i-1/2}) + O(h^3).$$

Окончательный вид дискретного аналога исходного уравнения

$$-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + p_i u_i = f_i, \quad 0 < i < N,$$

где u_i обозначает приближение к точному решению $u(x_i)$, в то время как $p_i = p(x_i)$ и $f_i = f(x_i)$ — значения известных функций в узлах сетки.

Для аппроксимации краевого условия третьего рода проинтегрируем исходное уравнение от 0 до $h/2$:

$$u'(h/2) - u'(0) = \int_0^{h/2} [p(x)u(x) - f(x)] dx.$$

Далее опять воспользуемся формулой прямоугольников для интеграла и заменим $u'(0)$ на $\alpha_1 u(0) + \beta_1$, а $u'(h/2)$ на $(u(h) - u(0))/h + O(h^2)$. В результате получим

$$\frac{u(h) - u(0)}{h} - \alpha_1 u_0 - \beta_1 + O(h^2) = \frac{h}{2} [p(h/4) u(h/4) - f(h/4)] + O(h^3).$$

Левая часть равенства содержит слагаемое $O(h^2)$, поэтому в его правой части значения функций в точке $x = h/4$ можно заменить их значениями в точке $x = 0$, сохранив тот же порядок аппроксимации $O(h^2)$. В результате получим

$$\frac{u_1 - u_0}{h} = \alpha_1 u_0 + \beta_1 + \frac{h}{2} (p_0 u_0 - f_0).$$

Окончательный вид разностной схемы

$$-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + p_i u_i = f_i, \quad 0 < i < N, \quad \frac{u_1 - u_0}{h} = \bar{\alpha}_1 u_0 + \bar{\beta}_1, \quad u_N = 0,$$

где $\bar{\alpha}_1 = \alpha_1 + h p_0/2$, $\bar{\beta}_1 = \beta_1 - h f_0/2$.

Интегральное тождество Марчука. Рассмотрим задачу

$$L u \equiv -(k(x) u')' + p(x) u = f(x), \quad x \in [0, 1], \quad u(0) = u(1) = 0,$$

у которой переменные коэффициенты удовлетворяют условиям

$$0 < k_0 \leq k(x) \leq k_1, \quad 0 \leq p(x) \leq p_1,$$

и $k(x), p(x), f(x)$ могут иметь конечное число разрывов первого рода. Построение разностной схемы для нее основано на аппроксимации интегрального тождества, которому удовлетворяет решение $u(x)$:

$$\begin{aligned} & -\frac{u(x_{i+1}) - u(x_i)}{\int_{x_i}^{x_{i+1}} \frac{dx}{k(x)}} + \frac{u(x_i) - u(x_{i-1}))}{\int_{x_{i-1}}^{x_i} \frac{dx}{k(x)}} + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} [p(x)u - f(x)] dx = \\ & = -\frac{1}{\int_{x_i}^{x_{i+1}} \frac{dx}{k(x)}} \int_{x_i}^{x_{i+1}} \frac{dx}{k(x)} \int_{x_{i+\frac{1}{2}}}^x [p(\xi)u(\xi) - f(\xi)] d\xi + \\ & + \frac{1}{\int_{x_{i-1}}^{x_i} \frac{dx}{k(x)}} \int_{x_{i-1}}^{x_i} \frac{dx}{k(x)} \int_{x_{i-\frac{1}{2}}}^x [p(\xi)u(\xi) - f(\xi)] d\xi. \end{aligned}$$

Метод Ритца. Пусть требуется найти решение задачи $Lu = f$, где L — самосопряженный и положительно определенный оператор в гильбертовом пространстве U со скалярным произведением (\cdot, \cdot) , т. е.

$$(Lu, v) = (u, Lv) \text{ и } (Lv, v) \geq \delta(v, v), \delta > 0.$$

Будем искать элемент $u \in U$, минимизирующий функционал

$$J(v) = (Lv, v) - 2(f, v).$$

Чтобы определить приближения к u , построим последовательность конечномерных подпространств U_h пространства U с известными базисами $\{\varphi_j^h\}$ и в каждом U_h найдем элемент u_h , минимизирующий $J(v)$. Для этого достаточно найти коэффициенты α_j разложения u_h по базису $\{\varphi_j^h\}$

$$u_h = \sum_{j=0}^N \alpha_j \varphi_j^h$$

из системы линейных алгебраических уравнений $A\alpha = \mathbf{b}$, где

$$a_{ij} = (L\varphi_j^h, \varphi_i^h), \quad b_i = (f, \varphi_i^h), \quad i, j = 0, 1, 2, \dots, N,$$

$N + 1$ — размерность U_h . Если последовательность U_h полна в U (т. е. $\forall v \in U$ существует последовательность $\{v_h\} \in U_h$ такая, что $\|v_h - v\|_U \rightarrow 0$ при $h \rightarrow 0$), то $\lim_{h \rightarrow 0} \|u_h - u\|_U = 0$.

При аппроксимации краевых задач для уравнений второго порядка в качестве базисных функций φ_j^h можно использовать кусочно-линейные функции. Для произвольной сетки $x_0 < x_1 < \dots < x_N$ они имеют вид

$$\varphi_0^h(x) = \begin{cases} \frac{x_1 - x}{x_1 - x_0} & \text{при } x_0 \leq x \leq x_1, \\ 0 & \text{при } x_1 \leq x \leq x_N; \end{cases}$$

$$\varphi_N^h(x) = \begin{cases} 0 & \text{при } x_0 \leq x \leq x_{N-1}, \\ \frac{x - x_{N-1}}{x_N - x_{N-1}} & \text{при } x_{N-1} \leq x \leq x_N; \end{cases}$$

$$\varphi_j^h(x) = \begin{cases} \frac{x - x_{j-1}}{x_j - x_{j-1}} & \text{при } x_{j-1} \leq x \leq x_j, \\ \frac{x_{j+1} - x}{x_{j+1} - x_j} & \text{при } x_j \leq x \leq x_{j+1}, \\ 0 & \text{при остальных } x, \end{cases}$$

для $j = 1, \dots, N - 1$.

В общем случае, если меры носителей базисных функций много меньше меры области, в которой решается задача, то метод Рунца часто называют *методом конечных элементов*.

Метод Галёркина. В отличие от метода Рунца, метод Галёркина не требует самосопряженности и положительной определенности оператора L из задачи $Lu = f$.

Для нахождения приближений к u в каждом из конечномерных подпространств U_h отыскивают элемент u_h такой, что для любого $v \in U_h$ справедливо равенство $(Lu_h - f, v) = 0$. Соответствующие коэффициенты α_j разложения u_h по базису подпространства U_h определяют из системы уравнений, имеющей тот же вид, что и в методе Рунца.

В качестве иллюстрации построим методом Галёркина разностную схему для дифференциальной несамосопряженной задачи

$$Lu \equiv -u'' + a(x)u' = f(x), \quad x \in [0, 1], \quad u(0) = u(1) = 0,$$

с коэффициентом $a(x) = 3x^2$, используя в качестве базисных кусочно-линейные функции $\varphi_j^h(x)$ на равномерной сетке $\bar{D}_h = \{x_j = jh, j = 0, \dots, N; Nh = 1\}$. Будем искать приближенное решение в виде

$$u_h = \sum_{j=1}^{N-1} \alpha_j \varphi_j^h,$$

неизвестные коэффициенты α_j определим из системы линейных алгебраических уравнений $A\alpha = \mathbf{b}$, в которой

$$a_{ij} = (L\varphi_j^h, \varphi_i^h), \quad b_i = (f, \varphi_i^h), \quad i, j = 1, 2, \dots, N-1.$$

Вычислим элементы матрицы

$$a_{ij} = (L\varphi_j^h, \varphi_i^h) = \int_0^1 [(\varphi_j^h)'(\varphi_i^h)' + a(x)(\varphi_j^h)'\varphi_i^h] dx.$$

Первое слагаемое в этой формуле имеет вид

$$\int_{x_{i-1}}^{x_{i+1}} (\varphi_j^h)'(\varphi_i^h)' dx = \begin{cases} -\frac{1}{h} & \text{при } j = i-1, \\ \frac{2}{h} & \text{при } j = i, \\ -\frac{1}{h} & \text{при } j = i+1, \\ 0 & \text{при } |j-i| > 1. \end{cases}$$

Для второго слагаемого в результате несложных вычислений получим

$$\int_{x_{i-1}}^{x_{i+1}} 3x^2(\varphi_j^h)' \varphi_i^h dx = \begin{cases} -\frac{1}{4} [2x_{i-1}^2 + (x_{i-1} + x_i)^2] & \text{при } j = i - 1, \\ hx_i & \text{при } j = i, \\ \frac{1}{4} [2x_{i+1}^2 + (x_{i+1} + x_i)^2] & \text{при } j = i + 1, \\ 0 & \text{при } |j - i| > 1. \end{cases}$$

Для компонент вектора правой части имеем

$$b_i = (f, \varphi_i^h) = \int_0^1 f \varphi_i^h dx = \int_{x_{i-1}}^{x_{i+1}} f \varphi_i^h dx \approx f(x_i) \int_{x_{i-1}}^{x_{i+1}} \varphi_i^h dx = hf_i.$$

После деления обеих частей уравнения на h запишем окончательный результат

$$-\frac{\alpha_{i+1} - 2\alpha_i + \alpha_{i-1}}{h^2} + \frac{d_i\alpha_{i+1} + e_i\alpha_i + c_i\alpha_{i-1}}{h} = f_i, \\ 0 < i < N, \alpha_0 = \alpha_N = 0,$$

где коэффициенты определены формулами:

$$d_i = [2x_{i+1}^2 + (x_{i+1} + x_i)^2] / 4, \quad e_i = hx_i, \\ c_i = -[2x_{i-1}^2 + (x_{i-1} + x_i)^2] / 4.$$

23.1. Справедливы ли следующие равенства:

$$1) \lim_{h \rightarrow 0} \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} = \\ = \lim_{h \rightarrow 0} \frac{\frac{u(x+2h) + u(x)}{2} - 2u(x) + \frac{u(x) + u(x-2h)}{2}}{h^2};$$

$$2) \lim_{h \rightarrow 0} \frac{u(x+h) - u(x-h)}{2h} = \\ = \lim_{h \rightarrow 0} \frac{\frac{u(x+2h) + u(x)}{2} - \frac{u(x) + u(x-2h)}{2}}{2h},$$

если $u(x) \in C^{(4)}$?

23.2. Рассмотрим дифференциальную задачу

$$u' + a(x)u(x) = f(x), \quad x \in [0, 1], \quad u(0) = 0 \quad \text{при } a(x), f(x) \in C^{(4)}[0, 1].$$

Считая, что функции u_i и f_i определены в узлах $x_i = ih$, $h = 1/N$, $i = 0, \dots, N$, определить порядок аппроксимации на решении разностной схемы:

$$1) \frac{u_{i+1} - u_i}{h} + a_i u_i = f_i, \quad 0 \leq i \leq N-1, \quad u_0 = 0;$$

$$2) \frac{u_{i+1} - u_{i-1}}{2h} + a_i u_i = f_i, \quad 1 \leq i \leq N-1, \quad u_0 = 0, \quad u_1 = hf_0,$$

где $a_i = a(x_i)$, $f_i = f(x_i)$.

23.3. Рассмотрим дифференциальную задачу

$$u' + a(x)u(x) = f(x), \quad x \in [0, 1], \quad u(0) = 0 \quad \text{при} \quad a(x), f(x) \in C^{(4)}[0, 1].$$

Считая, что функция u_i определена в узлах $x_i = ih$, $h = 1/N$, $i = 0, \dots, N$, а функция f_i — в узлах $x_{i+1/2} = (i + 1/2)h$, $i = 0, \dots, N-1$, определить порядок аппроксимации и порядок аппроксимации на решении разностной схемы

$$\frac{u_{i+1} - u_i}{h} + a_i \frac{u_{i+1} + u_i}{2} = f_i, \quad u_0 = 0, \quad i = 0, \dots, N-1,$$

где $a_i = a(x_{i+1/2})$, $f_i = f(x_{i+1/2})$.

23.4. Привести пример последовательности сеточных функций $\{\varphi_i^h\}$, $i = 0, 1, \dots, N$, $Nh = 1$, из семейства пространств $\{U_h\}$, которая сходилась бы при $h \rightarrow 0$ к некоторой функции

$$u \in U, \quad \text{если} \quad \|\varphi^h\| = \left(h \sum_{i=0}^N (\varphi_i^h)^2 \right)^{1/2}, \quad \text{и расходилась, если}$$

$$\|\varphi^h\| = \max_{0 \leq i \leq N} |\varphi_i^h|.$$

23.5. Сходится ли последовательность сеточных функций $\{\varphi_i^h\}$, $i = 0, 1, \dots, N$, $Nh = 1$, в норме $\|\varphi^h\| = \max_{0 \leq i \leq N} |\varphi_i^h|$ к функции $u(x)$ и с каким порядком, если

$$\varphi_i^h = \frac{1}{2} \left(u \left(x_i + \frac{h}{2} \right) + u \left(x_i - \frac{h}{2} \right) \right), \quad x_i = ih, \quad \varphi_0^h = u(0), \quad \varphi_N^h = u(1),$$

а $u(x)$ принадлежит одному из пространств $C^{(k)}$, $k \geq 0$?
Существуют ли функции $u(x)$, к которым $\{\varphi_i^h\}$ сходится с бесконечным порядком?

23.6. Для дифференциальной задачи

$$u'' = f(x), \quad x \in [0, 1], \quad u(0) = u(1) = 0$$

построить разностную схему методом неопределенных коэффициентов, используя трехточечный шаблон с постоянным шагом.

23.7. Для дифференциальной задачи

$$-u'' = f(x), \quad x \in [0, 1], \quad u(0) = a, \quad u(1) = b$$

на трехточечном шаблоне с переменным шагом построить разностные схемы первого и второго порядка аппроксимации на решении $u \in C^{(4)}[0, 1]$.

23.8. Для дифференциальной задачи

$$u' + cu = f(x), \quad x \in [0, 1], \quad u(0) = a, \quad c = \text{const}$$

построить интегро — интерполяционным методом разностную схему четвертого порядка аппроксимации на трехточечном шаблоне с постоянным шагом.

23.9. Для дифференциальной задачи

$$-(k(x)u')' = f, \quad x \in [0, 1], \quad u(0) = u(1) = 0,$$

$$k(x) = \begin{cases} 1 & \text{при } 0 \leq x \leq 1/2, \\ 2 & \text{при } 1/2 < x \leq 1, \end{cases}$$

построить разностную схему с помощью интегрального тождества Марчука, считая, что точка разрыва $k(x)$ является узлом равномерной сетки $\bar{D}_h = \{x_i = ih, i = 0, \dots, N, Nh = 1\}$.

23.10. Дана дифференциальная задача

$$-u'' + cu = f(x), \quad x \in [0, 1], \quad u(0) = u(1) = 0, \quad c = \text{const}.$$

При каких c для решения этой задачи можно применить метод Рунге?

23.11. Для дифференциальной задачи

$$Lu \equiv -(k(x)u')' + p(x)u = f(x), \quad x \in [0, 1], \quad u(0) = u(1) = 0,$$

с коэффициентами $k(x) = 1 + x$, $p(x) = 1$ построить разностную схему методом Рунге, используя кусочно-линейные базисные функции $\varphi_j^h(x)$ на равномерной сетке $\bar{D}_h = \{x_j = jh, j = 0, \dots, N, Nh = 1\}$.

23.12. Для дифференциальной задачи

$$-(k(x)u')' = 1, \quad x \in [0, 1], \quad u(0) = u(1) = 0,$$

$$k(x) = \begin{cases} 3/2 & \text{при } 0 \leq x < 1/4, \\ 2 & \text{при } 1/4 \leq x \leq 1, \end{cases}$$

построить разностную схему методом Рунта, используя кусочно-линейные базисные функции $\varphi_j^h(x)$ на равномерной сетке $\bar{D}_h = \{x_j = jh, j = 0, \dots, N, Nh = 1\}$ и считая, что точка разрыва $k(x)$ является узлом сетки.

23.13. Для дифференциальной задачи

$$-(k(x)u')' + p(x)u = f(x), \quad x \in [0, 1], \quad u(0) = u'(1) = 0,$$

переменные коэффициенты которой удовлетворяют условиям

$$0 < k_0 \leq k(x) \leq k_1, \quad 0 \leq p(x) \leq p_1,$$

методом Рунта построить аппроксимацию краевого условия $u'(1) = 0$, используя кусочно-линейные базисные функции $\varphi_j^h(x)$ на равномерной сетке $\bar{D}_h = \{x_j = jh, j = 0, \dots, N, Nh = 1\}$.

23.14. Показать, что для дифференциальной задачи

$$Lu \equiv -(k(x)u')' + p(x)u = f(x), \quad x \in [0, 1], \quad u(0) = 0, \quad u'(1) + \alpha u(1) = \beta,$$

переменные коэффициенты которой удовлетворяют условиям

$$0 < k_0 \leq k(x) \leq k_1, \quad 0 \leq p(x) \leq p_1,$$

функционал в методе Рунта имеет вид

$$J(v) = \int_0^1 (k(x)(v'(x))^2 + p(x)v^2(x) - 2f(x)v(x)) dx + \alpha k(1)v^2(1) - 2\beta k(1)v(1).$$

23.15. Для дифференциальной задачи

$$-u'' + a u' + p u = 1, \quad x \in [0, 1], \quad u(0) = u(1) = 1,$$

с постоянными коэффициентами a и $p \geq 0$ построить разностную схему методом Галёркина, используя кусочно-линейные базисные функции $\varphi_j^h(x)$ на равномерной сетке $\bar{D}_h = \{x_j = jh, j = 0, \dots, N, Nh = 1\}$.

§ 24. Задача Коши

Конкретизируем в случае задачи Коши для обыкновенного дифференциального уравнения

$$y' = f(x, y), \quad (1)$$

$$y(x_0) = y_0, \quad (2)$$

общие понятия метода сеток. Пусть для простоты рассматривается равномерная сетка $x_k = x_0 + kh$, $k \geq 0$. Тогда разностной схемой для задачи (1), (2) называют семейство разностных уравнений

$$\frac{1}{h} \sum_{i=0}^n a_{-i} y_{k-i} = \sum_{i=0}^n b_{-i} f_{k-i}, \quad k = n, n+1, \dots \quad (3)$$

с известными начальными условиями $y_0 = y(x_0)$, y_1, \dots, y_{n-1} , где a_{-i} , b_{-i} не зависят от h , $a_0 \neq 0$ и $f_{k-i} = f(x_{k-i}, y_{k-i})$.

Разностная схема (3) аппроксимирует дифференциальную задачу на решении $y(x)$ с порядком p , если выполнено условие нормировки

$$\lim_{h \rightarrow 0} \|f_h\|_{F_h} = \|f\|_F$$

и для функции погрешности

$$r_k^h = \frac{1}{h} \sum_{i=0}^n a_{-i} y(x_{k-i}) - \sum_{i=0}^n b_{-i} f(x_{k-i}, y(x_{k-i}))$$

справедлива оценка $\|r^h\|_{F_h} \leq ch^p$. Напомним, что положительные постоянные c и p не зависят от шага h .

В общем случае соотношение (3) — это нелинейная система, поэтому аппроксимацию левой и правой частей уравнения (1) нужно исследовать отдельно. При оценке порядка аппроксимации разностной схемы следует также учитывать порядок, с которым начальные условия аппроксимируют значения точного решения задачи (1), (2) в соответствующих узлах сетки. Если рассматривают только уравнение (1) без начального условия (2), то под разностной схемой понимают систему (3), а ее начальные условия во внимание не принимают.

Рассмотрим характеристическое уравнение разностной схемы (3) в однородном случае (т. е. в случае аппроксимации уравнения $y' = 0$):

$$F(\mu) \equiv \sum_{i=0}^n a_{-i} \mu^{n-i} = 0.$$

Схему называют α -устойчивой, если все корни характеристического уравнения принадлежат единичному кругу и на границе круга нет кратных корней. Это условие является необходимым. Известно, что для любой разностной схемы, не удовлетворяющей условию α -устойчивости, существует дифференциальное уравнение с бесконечно дифференцируемой правой частью, для которого даже при отсутствии округлений и погрешностей в начальных данных решение его разностного аналога не стремится к непрерывному решению при измельчении шага сетки.

Если в задаче не приводится конкретный вид правой части, то устойчивость понимается в смысле α -устойчивости.

- 24.1.** Показать, что необходимым и достаточным условием аппроксимации уравнения (1) разностными уравнениями (3) является выполнение равенств

$$\sum_{i=0}^n a_{-i} = 0, \quad -\sum_{i=0}^n i a_{-i} = 1, \quad \sum_{i=0}^n b_{-i} = 1.$$

- 24.2.** Проверить, аппроксимирует ли разностная схема уравнение (1):

$$1) \frac{1}{3h}(y_k - y_{k-3}) = f_{k-1};$$

$$2) \frac{1}{8h}(y_k - 3y_{k-2} + 2y_{k-3}) = \frac{1}{2}(f_{k-1} + f_{k-2});$$

$$3) \frac{1}{2h}(3y_k - 4y_{k-1} + y_{k-2}) = f_k.$$

- 24.3.** Для задачи $y' + y = x + 1$, $y(0) = 0$ рассматривается схема

$$\frac{y_{k+1} - y_{k-1}}{2h} + y_k = kh + 1, \quad y_0 = 0, \quad y_1 = 0.$$

Каков порядок аппроксимации на решении данной схемы? Можно ли его улучшить?

- 24.4.** Для задачи $y' + 5y = 5$, $y(0) = 2$ построена разностная схема

$$\frac{y_{k+1} - y_{k-1}}{2h} + 5y_k = 5, \quad y_0 = 2, \quad y_1 = 2 - 5h.$$

Исследовать ее аппроксимацию и сходимость.

24.5. Для задачи $y' + a(x)y = f(x)$, $y(0) = c$ рассматривается схема

$$\begin{aligned} \frac{y_{k+1} - y_k}{h} + (\alpha_1 a(x_k) + \alpha_2 a(x_{k+1})) (\beta_1 y_k + \beta_2 y_{k+1}) &= \\ &= \gamma_1 f(x_k) + \gamma_2 f(x_{k+1}), \\ y_0 &= c; \quad x_k = kh. \end{aligned}$$

Какими следует выбрать α_k , β_k и γ_k , чтобы получить второй порядок аппроксимации на решении?

24.6. Для уравнения $y' = f$ построить разностную схему

$$\frac{y_k - y_{k-2}}{2h} = a_1 f_k + a_0 f_{k-1} + a_{-1} f_{k-2}$$

с наивысшим порядком аппроксимации p на решении.

24.7. Исследовать устойчивость разностной схемы

$$\theta \frac{y_{k+1} - y_k}{h} + (1 - \theta) \frac{y_k - y_{k-1}}{h} = f_k \quad \text{при } \theta \in [0, 1].$$

24.8. При каких a , b и c схема

$$\frac{1}{h} (y_k + a y_{k-1} - a y_{k-3} - y_{k-4}) = b f_{k-1} + c f_{k-2} + b f_{k-3}$$

для уравнения $y' = f$ имеет максимальный порядок аппроксимации на решении? Выполнено ли условие α -устойчивости?

24.9. Исследовать сходимость решения разностной схемы

$$\begin{aligned} \frac{\varphi_k - \varphi_{k-1}}{h} + l \psi_{k-1} &= 0, \quad \varphi_0 = a, \quad h = \frac{1}{N}, \\ \frac{\psi_k - \psi_{k-1}}{h} - l \varphi_{k-1} &= 0, \quad \psi_0 = b, \quad k = 1, \dots, N, \end{aligned}$$

к решению дифференциальной задачи

$$\begin{aligned} u' + lv &= 0, \quad u(0) = a, \\ v' - lu &= 0, \quad v(0) = b \end{aligned}$$

на отрезке $x \in [0, 1]$ при $l = \text{const} \neq 0$, используя явные формулы для их решений.

24.10. Для задачи $y' = y$, $y(0) = 1$ рассмотреть схему

$$\frac{y_{k+1} - y_k}{h} = y_k, \quad y_0 = 1, \quad k \geq 0,$$

и в разложении ошибки $y(x_k) - y_k = c_1 h + c_2 h^2 + \dots$ найти постоянную c_1 для $x_k = 1$.

24.11. Для задачи $y' = y$, $y(0) = 1$ рассмотреть схему

$$\frac{y_{k+1} - y_k}{h} = \frac{y_{k+1} + y_k}{2}, \quad y_0 = 1, \quad k \geq 0,$$

и в разложении ошибки $y(x_k) - y_k = c_1 h + c_2 h^2 + \dots$ найти постоянную c_1 для $x_k = 1$.

24.12. Для задачи $y' = y$, $y(0) = 1$ рассмотреть схему

$$\frac{y_{k+1} - y_{k-1}}{2h} = y_k, \quad y_0 = 1, \quad y_1 = e^h, \quad k \geq 1,$$

и в разложении ошибки $y(x_k) - y_k = c_1 h + c_2 h^2 + \dots$ найти постоянную c_1 для $x_k = 1$.

24.13. Для задачи $y' = y$, $y(0) = 1$ рассмотреть схему

$$4 \frac{y_{k+1} - y_{k-1}}{2h} - 3 \frac{y_{k+1} - y_k}{h} = y_k, \quad y_0 = 1, \quad y_1 = e^h, \quad k \geq 1,$$

и в разложении ошибки $y(x_k) - y_k = c_1 h + c_2 h^2 + \dots$ найти постоянные c_1 и c_2 для $x_k = 1$.

Если шаблон сетки содержит l узлов, схему часто называют *l-точечной*.

24.14. Для задачи $y' + y = \cos 2x$, $y(0) = 0$ построить трехточечную разностную схему второго порядка сходимости.

24.15. Для задачи $y' + 5y = \sin 2x$, $y(0) = 2$ построить двухточечную разностную схему второго порядка сходимости.

24.16. Для задачи $y' - y = \exp 2x$, $y(0) = 1$ построить трехточечную разностную схему второго порядка сходимости.

24.17. Для задачи $y' - 2y = \exp x$, $y(0) = 1$ построить двухточечную разностную схему второго порядка сходимости.

24.18. Привести пример неустойчивой разностной схемы, аппроксимирующей на решении уравнение $y' = f$ строго: 1) с первым порядком; 2) со вторым порядком; 3) с третьим порядком.

24.19. Найти главный член погрешности аппроксимации на решении и исследовать устойчивость разностной схемы

$$1) \frac{y_k - y_{k-2}}{2h} = \frac{f_k + 4f_{k-1} + f_{k-2}}{6},$$

$$2) \frac{y_{k+1} - y_k}{h} = \frac{5f_k + 8f_{k+1} - f_{k+2}}{12},$$

$$3) \frac{y_{k+4} - y_k}{4h} = \frac{2f_{k+1} - f_{k+2} + 2f_{k+3}}{3},$$

$$4) \frac{y_k + 4y_{k-1} - 5y_{k-2}}{6h} = \frac{2f_{k-1} + f_{k-2}}{3}.$$

§ 25. Линейная краевая задача

Первая краевая задача для линейного обыкновенного дифференциального уравнения второго порядка имеет вид

$$-(k(x)u')' + p(x)u = f(x), \quad x \in [0, 1], \quad u(0) = u(1) = 0.$$

Предполагаем далее, что коэффициенты уравнения удовлетворяют условиям

$$0 < k_0 \leq k(x) \leq k_1, \quad 0 \leq p(x) \leq p_1.$$

В общем случае на любом из концов отрезка $[0, 1]$ краевое условие может быть задано в виде линейной комбинации функции и производной: $au + bu' = c$. Тогда следует обратить внимание на способ его аппроксимации.

Если это не оговаривается специально, то в задачах этого параграфа сетка выбирается равномерной: $x_i = ih, i = 0, \dots, N, Nh = 1$.

25.1. Определить локальный порядок аппроксимации в точке $x_i = ih$ для операторов L и L_h :

$$Lu = (k(x)u')',$$

$$(L_h u_h)_i = \frac{1}{h} \left[k(x_{i+1/2}) \frac{u_{i+1} - u_i}{h} - k(x_{i-1/2}) \frac{u_i - u_{i-1}}{h} \right],$$

считая коэффициент $k(x)$ достаточно гладким.

25.2. Для дифференциальной задачи $-u'' + u = f(x), u(0) = u(1) = 0$ рассмотреть разностную схему

$$-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + (\alpha u_{i+1} + \beta u_i + \gamma u_{i-1}) = f(x_i) + \frac{h^2}{12} f''(x_i),$$

$$1 \leq i \leq N-1, \quad u_0 = u_N = 0, \quad Nh = 1.$$

При каких α, β и γ аппроксимация на решении имеет четвертый порядок?

25.3. Достаточно гладкая функция $u(x)$ является решением уравнения $-u'' + p(x)u = f(x)$. Используя ее значения в точках $x_0 = 0$ и $x_1 = h$, построить аппроксимацию второго порядка для краевого условия $au(0) + bu'(0) = c$.

25.4. Для дифференциальной задачи $-u'' + p(x)u = f(x)$, $u'(0) = 1$, $u(1) = 0$ построить разностную схему второго порядка аппроксимации на сетке $\overline{D}_h = \{x_i = (i - 1/2)h, i = 0, \dots, N, h = 1/(N - 1/2)\}$.

25.5. Исследовать устойчивость разностной схемы

$$-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} = f_i, \quad 1 \leq i \leq N - 1, \quad u_0 = u_N = 0, \quad Nh = 1,$$

и показать, что при $h \rightarrow 0$ число обусловленности матрицы алгебраической системы для нахождения неизвестных u_i имеет порядок $O(h^{-2})$.

25.6. Для решения разностной схемы из **25.5** получить на основе принципа максимума оценку скорости сходимости

$$\max_{0 \leq i \leq N} |u(x_i) - u_i| \leq \frac{h^2}{96} \max_{[0,1]} |f''(x)|$$

к решению дифференциальной задачи $-u'' = f$, $u(0) = u(1) = 0$, при $f \in C^{(2)}[0, 1]$.

25.7. Функция $u(x)$ является решением уравнения $u'' = \cos x + 1$. Используя ее значения в точках $x_N = 1$ и $x_{N-1} = 1 - h$, построить аппроксимацию второго порядка для краевого условия $u'(1) - 3u(1) = 1$.

25.8. Функция $u(x)$ является решением уравнения $u'' - x^2 u = 1$. Используя ее значения в точках $x_0 = 0$ и $x_1 = h$, построить аппроксимацию второго порядка для краевого условия $u'(0) + 4u(0) = 1$.

25.9. Функция $u(x)$ является решением уравнения $u'' - 3u = \exp x$. Используя ее значения в точках $x_N = 1$ и $x_{N-1} = 1 - h$, построить аппроксимацию второго порядка для краевого условия $u'(1) = 0$.

25.10. Функция $u(x)$ является решением уравнения $u'' - 2u = \sin x - 1$. Используя ее значения в точках $x_0 = 0$ и $x_1 = h$, построить аппроксимацию второго порядка для краевого условия $u'(0) - u(0) = 0$.

25.11. Исследовать устойчивость разностной схемы

- 1) $-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + u_i = f_i, 0 < i < N, u_0 = u_1, u_{N-1} = u_N, Nh = 1;$
- 2) $-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} = f_i, 0 < i < N, u_0 = 0, u_{N-1} = u_N, Nh = 1;$
- 3) $-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} = f_i, 0 < i < N, u_0 = u_1, u_N = 0, Nh = 1;$
- 4) $-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} - (2 + \cos(2\pi x_i))u_i = f_i, 0 < i < N, u_0 = u_N = 0, Nh = 1.$

Вспомогательные формулы. Пусть φ_i — функция целочисленного аргумента i . Введем обозначения для разностей первого порядка:

$$\Delta\varphi_i = \varphi_{i+1} - \varphi_i, \quad \nabla\varphi_i = \varphi_i - \varphi_{i-1}.$$

Разности более высокого порядка определяются рекуррентно, например:

$$\Delta^m\varphi_i = \Delta(\Delta^{m-1}\varphi_i), \quad m \geq 1.$$

Известна формула дифференцирования произведения функций

$$(u(x)v(x))' = u(x)v'(x) + u'(x)v(x).$$

25.12. Показать, что формулы «разностного дифференцирования» имеют вид

$$\begin{aligned} \Delta(u_i v_i) &= u_i \Delta v_i + v_{i+1} \Delta u_i = u_{i+1} \Delta v_i + v_i \Delta u_i, \\ \nabla(u_i v_i) &= u_{i-1} \nabla v_i + v_i \nabla u_i = u_i \nabla v_i + v_{i-1} \nabla u_i. \end{aligned}$$

Рассмотрим формулу интегрирования по частям

$$\int_a^b u'(x)v(x) dx = u(x)v(x)|_a^b - \int_a^b u(x)v'(x) dx,$$

и введем обозначения

$$(\varphi, \psi) = \sum_{i=1}^{N-1} \varphi_i \psi_i, \quad [\varphi, \psi] = \sum_{i=1}^N \varphi_i \psi_i, \quad [\varphi, \psi] = \sum_{i=0}^{N-1} \varphi_i \psi_i$$

для сумм, являющихся аналогами интеграла $\int_a^b \varphi(x)\psi(x) dx$.

25.13. С помощью формулы Абеля

$$\sum_{i=0}^{N-1} (a_{i+1} - a_i) b_i = - \sum_{i=0}^{N-1} (b_{i+1} - b_i) a_{i+1} + a_N b_N - a_0 b_0$$

показать справедливость формулы суммирования по частям:

$$(\varphi, \Delta\psi) = -(\nabla\varphi, \psi] + \varphi_N \psi_N - \varphi_0 \psi_1.$$

Для дифференциального оператора $Lv(x) = (k(x)v'(x))' - p(x)v(x)$ формулы

$$\begin{aligned} \int_a^b u(x) Lv(x) dx &= - \int_a^b k(x) u'(x) v'(x) dx - \int_a^b p(x) u(x) v(x) dx + \\ &\quad + k(x) u(x) v'(x) \Big|_a^b, \\ \int_a^b [u(x) Lv(x) - Lu(x) v(x)] dx &= k(x) [u(x) v'(x) - u'(x) v(x)] \Big|_a^b, \end{aligned}$$

называют, соответственно, *первой* и *второй формулами Грина*.

25.14. Доказать справедливость соотношений

$$\begin{aligned} (\varphi, \Delta\nabla\psi) &= -(\nabla\varphi, \nabla\psi) + \varphi_{N-1} \nabla\psi_N - \varphi_0 \nabla\psi_1, \\ (\varphi, \Delta\nabla\psi) - (\Delta\nabla\varphi, \psi) &= \varphi_{N-1} \psi_N - \varphi_N \psi_{N-1} + \varphi_1 \psi_0 - \varphi_0 \psi_1, \end{aligned}$$

25.15. Вывести дискретные аналоги формул Грина для разностного оператора

$$\Lambda\varphi_i = \Delta(a_i \nabla\varphi_i) - d_i \varphi_i \equiv a_{i+1}(\varphi_{i+1} - \varphi_i) - a_i(\varphi_i - \varphi_{i-1}) - d_i \varphi_i.$$

Исследование устойчивости методом априорных оценок. Опишем этот метод на примере дифференциальной задачи

$$-u'' + p(x)u = f(x), \quad x \in [0, 1], \quad u(0) = u(1) = 0, \quad p(x) \geq 0.$$

Умножим уравнение на u и возьмем интеграл по отрезку $[0, 1]$. Проинтегрировав по частям, получим интегральное тождество

$$\int_0^1 (u'(x))^2 dx + \int_0^1 p u^2 dx = \int_0^1 f u dx.$$

Выведем теперь неравенство, связывающее интегралы от квадратов функции и ее производной. При условии $u(0) = 0$ для любого $x \in [0, 1]$ справедливо равенство $u(x) = \int_0^x u'(t) dt$, откуда

$$\begin{aligned} \int_0^1 u^2(x) dx &= \int_0^1 \left(\int_0^x u'(t) dt \right)^2 dx \leq \\ &\leq \int_0^1 \left(\int_0^x dt \right) \left(\int_0^x (u'(t))^2 dt \right) dx \leq \int_0^1 (u'(t))^2 dt. \end{aligned}$$

Применив к интегральному тождеству полученное неравенство, получим

$$\int_0^1 u^2 dx \leq \int_0^1 (u'(x))^2 dx + \int_0^1 p u^2 dx = \int_0^1 f u dx \leq \frac{1}{2} \left(\int_0^1 f^2 dx + \int_0^1 u^2 dx \right).$$

Отсюда следует априорная оценка для решения $u(x)$:

$$\|u\|_{L_2} \leq \|f\|_{L_2}, \quad \text{где} \quad \|u\|_{L_2}^2 = \int_0^1 u^2 dx,$$

означающая устойчивость дифференциальной задачи по правой части.

25.16. Исследовать методом априорных оценок при $p_i \geq 0$ устойчивость разностной схемы

$$\begin{aligned} -\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + p_i u_i &= f_i, \\ 1 \leq i \leq N - 1, \quad u_0 = u_N = 0, \quad Nh &= 1. \end{aligned}$$

Неравенства для сеточных функций. Учитывая определения из вспомогательных формул, введем следующие обозначения:

$$\begin{aligned} \|y\| &= \sqrt{h} [y, y]^{1/2}, \quad \|y\|_C = \max_{0 \leq i \leq N} |y_i|, \\ (y_{\bar{x}})_i &= (\nabla y)_i / h = (y_i - y_{i-1}) / h, \quad \|y_{\bar{x}}\| = \sqrt{h} (y_{\bar{x}}, y_{\bar{x}})^{1/2}, \end{aligned}$$

где h — постоянный шаг сетки $\bar{D}_h = \{x_i = ih, i = 0, \dots, N, h = l/N\}$.

25.17. Пусть $y_0 = y_N = 0$ и $l = 1$. Доказать неравенство $\|y\|_C \leq \|y_{\bar{x}}\|/2$.

Пусть $y_0 = y_N = 0$. Доказать неравенства

$$25.18. \|y\|_C \leq \sqrt{l} \|y_{\bar{x}}\|/2.$$

$$25.19. \|y\| \leq \sqrt{l} \|y\|_C.$$

$$25.20. h \|y_{\bar{x}}\|/2 \leq \|y\| \leq l \|y_{\bar{x}}\|/(2\sqrt{2}).$$

$$25.21. \text{ Пусть } y_0 = 0. \text{ Доказать неравенство } \|y\|_C \leq \sqrt{l} \|y_{\bar{x}}\|.$$

$$25.22. \text{ Для произвольной сеточной функции } y_i \text{ доказать неравенства } \|y\|_C^2 \leq 2(l \|y_{\bar{x}}\|^2 + y_0^2) \text{ и } \|y\|_C^2 \leq 2(l \|y_{\bar{x}}\|^2 + y_N^2).$$

§ 26. Гиперболические уравнения

Построение и исследование разностных схем для дифференциальных уравнений с частными производными, имеющих гиперболический тип, традиционно проводят в открытой полуплоскости

$$D = \{(x, t) : \infty > x > -\infty, t > 0\}$$

на примере линейного уравнения переноса

$$Lu \equiv u_t(x, t) + a(x, t) u_x(x, t) = f(x, t)$$

с начальным условием $u(x, 0) = \varphi(x)$.

Если это не оговаривается специально, то сетку выбирают равномерной по обеим переменным

$$x_m = mh, \quad m = 0, \pm 1, \dots; \quad t_n = n\tau, \quad n = 0, 1, \dots,$$

а для сеточной функции u в точке (x_m, t_n) используют обозначение u_m^n .

26.1. Определить порядок аппроксимации разностной схемы

$$\frac{u_m^{n+1} - u_m^n}{\tau} + a \frac{u_m^n - u_{m-1}^n}{h} = 0$$

для уравнения $u_t + au_x = 0$, $a = \text{const} > 0$. При каком соотношении τ и h решение дифференциального уравнения в узлах сетки совпадает с решением разностной схемы?

26.2. Определить порядок аппроксимации разностной схемы

$$\frac{u_m^{n+1} - \frac{u_{m+1}^n + u_{m-1}^n}{2}}{\tau} + a \frac{u_{m+1}^n - u_{m-1}^n}{2h} = 0$$

для уравнения $u_t + au_x = 0$.

26.3. Для однородного уравнения $u_t + a u_x = 0$, $a = \text{const}$, построить схемы первого и второго порядков аппроксимации на решении (если это возможно), используя шаблон (x_m, t_n) , (x_m, t_{n+1}) , (x_{m+1}, t_n) и условие $\tau = rh$, $r = \text{const}$.

26.4. Для уравнения $u_t - u_x = f$ построить разностную схему

$$a^0 u_m^{n+1} + a_{-1} u_{m-1}^n + a_0 u_m^n + a_1 u_{m+1}^n = \varphi_m^n$$

максимального порядка аппроксимации при условии $\tau = rh$, $r = \text{const}$.

26.5. Пусть для задачи $u_t - u_x = f$, $u(x, 0) = \varphi(x)$, используется схема

$$\frac{u_m^{n+1} - u_m^{n-1}}{2\tau} - \frac{u_{m+1}^n - u_{m-1}^n}{2h} = f_m^n, \quad u_m^0 = \varphi(mh).$$

Как определить значения функции u_m^1 , чтобы не ухудшить порядок аппроксимации на решении?

26.6. Для уравнения $u_t + a u_x = 0$, $a = \text{const}$, рассматривается схема с пересчетом

$$\frac{u_{m+1/2}^{n+1/2} - \frac{u_{m+1}^n + u_m^n}{2}}{\tau/2} + a \frac{u_{m+1}^n - u_m^n}{h} = 0,$$

$$\frac{u_m^{n+1} - u_m^n}{\tau} + a \frac{u_{m+1/2}^{n+1/2} - u_{m-1/2}^{n+1/2}}{h} = 0.$$

Определить ее порядок аппроксимации на решении.

26.7. Для уравнения $u_t + u_x = 0$ рассматривается семейство схем с параметром θ

$$\frac{u_m^{n+1} - u_m^n}{\tau} + \theta \frac{u_m^{n+1} - u_{m-1}^{n+1}}{h} + (1 - \theta) \frac{u_m^n - u_{m-1}^n}{h} = 0.$$

При каких значениях θ схема имеет на решении порядок аппроксимации $O(\tau^2 + h^2)$?

26.8. Для уравнения $u_t + u_x = 0$ построить схему с порядком аппроксимации $O(\tau^2 + h^4)$ на решении, используя шаблон из десяти точек: $(x_{m\pm 2}, t_k)$, $(x_{m\pm 1}, t_k)$, (x_m, t_k) , $k = n, n + 1$.

26.9. Для уравнения переноса $u_t + a(x, t)u_x = f(x, t)$ построить двухслойную схему порядка аппроксимации на решении:

1) $O(\tau^2 + h)$; 2) $O(\tau + h^2)$; 3) $O(\tau^2 + h^2)$; 4) $O(\tau + h)$

с минимальным, по возможности, числом узлов l в шаблоне.

Спектральный признак устойчивости. Разностные схемы для однородного уравнения переноса с постоянным коэффициентом a можно записать так:

$$L_h u_m^n \equiv \sum_{k,l} b_{lk} u_{m+l}^{n+k} = 0.$$

Рассмотрим их частные решения вида $u_m^n = (\lambda(\varphi))^n e^{im\varphi}$.

Спектральный признак устойчивости разностной схемы состоит в следующем: если при заданном законе стремления τ и h к нулю существует постоянная $0 \leq c < \infty$, такая, что для всех φ справедливо неравенство $|\lambda(\varphi)| \leq e^{c\tau}$, то спектральный признак выполнен, и схема может быть применена для численного решения соответствующей задачи Коши для уравнения $Lu = f$.

Известно, что если спектральный признак устойчивости не выполняется, то для решения задачи не существует априорной оценки вида $\|u^n\| \leq M$ с константой M , не зависящей от параметров сетки, в норме $\|\cdot\|$, которая не зависит от временного слоя n .

В следующих задачах требуется исследовать устойчивость разностной схемы с постоянным коэффициентом a с помощью спектрального признака.

$$26.10. \quad \frac{u_m^{n+1} - u_m^n}{\tau} + a \frac{u_m^n - u_{m-1}^n}{h} = 0.$$

$$26.11. \quad \frac{u_m^{n+1} - u_m^n}{\tau} + a \frac{u_{m+1}^n - u_{m-1}^n}{2h} = 0.$$

$$26.12. \quad \frac{u_m^{n+1} - u_m^n}{\tau} + a \frac{u_{m+1}^n - u_{m-1}^n}{2h} - \frac{h^2}{2\tau} \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2} = 0.$$

$$26.13. \quad \frac{u_m^{n+1} - u_m^n}{\tau} + a \frac{u_{m+1}^n - u_{m-1}^n}{2h} - \frac{a^2\tau}{2} \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2} = 0.$$

$$26.14. \quad \frac{u_m^{n+1} - u_m^n}{\tau} + a \frac{u_m^{n+1} - u_{m-1}^{n+1}}{h} = 0.$$

$$26.15. \frac{u_m^{n+1} - u_m^n}{\tau} + a \frac{u_{m+1}^{n+1} - u_{m-1}^{n+1}}{2h} = 0.$$

$$26.16. \frac{u_m^{n+1} - \frac{u_{m+1}^n + u_{m-1}^n}{2}}{\tau} + a \frac{u_{m+1}^n - u_{m-1}^n}{2h} = 0.$$

$$26.17. \frac{u_m^{n+1} - \frac{u_{m+1}^{n+1} + u_{m-1}^{n+1}}{2}}{\tau} + a \frac{u_{m+1}^{n+1} - u_{m-1}^{n+1}}{2h} = 0.$$

26.18. Для уравнения $u_t + u_x = 0$ рассматривается семейство схем с параметром θ

$$\frac{u_m^{n+1} - u_m^n}{\tau} + \theta \frac{u_{m+1}^n - u_m^n}{h} + (1 - \theta) \frac{u_m^n - u_{m-1}^n}{h} = 0.$$

При каких значениях $\theta \in [0, 1]$ схема устойчива?

26.19. Для уравнения $u_t + u_x = 0$ рассматривается семейство схем с параметром θ

$$\frac{u_m^{n+1} - u_m^n}{\tau} + \theta \frac{u_{m+1}^{n+1} - u_{m-1}^{n+1}}{h} + (1 - \theta) \frac{u_m^n - u_{m-1}^n}{h} = 0.$$

При каких значениях $\theta \in [0, 1]$ схема безусловно устойчива?

§ 27. Параболические уравнения

Построение и исследование разностных схем для дифференциальных уравнений с частными производными, имеющих параболический тип, как правило, проводят в открытой полуполосе

$$D = \{(x, t) : 1 > x > 0, t > 0\}$$

на примере простейшего уравнения теплопроводности

$$Lu \equiv u_t(x, t) - u_{xx}(x, t) = f(x, t)$$

с начальным $u(x, 0) = \varphi(x)$ при $t = 0$ и краевыми условиями $u(0, t) = u(1, t) = 0$ при $\forall t \geq 0$. Предполагается, что начальная функция $\varphi(x)$ удовлетворяет краевым условиям.

В общем случае на любом из концов отрезка $[0, 1]$ краевое условие может быть задано в виде линейной комбинации функции u и ее производной u_x . Тогда необходимо обратить внимание на способ его аппроксимации.

Характерной особенностью задач с уравнениями параболического типа является смешанный тип граничных данных: краевые условия по переменной x и начальные по $-t$. Поэтому исследование аппроксимации такое же, как в гиперболическом и эллиптическом случаях, а исследование устойчивости проводят специальным образом.

Если это не оговаривается специально, то сетку выбирают равнономерной по обеим переменным

$$x_m = mh, \quad m = 0, 1, \dots, M, \quad Mh = 1; \quad t_n = n\tau, \quad n = 0, 1, \dots,$$

для сеточной функции u в точке (x_m, t_n) используют обозначение u_m^n , а краевые условия берут однородными: $u_0^n = u_M^n = 0 \forall n$.

27.1. При каком соотношении τ и h разностная схема

$$\frac{u_m^{n+1} - u_m^n}{\tau} = \frac{u_{m-1}^n - 2u_m^n + u_{m+1}^n}{h^2}$$

имеет на решении порядок аппроксимации $O(\tau^2 + h^4)$?

27.2. При каких θ разностная схема

$$\frac{u_m^{n+1} - u_m^n}{\tau} = \theta \frac{u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}}{h^2} + (1-\theta) \frac{u_{m-1}^n - 2u_m^n + u_{m+1}^n}{h^2}$$

имеет на решении порядок аппроксимации $O(\tau^2 + h^4)$?

27.3. Определить порядок аппроксимации на решении схемы

$$\begin{aligned} & \frac{1}{12} \frac{u_{m+1}^{n+1} - u_{m+1}^n}{\tau} + \frac{5}{6} \frac{u_m^{n+1} - u_m^n}{\tau} + \frac{1}{12} \frac{u_{m-1}^{n+1} - u_{m-1}^n}{\tau} = \\ & = \frac{1}{2} \left(\frac{u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}}{h^2} + \frac{u_{m-1}^n - 2u_m^n + u_{m+1}^n}{h^2} \right). \end{aligned}$$

27.4. Для уравнения теплопроводности построить схему наивысшего порядка аппроксимации на решении, используя шаблон из перечисленных точек сетки:

- 1) $(x_{m-1}, t_{n-1}), (x_{m+1}, t_{n+1}), (x_m, t_k), k = n-1, n, n+1$;
- 2) $(x_{m\pm 1}, t_k), (x_m, t_k), k = n, n+1$;
- 3) $(x_{m\pm 1}, t_n), (x_m, t_k), k = n-1, n, n+1$;
- 4) $(x_{m-1}, t_{n+1}), (x_{m+1}, t_{n-1}), (x_m, t_k), k = n-1, n, n+1$.

Анализ устойчивости схем в равномерной метрике. Определим норму сеточной функции u_m^n на n -м временном слое

$$\|u^n\|_C = \max_{0 \leq m \leq M} |u_m^n|.$$

Схему для простейшего уравнения теплопроводности называют *устойчивой в равномерной метрике* на отрезке $[0, T]$, $T = N\tau$, если имеет место неравенство

$$\max_{1 \leq n \leq N} \|u^n\|_C \leq \|u^0\| + K \max_{0 \leq n \leq N} \|f^n\|_C,$$

где величина K не зависит от шагов сетки τ и h , но может зависеть от T .

При исследовании устойчивости разностных схем в случае краевых условий первого рода важную роль играют сеточные функции

$$y_m^{(k)} = \sin(\pi k m h), \quad m = 0, \dots, M, \quad k = 1, \dots, M-1,$$

являющиеся решениями задачи на собственные значения

$$\frac{y_{m+1} - 2y_m + y_{m-1}}{h^2} = -\lambda y_m, \quad y_0 = y_M = 0, \quad h = 1/M.$$

С их помощью несложно строить частные решения разностных схем для однородного уравнения $u_t - u_{xx} = 0$ с однородными краевыми условиями:

$$u_m^n = \mu_h^n(k) y_m^{(k)}, \quad m = 0, \dots, M, \quad k = 1, \dots, M-1.$$

27.5. Найти порядок аппроксимации явной разностной схемы

$$\frac{u_m^{n+1} - u_m^n}{\tau} = \frac{u_{m-1}^n - 2u_m^n + u_{m+1}^n}{h^2} + f_m^n, \quad 1 \leq m \leq M-1,$$

$$u_m^0 = \varphi(mh), \quad u_0^n = u_M^n = 0 \quad \forall n \geq 0,$$

и исследовать ее устойчивость при $\tau \leq h^2/2$.

27.6. Исследовать аппроксимацию и устойчивость полностью неявной разностной схемы

$$\frac{u_m^{n+1} - u_m^n}{\tau} = \frac{u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}}{h^2} + f_m^{n+1}, \quad 1 \leq m \leq M-1,$$

$$u_m^0 = \varphi(mh), \quad u_0^n = u_M^n = 0 \quad \forall n \geq 0.$$

27.7. Первая краевая задача для однородного уравнения теплопроводности $u_t = u_{xx}$ аппроксимируется явной разностной схемой

$$\frac{u_m^{n+1} - u_m^n}{\tau} = \frac{u_{m-1}^n - 2u_m^n + u_{m+1}^n}{h^2}, \quad 1 \leq m \leq M-1,$$

$$u_m^0 = \varphi(mh), \quad u_0^n = u_M^n = 0 \quad \forall n \geq 0.$$

Определить порядок сходимости решения разностной схемы к решению дифференциальной задачи при различных $\rho = \tau/h^2$.

27.8. Доказать, что явная разностная схема

$$\frac{u_m^{n+1} - u_m^n}{\tau} = \frac{u_{m-1}^n - 2u_m^n + u_{m+1}^n}{h^2}, \quad 1 \leq m \leq M-1,$$

$$u_m^0 = \varphi(mh), \quad u_0^n = u_M^n = 0 \quad \forall n \geq 0$$

неустойчива, если $\limsup_{\tau, h \rightarrow 0} \frac{\tau/h^2 - 1/2}{\tau} = \infty$.

27.9. Исследовать устойчивость разностной схемы

$$\frac{u_m^{n+1} - u_m^{n-1}}{2\tau} = \frac{u_{m-1}^n - 2u_m^n + u_{m+1}^n}{h^2}, \quad 1 \leq m \leq M-1,$$

$$u_0^n = u_M^n = 0 \quad \forall n \geq 0.$$

Анализ устойчивости схем в интегральной метрике. Положим

$$\|u^n\|_{L_{2,h}} = \left(h \sum_{m=1}^{M-1} (u_m^n)^2 \right)^{1/2}$$

и будем называть однородную разностную схему *устойчивой по начальным данным в метрике $L_{2,h}$* на отрезке $[0, T]$, $T = N\tau$, если справедливо неравенство

$$\max_{1 \leq n \leq N} \|u^n\|_{L_{2,h}} \leq K \|u^0\|_{L_{2,h}},$$

где величина K не зависит от шагов сетки τ и h .

27.10. При каких $\theta \in [0, 1]$ разностная схема

$$\frac{u_m^{n+1} - u_m^n}{\tau} = \theta \frac{u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}}{h^2} + (1-\theta) \frac{u_{m-1}^n - 2u_m^n + u_{m+1}^n}{h^2},$$

$$1 \leq m \leq M-1, \quad u_m^0 = u_0(mh), \quad u_0^n = u_M^n = 0 \quad \forall n \geq 0,$$

является устойчивой по начальным данным в интегральной метрике?

- 27.11.** Однородное уравнение теплопроводности $u_t = u_{xx}$ аппроксимируется схемой Дюфорта–Франкела (другое название схемы — «ромб»)

$$\frac{u_m^{n+1} - u_m^{n-1}}{2\tau} = \frac{u_{m+1}^n - u_m^{n-1} - u_m^{n+1} + u_{m-1}^n}{h^2},$$

$$1 \leq m \leq M - 1, \quad u_0^n = u_M^n = 0 \quad \forall n \geq 1.$$

Выяснить условия ее устойчивости по начальным данным в интегральной метрике и показать, что если $h \rightarrow 0, \tau \rightarrow 0$ так, что $\tau/h = r \neq 0$, то эта схема аппроксимирует гиперболическое уравнение $u_t + r^2 u_{tt} = u_{xx}$.

В следующих задачах требуется исследовать устойчивость разностной схемы по начальным данным в интегральной метрике.

27.12.
$$\frac{u_m^{n+1} - u_m^{n-1}}{2\tau} = \frac{u_{m-1}^{n+1} - 2u_m^{n+1} + u_{m+1}^{n+1}}{h^2}, \quad 1 \leq m \leq M - 1,$$

$$u_0^n = u_M^n = 0 \quad \forall n \geq 1.$$

27.13.
$$\frac{u_m^{n+1} - u_m^{n-1}}{2\tau} = \frac{u_{m-1}^{n-1} - 2u_m^{n-1} + u_{m+1}^{n-1}}{h^2}, \quad 1 \leq m \leq M - 1,$$

$$u_0^n = u_M^n = 0 \quad \forall n \geq 1.$$

27.14.
$$\frac{u_m^{n+1} - u_m^{n-1}}{2\tau} = \frac{u_{m-1}^n - 2u_m^{n+1} + u_{m+1}^n}{h^2}, \quad 1 \leq m \leq M - 1,$$

$$u_0^n = u_M^n = 0 \quad \forall n \geq 1.$$

§ 28. Эллиптические уравнения

Построение и исследование разностных схем для дифференциальных уравнений с частными производными, имеющих эллиптический тип, в простейшем случае проводят в области прямоугольной формы

$$D = \{(x, y): X > x > 0, Y > y > 0\}$$

на примере уравнения с переменными коэффициентами $a_1, a_2 \geq \geq a_0 > 0$

$$Lu \equiv (a_1(x, y)u_x(x, y))_x + (a_2(x, y)u_y(x, y))_y = f(x, y)$$

и заданными однородными краевыми условиями первого рода: $u(0, y) = u(X, y) = 0$ при $Y \geq y \geq 0$ и $u(x, 0) = u(x, Y) = 0$ при $X \geq x \geq 0$.

Наиболее часто рассматривается уравнение Пуассона ($a_1(x, y) = a_2(x, y) \equiv 1$):

$$\Delta u \equiv u_{xx}(x, y) + u_{yy}(x, y) = f(x, y).$$

В общем случае на любой части границы краевое условие для него может быть задано в виде линейной комбинации функции u и ее производной первого порядка по направлению нормали к границе. Тогда необходимо обратить внимание на способ его аппроксимации.

Типичным примером эллиптического оператора четвертого порядка является бигармонический оператор:

$$\Delta^2 u = u_{xxxx}(x, y) + 2u_{xxyy}(x, y) + u_{yyyy}(x, y),$$

для которого краевое условие может содержать линейную комбинацию производных неизвестной функции до третьего порядка включительно.

Особенность постановки эллиптических задач — наличие только краевых условий. Поэтому аппроксимацию и устойчивость исследуют как в случае линейных краевых задач для обыкновенных дифференциальных уравнений.

Для удобства независимые переменные x_1, x_2 в двумерном случае будем также обозначать через x, y .

Введем обозначение $\Lambda_\alpha u(x_1, x_2)$, $\alpha = 1, 2$, для разностного аналога оператора второй производной $L_\alpha u = u_{x_\alpha x_\alpha}$ по переменной x_α , например,

$$\Lambda_1 u(x_1, x_2) = \frac{u(x_1 + h_1, x_2) - 2u(x_1, x_2) + u(x_1 - h_1, x_2)}{h_1^2}.$$

Аналогичный смысл имеет выражение

$$\Lambda_1 u_{i,j} = \frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h_1^2}.$$

28.1. Оценить погрешность аппроксимации оператора Лапласа Δ оператором $\Delta^h = \Lambda_1 + \Lambda_2$ (стандартная аппроксимация на шаблоне «крест»).

- 28.2.** Построить аппроксимацию оператора Лапласа и оценить ее погрешность на шаблоне «косой крест» при $h_1 = h_2 = h$:

$$\Delta^h u(x_1, x_2) = h^{-2} [a_{0,0}u(x_1, x_2) + a_{1,1}u(x_1 + h, x_2 + h) + a_{1,-1}u(x_1 + h, x_2 - h) + a_{-1,1}u(x_1 - h, x_2 + h) + a_{-1,-1}u(x_1 - h, x_2 - h)],$$

где коэффициенты $a_{i,j}$ не зависят от h .

- 28.3.** Построить аппроксимацию оператора Лапласа и оценить ее погрешность на треугольной сетке (область разбита на непересекающиеся правильные треугольники со стороной h).
- 28.4.** Используя значения функции u в центре A_0 и в вершинах A_k правильного n -угольника со стороной h , получить аппроксимацию оператора Лапласа в центре многоугольника. Оценить ее порядок для различных n .
- 28.5.** Построить все девятиточечные разностные аппроксимации оператора Лапласа с независимыми от h коэффициентами a_{ij} , которые имеют вид

$$\Delta^h u(x_1, x_2) = h^{-2} [a_{0,0}u(x_1, x_2) + a_{1,0}u(x_1 + h, x_2) + a_{-1,0}u(x_1 - h, x_2) + a_{0,1}u(x_1, x_2 + h) + a_{0,-1}u(x_1, x_2 - h) + a_{1,1}u(x_1 + h, x_2 + h) + a_{1,-1}u(x_1 + h, x_2 - h) + a_{-1,1}u(x_1 - h, x_2 + h) + a_{-1,-1}u(x_1 - h, x_2 - h)]$$

и обладают вторым порядком аппроксимации, т. е. $\Delta^h u(x_1, x_2) - \Delta u(x_1, x_2) = O(h^2)$ при $u \in C^{(4)}$.

- 28.6.** Какие из разностных операторов в **28.5** отрицательно определены?
- 28.7.** Построить тринадцатиточечную разностную аппроксимацию бигармонического оператора Δ^2 , использующую узлы (x_1, x_2) , $(x_1 \pm h, x_2)$, $(x_1, x_2 \pm h)$, $(x_1 \pm 2h, x_2)$, $(x_1, x_2 \pm 2h)$, $(x_1 \pm h, x_2 \pm h)$, и оценить погрешность аппроксимации на функциях $u \in C^{(6)}$.
- 28.8.** Если u — гармоническая функция в ограниченной области D с границей Γ , то $\int_{\Gamma} u_n d\Gamma = 0$, где u_n — производная по направлению нормали к границе Γ . Сформулировать и доказать

аналог этого равенства для решений разностного уравнения

$$\Delta^h u_{i,j} \equiv (\Lambda_1 + \Lambda_2) u_{i,j} = 0, \quad 1 \leq i < N_1, \quad 1 \leq j < N_2,$$

в прямоугольной области, покрытой равномерной сеткой с шагом h ($N_1 h = X$, $N_2 h = Y$).

- 28.9.** Построить разностную схему во внутренних узлах сетки для уравнения Пуассона с аппроксимацией на решении $O(h^4)$.
- 28.10.** Построить разностную схему во внутренних узлах сетки для уравнения Пуассона с аппроксимацией на решении $O(h^6)$.
- 28.11.** Для уравнения $\Delta u = f$ построить аппроксимацию на решении с порядком $O(h^2)$ граничного условия $\partial u / \partial x_1 - \alpha u = 0$ при $x_1 = 0$, используя минимальное количество узлов вдоль оси x_1 .
- 28.12.** Для уравнения $\Delta u = f$ построить аппроксимацию на решении с порядком $O(h^4)$ граничного условия $\partial u / \partial x_1 - \alpha u = 0$ при $x_1 = 0$, используя минимальное количество узлов вдоль оси x_1 .
- 28.13.** Пусть в прямоугольной области, покрытой равномерной сеткой с шагом h , определен разностный аналог оператора Лапласа:

$$\Delta^h u_{i,j} \equiv (\Lambda_1 + \Lambda_2) u_{i,j}, \quad 1 \leq i < N_1, \quad 1 \leq j < N_2.$$

Показать, что если справедливо неравенство $\Delta^h u_{i,j} \leq 0$ при всех $1 \leq i < N_1$, $1 \leq j < N_2$, то функция $u_{i,j}$ достигает наименьшего значения хотя бы в одной точке границы, т. е. при $i = 0$ или $i = N_1$, либо при $j = 0$ или $j = N_2$.

- 28.14.** Доказать, что если в обозначениях **28.13** справедливо неравенство $\Delta^h u_{i,j} \geq 0$ при всех $1 \leq i < N_1$, $1 \leq j < N_2$, то функция $u_{i,j}$ достигает наибольшего значения хотя бы в одной точке границы.
- 28.15.** Пусть в прямоугольной области, покрытой равномерной сеткой с шагом h , разностный аналог оператора Лапласа

$$\Delta^h u_{i,j} \equiv (\Lambda_1 + \Lambda_2) u_{i,j}, \quad 1 \leq i < N_1, \quad 1 \leq j < N_2,$$

определен на сеточных функциях $u_{i,j} \equiv u_h$, обращающихся в нуль на границе, т. е. при $i = 0, N_1$ и при $j = 0, N_2$. Доказать, что оператор $(-\Delta^h)$ является симметричным, положительно определенным и для него справедливы оценки

$c_1(u_h, u_h) \leq (-\Delta^h u_h, u_h) \leq c_2(u_h, u_h)$, в которых постоянная $c_1 > 0$ не зависит от сеточного параметра h , а постоянная c_2 может быть выбрана равной $8/h^2$.

- 28.16.** Показать, что для решения методом Гаусса разностного уравнения Пуассона $\Delta^h u_h = f_h$ с однородными условиями Дирихле на границе (см. **28.15**) при естественной нумерации неизвестных (по строкам, т. е. $u_{1,1}, u_{2,1}, \dots, \dots, u_{N_1-1,1}, u_{1,2}, \dots, u_{N_1-1, N_2-1}$) требуется порядка $O(h^{-4})$ арифметических действий.
- 28.17.** Упорядочить неизвестные задачи **28.16** так, чтобы количество арифметических действий при решении методом Гаусса имело порядок $O(h^{-3})$.
- 28.18.** Пусть в единичном квадрате D задана регулярная («северо-восточная») триангуляция с шагом h и в качестве базисных функций используются кусочно-линейные над треугольниками функции. Записать систему уравнений метода Рунге (конечных элементов) для решения уравнения Пуассона с однородными краевыми условиями.
- 28.19.** Пусть единичный квадрат D разбит на элементарные квадраты со стороной h и в качестве базисных функций используются билинейные на квадратах функции. Записать систему уравнений метода Рунге (конечных элементов) для решения уравнения Пуассона с однородными краевыми условиями.
- 28.20.** Построить аппроксимацию порядка $O(h^2)$ в точке (x, y) для смешанной производной $u_{xy}(x, y)$ на семиточечном шаблоне:
- 1) $(x, y), (x \pm h, y), (x, y \pm h), (x - h, y + h), (x + h, y - h)$;
 - 2) $(x, y), (x \pm h, y), (x, y \pm h), (x + h, y + h), (x - h, y - h)$.

Решение интегральных уравнений

Обозначим через G_b интегральный оператор

$$G_b y(x) = \int_a^b K(x, s)y(s) ds,$$

где заданная функция $K(x, s)$ называется *ядром*, а верхний предел интегрирования b в общем случае может быть переменным. Типичными примерами интегральных уравнений являются *уравнения Фредгольма и Вольтерры*, каждое из которых может быть первого или второго рода. В уравнениях Фредгольма b является постоянной заданной величиной, а уравнения первого и второго рода имеют вид

$$G_b y(x) = f(x) \quad \text{и} \quad y(x) - \lambda G_b y(x) = f(x).$$

В уравнение второго рода входит дополнительный числовой параметр λ , который может быть задан или подлежит определению в зависимости от постановки задачи. В уравнениях Вольтерры верхний предел интегрирования совпадает с текущим значением независимой переменной: $b = x$, поэтому уравнения первого и второго рода имеют вид

$$G_x y(x) = f(x) \quad \text{и} \quad y(x) - \lambda G_x y(x) = f(x).$$

Функция $f(x)$ задана, а $y(x)$ подлежит определению в уравнениях всех типов.

§ 29. Методы замены интеграла и ядра

Наиболее распространенной модельной постановкой является следующая задача для уравнения Фредгольма второго рода: на множествах $[a, b]$ и $[a, b] \times [a, b]$ заданы квадратично интегрируемые функции $f(x)$ и $K(x, s)$, требуется для заданного значения параметра λ определить квадратично интегрируемую функцию $y(x)$, удовлетворяющую интегральному уравнению

$$y(x) - \lambda \int_a^b K(x, s)y(s) ds = f(x). \quad (1)$$

Предполагается, что (1) имеет единственное решение. Это справедливо, например, если выполнено условие

$$|\lambda| < \left(\int_a^b \int_a^b |K(x, s)|^2 dx ds \right)^{-1/2}.$$

Метод замены интеграла. Применим для определения приближенного значения интеграла $I(\varphi) = \int_a^b \varphi(s) ds$ квадратурную формулу $S_n(\varphi) = \sum_{j=1}^n c_j \varphi(s_j)$. Тогда дискретный аналог уравнения (1) в узлах $x_i = s_i$, $1 \leq i \leq n$, примет вид

$$y_i - \lambda \sum_{j=1}^n c_j K(x_i, s_j) y_j = f(x_i), \quad 1 \leq i \leq n. \quad (2)$$

Это система линейных алгебраических уравнений $Ay = f$ относительно неизвестных y_1, y_2, \dots, y_n , являющихся приближениями к точным значениям $y(x_1), y(x_2), \dots, y(x_n)$.

Метод замены ядра. Если в уравнении (1) ядро $K(x, s)$ вырождено, т. е. представимо как $K(x, s) = \sum_{i=1}^p A_i(x) B_i(s)$, где $\{A_i(x)\}_{i=1}^p$ и $\{B_i(x)\}_{i=1}^p$ — системы линейно независимых на отрезке $[a, b]$ функций, то (1) можно записать в виде

$$y(x) - \lambda \sum_{i=1}^p A_i(x) \int_a^b B_i(s) y(s) ds = f(x). \quad (3)$$

Решение уравнения такой структуры удобно представить формулой

$$y(x) = f(x) + \lambda \sum_{i=1}^p D_i A_i(x),$$

где D_i — некоторые постоянные, подлежащие определению. В результате подстановки в уравнение (3) формулы для решения $y(x)$ и сокращения на λ получим

$$\sum_{i=1}^p D_i A_i(x) - \lambda \sum_{i=1}^p A_i(x) \sum_{j=1}^p D_j \int_a^b A_j(s) B_i(s) ds = \sum_{i=1}^p A_i(x) \int_a^b f(s) B_i(s) ds.$$

Обозначим

$$f_i = \int_a^b f(s)B_i(s) ds, \quad a_{ij} = \int_a^b A_j(s)B_i(s) ds$$

и на основании линейной независимости функций $\{A_i(x)\}_{i=1}^p$ перепишем полученное равенство в виде

$$D_i - \lambda \sum_{j=1}^p a_{ij}D_j = f_i, \quad 1 \leq i \leq n. \quad (4)$$

Если определитель системы (4) отличен от нуля, то система имеет единственное решение D_1, \dots, D_p , и решение интегрального уравнения $y(x)$ будет найдено в явном виде. Если при заданном λ определитель (4) равен нулю, то λ является *характеристическим числом* уравнения (ядра $K(x, s)$). В этом случае, находя все линейно независимые решения соответствующей однородной системы, можно записать в явном виде собственные функции ядра $K(x, s)$, отвечающие этому характеристическому числу λ .

Способы построения вырожденных ядер, близких к данному, могут быть самыми различными. Например, $K(x, s)$ можно приближать частичными суммами степенного или двойного тригонометрического ряда, если оно разлагается в соответствующий равномерно сходящийся в прямоугольнике $a \leq x, s \leq b$ ряд, или приближать ядро алгебраическими или тригонометрическими интерполяционными многочленами.

Найти приближенное решение интегрального уравнения методом замены интеграла квадратурной формулой Симпсона и оценить его погрешность.

29.1. $y(x) + \int_0^1 x(e^{xs} - 1)y(s) ds = e^x - x.$

29.2. $y(x) = \frac{5}{6}x + \frac{1}{2} \int_0^1 xsy(s) ds.$ **29.3.** $y(x) = e^{-x} + \frac{1}{2} \int_0^1 xe^s y(s) ds.$

29.4. Найти в узлах $a = x_1 < x_2 < \dots < x_n = b$ приближенное решение уравнения Вольтерры второго рода

$$y(x) - \lambda \int_a^x K(x, s) y(s) ds = f(x)$$

методом замены интеграла составной квадратурной формулой трапеций.

- 29.5.** Найти в точках $x=0, 1/2, 1$ приближенное решение интегрального уравнения

$$y(x) = e^{-x} + \int_0^x e^{-(x-s)} y(s) ds$$

методом замены интеграла составной квадратурной формулой трапеций.

- 29.6.** Найти в точках $x=0, 1/2, 1$ приближенное решение интегрального уравнения

$$y(x) = e^x + \int_0^x e^{x-s} y(s) ds$$

методом замены интеграла составной квадратурной формулой трапеций.

- 29.7.** Пусть $K(x, s) \equiv K = \text{const} > 0$. Записать в явном виде элементы матрицы в системе (2), если в качестве квадратурной формулы используется составная формула прямоугольников с узлами в центральных точках подотрезков равномерного разбиения.

- 29.8.** Исследовать сходимость метода простой итерации для решения системы линейных уравнений, полученной в **29.7**.

- 29.9.** Пусть в задаче (2) ядро $K(x, s) \equiv K = \text{const} > 0$, $|\lambda| K(b-a) < \alpha < 1$. Оценить погрешность приближенного решения при замене интеграла составной квадратурной формулой прямоугольников с узлами в центральных точках подотрезков равномерного разбиения.

- 29.10.** Показать, что для решения системы уравнений, полученной в **29.7**, метод Гаусса является устойчивым, т. е. все элементы матриц, возникающих в процессе треугольной LR — факторизации, равномерно ограничены.

- 29.11.** Найти приближенное решение интегрального уравнения

$$y(x) + \int_0^1 x(e^{xs} - 1)y(s) ds = e^x - x$$

с помощью замены ядра вырожденным ядром $H(x, s) = x^2s + x^3s^2/2 + x^4s^3/6$, являющимся суммой первых трех членов разложения $K(x, s)$ в ряд Тейлора.

29.12. Найти приближенное решение интегрального уравнения с вырожденным ядром

$$y(x) = -x/6 - 1/2 + \int_0^1 (1 + 2xs)y(s) ds.$$

§ 30. Проекционные методы

Метод наименьших квадратов. Будем искать приближенное решение $z(x)$ интегрального уравнения (29.1) в виде

$$z(x) = \sum_{j=1}^n c_j \varphi_j(x),$$

где $\{\varphi_j(x)\}_{j=1}^n$ — известные линейно независимые функции, которые часто называют *координатными*. Определим невязку уравнения

$$Rz(x) = z(x) - \lambda \int_a^b K(x, s)z(s) ds - f(x). \quad (1)$$

Неизвестные постоянные коэффициенты c_1, \dots, c_n найдем из условия минимума интеграла $J = \int_a^b [Rz(x)]^2 dx$, т. е. из условий

$$\frac{\partial J}{\partial c_i} = 0, \quad 1 \leq i \leq n.$$

Используя выражение для невязки (1), получим для отыскания коэффициентов систему линейных алгебраических уравнений $A\mathbf{c} = \mathbf{f}$, где

$$a_{ij} = \int_a^b \left[\varphi_j(x) - \lambda \int_a^b K(x, s)\varphi_j(s) ds \right] \left[\varphi_i(x) - \lambda \int_a^b K(x, s)\varphi_i(s) ds \right] dx,$$

$$f_i = \int_a^b f(x) \left[\varphi_i(x) - \lambda \int_a^b K(x, s)\varphi_i(s) ds \right] dx, \quad 1 \leq i, j \leq n.$$

Матрица A симметрична, а также положительно определена, если λ не является характеристическим числом исходного уравнения. В этом случае алгебраическая система имеет единственное решение, а $z(x)$ стремится к $y(x)$ с ростом n .

Метод Петрова—Галёркина. Пусть имеются две различные системы линейно независимых функций. Будем искать приближенное решение $z(x)$ интегрального уравнения (29.1) в виде

$$z(x) = f(x) + \sum_{j=1}^n c_j \varphi_j(x),$$

где $\{\varphi_j(x)\}_{j=1}^n$ — первая система линейно независимых функций. Подлежащие определению коэффициенты c_1, \dots, c_n найдем из условий ортогональности невязки (1) каждой функции из второй системы линейно независимых функций $\{\psi_i(x)\}_{i=1}^n$:

$$\int_a^b Rz(x) \psi_i(x) dx = 0, \quad 1 \leq i \leq n.$$

Эти условия представляют собой систему линейных алгебраических уравнений $A\mathbf{c} = \mathbf{f}$ для отыскания неизвестных постоянных коэффициентов, где

$$a_{ij} = \int_a^b \varphi_j(x) \psi_i(x) dx - \lambda \int_a^b \psi_i(x) \int_a^b K(x, s) \varphi_j(s) ds dx,$$

$$f_i = \lambda \int_a^b \psi_i(x) \int_a^b K(x, s) f(s) ds dx, \quad 1 \leq i, j \leq n.$$

Метод Бубнова—Галёркина. Данный метод является частным случаем метода Петрова—Галёркина, когда обе системы функций совпадают: $\varphi_j(x) = \psi_j(x)$, $1 \leq j \leq n$. Иногда удобно для приближенного решения использовать представление

$$z(x) = \sum_{j=1}^n c_j \varphi_j(x)$$

(без учета правой части $f(x)$), например, при определении характеристических чисел уравнения.

Метод моментов. Пусть задана ортонормированная система координатных функций $\{\varphi_j(x)\}_{j=1}^{\infty}$. Рассмотрим вспомогательные функции

$$u_j(x) = \int_a^b K(x, s)\varphi_j(s) ds, \quad j = 1, 2, \dots$$

и построим вырожденное ядро $H_n(x, s) = \sum_{j=1}^n u_j(s)\varphi_j(x)$. Определим приближенное решение по методу моментов как точное решение $z(x)$ интегрального уравнения

$$z(x) - \lambda \int_a^b H_n(x, s)z(s) ds = f(x).$$

Метод коллокации. Будем искать приближенное решение $z(x)$ интегрального уравнения (29.1) в виде

$$z(x) = \sum_{j=1}^n c_j \varphi_j(x),$$

где $\{\varphi_j(x)\}_{j=1}^n$ — известные линейно независимые функции. Искомые коэффициенты определим из условия обращения в нуль невязки $Rz(x)$ в заданных точках (точках коллокации) отрезка $[a, b]$.

30.1. Найти приближенное решение интегрального уравнения

$$y(x) = x + \int_{-1}^1 xsy(s) ds$$

методом наименьших квадратов, используя в качестве координатных функций $\varphi_1(x) = 1$, $\varphi_2(x) = x$.

30.2. Найти методом Петрова — Галёркина приближенное решение интегрального уравнения

$$y(x) = 1 + \int_{-1}^1 (xs + x^2)y(s) ds,$$

используя в качестве первой системы функций $\varphi_1(x) = x$, $\varphi_2(x) = x^2$, а в качестве второй системы $\psi_1(x) = 1$, $\psi_2(x) = x$.

- 30.3.** Найти методом Бубнова—Галёркина приближенное решение интегрального уравнения

$$y(x) = 1 + \int_{-1}^1 (xs + x^2)y(s) ds,$$

используя в качестве системы функций $\varphi_1(x) = x$, $\varphi_2(x) = x^2$.

- 30.4.** Найти методом Бубнова—Галёркина два младших характеристических числа и соответствующие им собственные функции однородного интегрального уравнения

$$y(x) = \lambda \int_0^1 K(x, s)y(s) ds,$$

где

$$K(x, s) = \begin{cases} x(1-s) & \text{при } 0 \leq x \leq s \leq 1, \\ s(1-x) & \text{при } 0 \leq s \leq x \leq 1, \end{cases}$$

если координатные функции имеют вид: $\varphi_1(x) = 1$, $\varphi_2(x) = x(1-x)$, $\varphi_3(x) = x(1-x)(1-2x)$.

- 30.5.** Найти методом Бубнова—Галёркина два младших характеристических числа и соответствующие им собственные функции однородного интегрального уравнения

$$y(x) = \lambda \int_0^\pi K(x, s)y(s) ds,$$

с вырожденным ядром $K(x, s) = \cos^2 x \cos 2s + \cos 3x \cos^3 s$.

- 30.6.** Найти приближенное решение интегрального уравнения

$$y(x) = 1 + \int_{-1}^1 (xs + x^2)y(s) ds$$

методом Бубнова—Галёркина, используя в качестве координатной системы функций три первых многочлена Лежандра.

- 30.7.** Найти приближенное решение интегрального уравнения

$$y(x) = 1 + \int_{-1}^1 (xs + x^2)y(s) ds$$

методом моментов, используя в качестве координатных функций два первых многочлена Лежандра.

30.8. Найти приближенное решение интегрального уравнения

$$y(x) = 1 + \frac{4}{3}x + \int_{-1}^1 (xs^2 - x)y(s) ds$$

методом коллокации с точками $-1, 0, 1$, используя в качестве координатной системы функций три первых многочлена Лежандра.

§ 31. Некорректные задачи

Типичным примером некорректной задачи является интегральное уравнение Фредгольма первого рода $G_b y(x) = f(x)$, т. е.

$$\int_a^b K(x, s)y(s) ds = f(x). \quad (1)$$

Под некорректностью здесь понимается либо несуществование решения $y(x)$, либо его неединственность, которая, как правило, приводит к неустойчивости решения относительно возмущения правой части. Эти эффекты проявляются (или не проявляются) в зависимости от свойств ядра и правой части уравнения.

Напомним полезные свойства ядра $K(x, s)$ интегрального оператора G_b . Пусть ядро $K(x, s)$ вещественное, симметричное и непрерывное. Тогда существует полная, ортонормированная в метрике пространства $L_2(a, b)$ система собственных функций $\{z_i(x)\}_{i=1}^{\infty}$ оператора G_b :

$$G_b z_i(x) = \lambda_i z_i(x), \quad (z_i, z_j) = \int_a^b z_i(x) z_j(x) dx = \delta_i^j.$$

При этом само ядро определяется сходящимся рядом

$$K(x, s) = \sum_{i=1}^{\infty} \lambda_i z_i(x) z_i(s) \quad \text{и} \quad \|K(x, s)\|^2 = \sum_{i=1}^{\infty} |\lambda_i|^2,$$

где все характеристические числа λ_i вещественные. Далее по умолчанию все ядра $K(x, s)$ обладают такими свойствами.

Для решения некорректных задач используют различные способы *регуляризации*. Например, в качестве приближения к решению y уравнения Фредгольма первого рода (1) рассматривают некоторое решение y_α корректного уравнения Фредгольма второго рода (29.1), зависящее от параметра регуляризации $\alpha > 0$.

Простейшая регуляризация. Этот способ заключается в том, что исходное уравнение $G_b y = f$ заменяют возмущенным уравнением $\alpha y_\alpha + G_b y_\alpha = f_\delta$, где f_δ возникает в силу незнания точной правой части f : $\|f_\delta - f\| \leq \delta$.

Метод регуляризации Тихонова. Приближенное решение y_α находят из условия минимума функционала $J(v) = \|G_b v - f_\delta\|^2 + \alpha \|v\|^2$, что приводит к уравнению Эйлера: $\alpha y_\alpha + G_b^* G_b y_\alpha = G_b^* f_\delta$, где $G_b^* \varphi(x) = \int_a^b K(s, x) \varphi(s) ds$ — интегральный оператор, сопряженный исходному.

Основным моментом в методах регуляризации является выбор параметра α и его согласование с величиной погрешности входных данных δ : необходимо, чтобы при $\delta \rightarrow 0$ выполнялось $y_\alpha \rightarrow y$.

- 31.1.** Доказать, что решение интегрального уравнения (1) существует и единственно, если все его характеристические числа отличны от нуля, а $f(x)$ — достаточно гладкая вещественная функция.
- 31.2.** Пусть характеристические числа интегрального уравнения (1) обладают следующим свойством: $\lambda_i \neq 0$ при $1 \leq i \leq p$ и $\lambda_i = 0$ при $i > p$. Показать, что при сколь угодно гладкой правой части $f(x)$ решение интегрального уравнения (1) может не существовать, а если оно существует, то не является единственным.
- 31.3.** Пусть для решения y уравнения (1) с оператором $G_b = G_b^* > 0$ справедливо неравенство $\|G_b^{-1} y\| \leq M = \text{const}$. Показать, что метод простейшей регуляризации при $\alpha = \sqrt{\delta/M}$ сходится: $y_\alpha \rightarrow y$ при $\delta \rightarrow 0$.
- 31.4.** Записать регуляризованное интегральное уравнение в методе Тихонова для некорректного уравнения (1).
- 31.5.** Пусть в задаче определения характеристических чисел уравнения $y(x) = \lambda G_b y(x)$ ядро $K(x, s)$ интегрального оператора

симметрично и положительно определено и по заданной начальной функции $\varphi_0(x)$ строится последовательность функций $\varphi_i(x) : \varphi_{i+1}(x) = G_b \varphi_i(x)$, $i = 0, 1, \dots$. Показать, что для минимального характеристического числа λ_{\min} можно получить приближения: $\lambda_{\min} \approx \|\varphi_i\|/\|\varphi_{i+1}\|$ или $\lambda_{\min} \approx \|\varphi_i\|^{-1/i}$.

- 31.6.** Пусть в задаче определения характеристических чисел уравнения $y(x) = \lambda G_b y(x)$ с симметричным ядром найдены числа S_i (i -е следы ядра $K(x, s)$). Показать, что для минимального характеристического числа λ_{\min} при больших i имеется приближение следующего вида:

$$|\lambda_{\min}| \approx \sqrt{S_{2i}/S_{2i+2}}.$$

- 31.7.** Пусть уравнение $G_b y(x) = f(x)$ имеет единственное решение и ядро $K(x, s)$ симметричное и положительно определенное. Показать, что метод простой итерации

$$\frac{y^{k+1}(x) - y^k(x)}{\tau} + \int_a^b K(x, s) y^k(s) ds = f(x)$$

сходится при $0 < \tau < 2/\lambda_{\max}$, где λ_{\max} — максимальное характеристическое число.

- 31.8.** Вычислить методом из **31.5** наименьшее характеристическое число ядра $K(x, s) = xs$, $0 \leq x, s \leq 1$, начиная с $\varphi_0(x) \equiv 1$.

Ответы, указания, решения



К главе I. Погрешность решения задачи

1.1. Каждый элемент $x \in F$ имеет вид

$$x = \pm \left(\frac{d_1}{2} + \frac{d_2}{4} + \frac{d_3}{8} \right) 2^\alpha, \quad \alpha \in \{-1, 0, 1, 2\}, \quad d_i \in \{0, 1\}$$

и $d_1 \neq 0$ для $x \neq 0$.

Зафиксируем различные значения мантисс m_i для ненулевых элементов множества:

$$\frac{1}{2}, \quad \frac{1}{2} + \frac{1}{8} = \frac{5}{8}, \quad \frac{1}{2} + \frac{1}{4} = \frac{3}{4}, \quad \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8},$$

или $m_i \in \left\{ \frac{1}{2}, \frac{5}{8}, \frac{3}{4}, \frac{7}{8} \right\}$. Далее, умножая m_i на 2^α с $\alpha \in \{-1, 0, 1, 2\}$ и добавляя знаки \pm , получим все ненулевые элементы множества F : $\pm \frac{1}{4}, \pm \frac{5}{16}, \pm \frac{3}{8}, \pm \frac{7}{16}, \pm \frac{1}{2}, \pm \frac{5}{8}, \pm \frac{3}{4}, \pm \frac{7}{8}, \pm 1, \pm \frac{5}{4}, \pm \frac{3}{2}, \pm \frac{7}{4}, \pm 2, \pm \frac{5}{2}, \pm 3, \pm \frac{7}{2}$. После добавления к ним числа *ноль* имеем искомое множество.

1.2. 1) $\frac{3}{4}$, 2) 0, 3) « ∞ » ($x > \frac{7}{2}$), 4) $\frac{5}{4}$, 5) $\frac{3}{2}$ или $\frac{7}{4}$, 6) « ∞ », 7) 0, 8) 0.

1.3. Нет (см. 1.2).

1.4. Рассмотрим вычисления по первой формуле. Так как $\sqrt{4899} = 69,992\dots$, то после округления получаем $\sqrt{4899} \approx 69,99$, $y_1 \approx 70 - 69,99 = 0,01$.

Вторая формула представляет собой результат «избавления от иррациональности в числителе» первой формулы. Последовательно вычисляя, получаем $70 + 69,99 = 139,99 \approx 140,0$, $1/140 = 0,00714285\dots$ Наконец, после последнего округления имеем $y_2 = 0,007143$.

Если провести вычисления с большим количеством разрядов, то можно убедиться, что в y_1 и y_2 все подчеркнутые цифры результата верные; однако во втором случае точность результата значительно выше. В первом случае пришлось вычитать близкие числа, что привело к эффекту *пропадания значащих цифр*, часто существенно искажающему конечный результат вычислений.

1.5. В данном случае имеется два источника погрешности: *погрешность метода* и *вычислительная погрешность*. Первая связана с неточностью формулы в правой части при отсутствии ошибок округления. Разложим $f(x+h)$ в ряд Тейлора в окрестности точки x : $f(x+h) = f(x) + hf'(x) + \frac{h^2}{2} f''(\xi)$. Отсюда имеем оценку для погрешности метода

$$\left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| \leq \frac{h}{2} M_2.$$

С другой стороны, в силу наличия ошибок округлений в вычислениях участвуют не точные значения $f(x+h)$, $f(x)$, а их приближения $f^*(x+h)$, $f^*(x)$ с заданной абсолютной погрешностью Δ . Поэтому полная погрешность равна

$$R_1(h, \Delta) = \left| \frac{f^*(x+h) - f^*(x)}{h} - f'(x) \right|.$$

Добавляя в числитель дроби $\pm f(x+h)$ и $\pm f(x)$, получим

$$\begin{aligned} R_1(h, \Delta) &= \\ &= \left| \frac{f^*(x+h) - f(x+h)}{h} - \frac{f^*(x) - f(x)}{h} + \left(\frac{f(x+h) - f(x)}{h} - f'(x) \right) \right|. \end{aligned}$$

Оценка вычислительной погрешности для каждого из двух первых слагаемых имеет вид Δ/h , а погрешность метода в предположении ограниченности второй производной $|f''(\xi)| \leq M_2$ равна $h M_2/2$. Следовательно, оценка полной погрешности

$$R_1(h, \Delta) \leq 2\Delta/h + h M_2/2.$$

Для определения значения h_0 , при котором минимизируется найденная оценка, следует приравнять к нулю производную по h от правой части полученного выражения. Окончательно получим:

$$h_0 = 2\sqrt{\Delta/M_2} \quad \text{и} \quad R_1(h_0, \Delta) = 2\sqrt{\Delta M_2}.$$

1.6. Используя аксиому

$$fl(a+b) = (a+b)(1+\eta), \quad |\eta| \leq \frac{1}{2} p^{1-t},$$

имеем

$$\begin{aligned} fl(S) &= (\dots((x_1 + x_2)(1 + \eta_2) + x_3)(1 + \eta_3) + \dots + x_n)(1 + \eta_n) = \\ &= (x_1 + x_2) \prod_{j=1}^{n-1} (1 + \eta_{j+1}) + x_3 \prod_{j=2}^{n-1} (1 + \eta_{j+1}) + \dots + x_n \prod_{j=n-1}^{n-1} (1 + \eta_{j+1}). \end{aligned}$$

Перепишем полученное выражение в виде $fl(S) = \sum_{j=1}^n x_j (1 + E_j)$, где для E_j справедливы оценки

$$|E_1| = \frac{n-1}{2} p^{1-t} + O\left(p^{2(1-t)}\right),$$

$$|E_i| = \left| \prod_{j=i-1}^{n-1} (1 + \eta_{j+1}) \right| = \frac{n+1-i}{2} p^{1-t} + O\left(p^{2(1-t)}\right) \text{ при } 2 \leq i \leq n.$$

Найденное представление означает, что суммирование чисел на компьютере в режиме с плавающей запятой эквивалентно точному суммированию с относительным возмущением E_j в слагаемом x_j . При этом относительные возмущения неодинаковы: они максимальны в первых слагаемых и минимальны в последних. Абсолютная погрешность Δ вычисления суммы равна $\Delta = \sum_{j=1}^n |x_j| |E_j|$.

Оценки E_j не зависят от x_j , поэтому в общем случае погрешность Δ будет наименьшей, если числа суммировать в порядке возрастания их абсолютных значений, начиная с наименьшего.

1.7. Следует использовать второй алгоритм (см. решение **1.6**).

1.8. Воспользуемся оценками величин E_j из решения **1.6**:

$$|E_1| = \frac{n-1}{2} p^{1-t} + O\left(p^{2(1-t)}\right),$$

$$|E_i| = \left| \prod_{j=i-1}^{n-1} (1 + \eta_{j+1}) \right| = \frac{n+1-i}{2} p^{1-t} + O\left(p^{2(1-t)}\right), \quad 2 \leq i \leq n.$$

Из них следует, что $|E_1/E_n| \approx n$, т. е. первое слагаемое вносит возмущение примерно в n раз большее, чем последнее. Неравноправие слагаемых объясняется тем, что в образовании погрешностей каждое слагаемое участвует столько раз, сколько суммируются зависящие от него частичные суммы.

Влияние всех слагаемых можно уравнивать с помощью следующего приема. Пусть для простоты количество слагаемых равно $n = 2^k$. На первом этапе разобьем близкие слагаемые x_j на пары и сложим каждую из них. При этом в каждое слагаемое вносится относительное возмущение одного порядка. Далее будем складывать уже полученные суммы. Для этого повторяем процесс разбиения и попарного суммирования до тех пор, пока получающиеся суммы не превратятся в одно число (степень двойки 2^k нужна только здесь). Абсолютная погрешность по-прежнему имеет вид

$\Delta = \sum_{j=1}^n |x_j| |\tilde{E}_j|$, но теперь для всех \tilde{E}_j справедлива оценка

$$\left| \tilde{E}_j \right| = \frac{1 + \log_2 n}{2} p^{1-t} + O\left(p^{2(1-t)}\right), \quad 1 \leq j \leq n.$$

Таким образом, меняя только порядок суммирования, можно уменьшить вычислительную погрешность примерно в $n/\log_2 n$ раз.

1.10. Воспользоваться решением **1.6** с учетом знаконеопределенности a_i и с точностью до слагаемых $O(\eta^2)$ получить

$$|P_n(1) - P_n^*(1)| \leq n \eta (|a_0| + |a_1| + \dots + |a_n|).$$

1.11. С точностью до слагаемых $O(\eta^2)$

$$|S - S^*| \leq n \eta \|x\|_2 \|y\|_2, \quad \|z\|_2^2 = \sum_{j=1}^n z_j^2.$$

1.12. Пусть в результате округления значения I_1 получено значение I_1^* , использование которого приводит к величинам $I_n^* = 1 - n I_{n-1}^*$. Для погрешности $\Delta_n = I_n - I_n^*$ имеем соотношение $\Delta_n = -n \Delta_{n-1}$, откуда следует $\Delta_n = (-1)^{n+1} n! \Delta_1$. Полученная формула гарантирует факториальный рост погрешности и ее знакопеременность. Поскольку точные значения интегралов удовлетворяют неравенству $0 < I_n < \int_0^1 x^n dx = 1/(n+1)$, то, начиная с некоторого n , модуль погрешности Δ_n станет больше результата I_n . Поэтому данный алгоритм является *неустойчивым* и пользоваться им нельзя.

1.13. Формулы, приведенные в условии, являются численными алгоритмами решения задачи Коши для уравнения $y'' = f(x)$. Рассмотрим модельную задачу с постоянной M

$$y'' = M, \quad y(0) = y'(0) = 0,$$

имеющую точное решение $y(x) = x^2 M/2$. Введем сетку $x_n = n h$ с шагом $h > 0$ и будем искать приближенное решение по формуле

$$y_{n+1} = 2y_n - y_{n-1} + h^2 M, \quad n = 1, 2, \dots, \quad y_0 = 0, \quad y_1 = h^2 M/2.$$

При отсутствии ошибок округлений получим $y_n = (n h)^2 M/2$, т. е. проекцию точного решения на сетку.

Вычисления с погрешностями приводят к соотношениям

$$\begin{aligned} y_0^* &= 0, \quad y_1^* = h^2 M/2 + \Delta_1, \\ y_{n+1}^* &= 2y_n^* - y_{n-1}^* + h^2 M + \Delta_{n+1}, \quad n = 1, 2, \dots \end{aligned}$$

Отсюда для погрешности $r_n = y_n^* - y_n$ имеем

$$r_{n+1} = 2r_n - r_{n-1} + \Delta_{n+1}, \quad n = 1, 2, \dots, \quad r_0 = 0, \quad r_1 = \Delta_1.$$

Для простоты вычислений предположим, что все Δ_n постоянны и равны Δ , тогда $r_n = \Delta(n^2 + n)/2$. Сопоставляя точное решение y_n и погрешность, приходим к относительной погрешности порядка $\Delta/(Mh^2)$. Требование малости этой величины налагает ограничение на шаг интегрирования h снизу, так как обычно $\Delta \sim p^{1-t}$, т. е. имеет порядок машинной точности.

Аналогичные рассуждения для второго способа расчетов дают оценку относительной погрешности порядка $\Delta/(Mh)$, т. е. этот способ приводит к более слабым ограничениям на h при том же Δ . Таким образом, использование формул

$$\frac{z_{n+1} - z_n}{h} = f_n, \quad \frac{y_n - y_{n-1}}{h} = z_n,$$

как правило, дает меньшую вычислительную погрешность.

1.14. Приближенное значение равно $1/2$.

2.1. Используя определения величин y_1 и y_2 , выражение для $A(y^*)$ перепишем в виде

$$A(y^*) = \sup_{y(x_1, x_2, \dots, x_n) \in [y_1, y_2]} |y(x_1, x_2, \dots, x_n) - y^*|,$$

при этом $A(y_1) = A(y_2) = y_2 - y_1$. Обозначим $A = y_2 - y_1$. Заметим, что для $y^* \notin [y_1, y_2]$ справедливо неравенство $A(y^*) > A$. Так как нас интересует минимальное значение величины $A(y^*)$, то достаточно проанализировать только случай $y^* \in [y_1, y_2]$. Введем для y^* параметризацию $y^* = \alpha y_1 + (1 - \alpha) y_2$ с $\alpha \in [0, 1]$ и рассмотрим предельную абсолютную погрешность

$$\begin{aligned} A(y^*) &= \sup_{y \in [y_1, y_2]} |y - [\alpha y_1 + (1 - \alpha) y_2]| = \\ &= \max\{\alpha A(y_1), (1 - \alpha) A(y_2)\} = A \max\{\alpha, 1 - \alpha\}. \end{aligned}$$

Минимум величины $\max\{\alpha, 1 - \alpha\}$ равен $1/2$ и достигается при $\alpha = 1/2$, т. е. минимум $A(y^*)$ имеет место при $y^* = \frac{y_1 + y_2}{2}$.

2.2. Если известны оценки $|x_j - x_j^*| \leq \Delta(x_j^*)$, $j = 1, 2$, то можно определить область

$$G = \{(x_1, x_2) : x_j^* - \Delta(x_j^*) \leq x_j \leq x_j^* + \Delta(x_j^*), j = 1, 2\}.$$

Предельные абсолютные погрешности для функций $y_{\pm} = x_1 \pm x_2$ в этой области имеют вид

$$\begin{aligned} A(y^*) &= \sup_{(x_1, x_2) \in G} |y_{\pm} - y_{\pm}^*| = \sup_{(x_1, x_2) \in G} |(x_1 \pm x_2) - (x_1^* \pm x_2^*)| \leq \\ &\leq \sum_{j=1}^2 \sup_{x_j} |x_j - x_j^*| = \Delta(x_1^*) + \Delta(x_2^*). \end{aligned}$$

2.3. Если известны оценки $|x_j - x_j^*|/|x_j^*| \leq \delta(x_j^*)$, $j = 1, 2$, то можно определить область

$$G = \{(x_1, x_2) : x_j^* - \Delta(x_j^*) \leq x_j \leq x_j^* + \Delta(x_j^*), j = 1, 2\},$$

где $\Delta(x_j^*) = |x_j^*| \delta(x_j^*)$. Рассмотрим в этой области функцию $y = x_1 x_2$ и оценим ее предельную относительную погрешность:

$$\begin{aligned} R(y^*) &= \frac{A(y^*)}{|y^*|} = \frac{1}{|x_1^* x_2^*|} \sup_{(x_1, x_2) \in G} |x_1 x_2 - x_1^* x_2^*| \leq \\ &\leq \frac{1}{|x_1^* x_2^*|} (\Delta(x_1^*) x_2^* + \Delta(x_2^*) x_1^* + \Delta(x_1^*) \Delta(x_2^*)). \end{aligned}$$

Отбрасывая члены второго порядка малости, получим

$$R(y^*) \leq \frac{\Delta(x_1^*)}{|x_1^*|} + \frac{\Delta(x_2^*)}{|x_2^*|} = \delta(x_1^*) + \delta(x_2^*).$$

Аналогично исследуется случай функции $y = x_1/x_2$.

2.4. Используем формулу конечных приращений Лагранжа

$$y(x_1, x_2, \dots, x_n) - y^* = \sum_{j=1}^n b_j(\theta)(x_j - x_j^*), \text{ где}$$

$$b_j(\theta) = \left. \frac{\partial y(x_1, x_2, \dots, x_n)}{\partial x_j} \right|_{\mathbf{x}=\mathbf{x}(\theta)}, \quad \mathbf{x}(\theta) = \mathbf{x}^* + \theta(\mathbf{x} - \mathbf{x}^*), \quad \theta \in [0, 1].$$

Отсюда следует $A(y^*) \leq A_{\text{sup}}(y^*)$, так как $|b_j(\theta)| \leq B_j$.

В силу непрерывности производных $\partial y/\partial x_j$, справедливо представление $B_j = |b_j(0)| + o(1)$ при $\rho \rightarrow 0$. Так как $b_j \equiv b_j(0)$, то величина $A_{\text{sup}}(y^*)$ может быть записана в виде $A_{\text{sup}}(y^*) = A_{\text{lin}}(y^*) + o(\rho)$.

На практике часто используют, вообще говоря, неверную «оценку» $|y(x_1, x_2, \dots, x_n) - y^*| \leq A_{\text{lin}}(y^*)$, называемую *линейной оценкой погрешности*. Величина $A_{\text{lin}}(y^*)$ вычисляется значительно проще, чем $A_{\text{sup}}(y^*)$ или $A(y^*)$, но не следует забывать о требуемой малости ρ .

2.5. 1) Здесь $y^* = 1$, $\frac{\partial y}{\partial x} = 10 \cdot x^9$, $b(0) = 10$.

Пусть $\Delta(x^*) = 0,001$, тогда в обозначениях **2.4**

$$B = \sup_{|x-1| \leq 0,001} |10 \cdot x^9| = 10,09 \dots,$$

$$A_{\text{sup}}(y^*) = B \Delta(x^*) = 0,01009 \dots,$$

$$A_{\text{lin}}(y^*) = |b(0)| \Delta(x^*) = 0,01,$$

$$A(y^*) = \sup_{|x-1| \leq 0,001} |x^{10} - 1| = 1,001^{10} - 1 = 0,010045 \dots$$

В этом случае верхняя, линейная и предельно точная оценки отличаются несущественно.

2) Здесь

$$B = \sup_{|x-1| \leq 0,1} |10 \cdot x^9| = 10 \cdot (1,1)^{10} = 23, \dots,$$

$$A_{\text{sup}}(y^*) = B \Delta(x^*) = 2,3 \dots,$$

$$A_{\text{lin}}(y^*) = |b(0)| \Delta(x^*) = 1,$$

$$A(y^*) = \sup_{|x-1| \leq 0,1} |x^{10} - 1| = 1,1^{10} - 1 = 1,5 \dots$$

Различие между рассматриваемыми величинами здесь более заметно.

2.6. Дифференцируя по x_j , имеем $\frac{\partial F}{\partial y} \frac{\partial y}{\partial x_j} + \frac{\partial F}{\partial x_j} = 0$, откуда $\frac{\partial y}{\partial x_j} =$

$$= - \frac{\partial F}{\partial x_j} \left(\frac{\partial F}{\partial y} \right)^{-1}.$$

При фиксированных x_1^*, \dots, x_n^* можно найти y^* как решение нелинейного уравнения с одним неизвестным y : $F(y, x_1^*, \dots, x_n^*) = 0$. Далее вычисляем значения $b_j =$

$$= - \frac{\partial F}{\partial x_j} \left(\frac{\partial F}{\partial y} \right)^{-1} \Bigg|_{(y^*, x_1^*, \dots, x_n^*)},$$

приводящие к искомой величине

$$A_{\text{lin}}(y^*) = \sum_{j=1}^n |b_j| \Delta(x_j^*).$$

2.7. Воспользоваться решением **2.6**, где $F(y, b, c) \equiv y^2 + by + c = 0$ — неявная функция, и вычислить следующие величины:

$$b_1 = - \frac{\partial F}{\partial b} \left(\frac{\partial F}{\partial y} \right)^{-1} \Bigg|_{(y^*, b^*, c^*)} = - \frac{y^*}{2y^* + b^*}, \quad b_2 = - \frac{1}{2y^* + b^*}.$$

2.8. Пусть y^* — двукратный корень уравнения $F(y, b, c) \equiv y^2 + by + c = 0$ при $b = b^*$, $c = c^*$. Разложим F в ряд Тейлора в окрестности точки (y^*, b^*, c^*) :

$$F(y, b, c) = F(y^*, b^*, c^*) + F_y(y^*, b^*, c^*)(y - y^*) + F_b(y^*, b^*, c^*)(b - b^*) + F_c(y^*, b^*, c^*)(c - c^*) + \frac{1}{2} F_{yy}(y^*, b^*, c^*)(y - y^*)^2 + o(\rho) = 0.$$

Из условия имеем

$$F(y^*, b^*, c^*) = F_y(y^*, b^*, c^*) = 0, \quad \frac{1}{2} F_{yy}(y^*, b^*, c^*) = 1,$$

что приводит к неравенству

$$(y - y^*)^2 \leq |F_b(y^*, b^*, c^*)| |b - b^*| + |F_c(y^*, b^*, c^*)| |c - c^*| + o(\rho),$$

или $|y - y^*| = O(\sqrt{\rho})$.

- 2.9.** Рассмотрим более общее уравнение $f(y) = a$ и вычислим $a^* = f(y^*)$. При малых $y^* - y$ из равенства $f(y) - f(y^*) = a - a^*$ следует $f'(y^*)(y - y^*) \approx a - a^*$, откуда $y - y^* \approx \frac{a - a^*}{f'(y^*)} = \frac{a - f(y^*)}{f'(y^*)}$.

Заметим, что $f'(y^*) \neq 0$, так как y^* — простой корень. Полагая $a = 0$ (по условию), приходим к искомой формуле.

- 2.10.** Если взять $\lg 2 = 0,30102999566\dots$, то корни примут значения $y_1 = 1,83604425979\dots$ и $y_2 = 0,16395574020\dots$. По условию задачи это означает, что требуется найти приближение к числу $\lg 2$, обеспечивающее значения корней $y_1^* = 1,836$ и $y_2^* = 0,164$. Теперь воспользуемся решением **2.7** при $b = -2$, $\Delta(b^*) = 0$ и $c = \lg 2$. После подстановки в $A_{\text{lin}}(y^*) = \frac{|y^* \Delta(b^*) + \Delta(c^*)}{|2y^* + b^*|}$ имеем $A_{\text{lin}}(y_{1,2}^*) = \frac{\Delta(c^*)}{2\sqrt{1 - c^*}} = \Delta(c^*) \cdot 0,5980544\dots$. Из этой формулы следует: если требуется в решении получить n верных знаков, то достаточно в c^* взять также n верных знаков, так как постоянная, связывающая величины погрешностей, не превосходит единицы. Таким образом, требуется взять $\lg 2$ с четырьмя верными знаками, т. е. $\lg 2 \approx 0,301$.

Если провести аккуратные вычисления, то при $\lg 2 \approx 0,301$ получим $y_1^* = 1,83606\dots \approx 1,836$ и $y_2 = 0,16393\dots \approx 0,164$. Меньшее количество верных знаков брать нельзя, так как, например, при $\lg 2 \approx 0,30$ имеем $y_1^* = 1,83666\dots \approx 1,837$ и $y_2 = 0,16333\dots \approx 0,163$.

- 2.11.** Воспользоваться решением **2.7** и **2.8**.

К главе II. Приближение функций и производных

- 3.1.** 1) $L_3(x) = 2x^2 + x + 2$; 2) $L_3(x) = x + 2$.

- 3.2.** Определитель данной системы уравнений есть определитель Вандермонда, следовательно, задача вычисления коэффициентов искомого многочлена имеет единственное решение. Пусть узлы интерполяции принадлежат отрезку $[0, 1]$. Функции x^{n-2} и x^{n-1} при больших n на этом отрезке почти неразличимы, поэтому столбцы

матрицы $(x_1^{n-2}, \dots, x_n^{n-2})^T$ и $(x_1^{n-1}, \dots, x_n^{n-1})^T$ получатся близкими и матрица близка к вырожденной.

3.3. x^p при $p = 0, \dots, n-1$, и $x^n - \omega_n(x)$ при $p = n$.

3.4. Пусть $n = 3$, $\omega_3(x) = (x-a)(x-(a+b)/2)(x-b)$. После замены переменных $x = \frac{a+b}{2} + \frac{b-a}{2}y$ получим

$$\omega_3(y) = \left(\frac{b-a}{2}\right)^3 (y^3 - y), \quad \text{где } y \in [-1, 1].$$

Точки экстремума кубического многочлена $y^3 - y$ на отрезке $[-1, 1]$ равны, соответственно, $y_{1,2} = \pm \frac{1}{\sqrt{3}}$. Следовательно,

$$\|\omega_3(x)\| = |\omega_3(y_{1,2})| = \frac{(b-a)^3}{12\sqrt{3}}.$$

Аналогичные рассуждения для случаев $n = 2$ и $n = 4$ дают

$$\|\omega_2(x)\| = \frac{(b-a)^2}{4}, \quad \|\omega_4(x)\| = \frac{(b-a)^4}{81}.$$

3.5. 1) $p = 3$; 2) $p = 2$.

3.6. Использовать выпуклость функции $\ln x$ и представление погрешности (но не оценку погрешности!).

3.7. Поскольку $f^{(4)}(x) = \frac{4!}{(A^2-x)^5}$ и $\|\omega_4(x)\| = 1$ (см. решение **3.4**), то для оценки погрешности имеем $\|f(x) - L_4(x)\| \leq \left\| \frac{1}{(A^2-x)^5} \right\| = \frac{1}{(A^2+1)^5}$. Следовательно, $\|f(x) - L_4(x)\| \leq 10^{-5}$ при $|A| \geq 3$.

3.8. Воспользоваться представлением для вспомогательных многочленов $\Phi_i(x) = \frac{\omega_n(x)}{(x-x_i)\omega_n'(x_i)}$.

3.23. Например, функция Рунге $f(x) = 1/(25x^2 + 1)$ или $f(x) = |x|$.

3.34. По определению, вспомогательные многочлены $(n-1)$ -й степени $\Phi_i(y)$, $i = 1, \dots, n$ обладают свойством $\Phi_i(y_k) = \delta_i^k$. Положим в формуле для $\Phi_i(x)$, обладающей теми же свойствами, $x = x(y)$. Линеинное преобразование не меняет степени многочлена. Кроме того, $\Phi_i(x_k) = \Phi_i(x(y_k)) = \Phi_i(y_k) = \delta_i^k$, т. е. два многочлена $(n-1)$ -й степени совпадают в n точках. Отсюда следует равенство констант Лебега $\lambda_n^{[a,b]}$ и $\lambda_n^{[-1,1]}$.

Таким образом, величина λ_n не зависит от длины и расположения отрезка интерполяции $[a, b]$, а определяется только взаимным расположением узлов.

3.35. По определению λ_n на отрезке $[1, n]$ имеем

$$\lambda_n = \max_{x \in [1, n]} \sum_{i=1}^n \left| \prod_{\substack{j=1 \\ j \neq i}}^n \frac{x-j}{i-j} \right|.$$

Справедливы следующие соотношения:

$$\prod_{\substack{j=1 \\ j \neq i}}^n |i-j| = (i-1)!(n-i)!, \quad \prod_{j=1}^n (j-1/2) \geq \frac{n!}{2\sqrt{n}}, \quad n \geq 1,$$

первое из которых легко проверяется, а второе доказывается по индукции. Проведем с их помощью оценку снизу для λ_n :

$$\lambda_n = \max_{x \in [1, n]} \sum_{i=1}^n \frac{1}{(i-1)!(n-i)!} \prod_{\substack{j=1 \\ j \neq i}}^n |x-j| \geq \sum_{i=1}^n \frac{1}{(i-1)!(n-i)!} \prod_{\substack{j=1 \\ j \neq i}}^n \left| \frac{3}{2} - j \right|$$

(использовано неравенство $\max_{x \in [1, n]} |f(x)| \geq |f(3/2)|$). Для оценки произведения в правой части выполним преобразования:

$$\begin{aligned} \prod_{\substack{j=1 \\ j \neq i}}^n \left| \frac{3}{2} - j \right| &= \frac{1}{|i-3/2|} \prod_{j=1}^n \left| \frac{3}{2} - j \right| = \frac{1}{2|i-3/2|} \prod_{j=1}^{n-1} \left| \frac{1}{2} - j \right| \geq \\ &\geq \frac{(n-1)!}{4(n-3/2)\sqrt{n-1}}. \end{aligned}$$

Наконец, получим искомое неравенство с $K_1 = 1/8$:

$$\begin{aligned} \lambda_n &\geq \frac{1}{4(n-3/2)\sqrt{n-1}} \sum_{i=1}^n \frac{(n-1)!}{(i-1)!(n-i)!} \geq \\ &\geq \frac{1}{4n^{3/2}} \sum_{i=1}^n C_{n-1}^{i-1} = \frac{1}{8} \frac{2^n}{n^{3/2}}. \end{aligned}$$

3.37. Константа Лебега не зависит от отрезка (см. **3.34**), поэтому будем считать, что $x \in [-1, 1]$. Тогда $x_1 = -\xi$, $x_2 = 0$, $x_3 = \xi$, где ξ — произвольное число из отрезка $[\sqrt{8}/3, 1]$; $\lambda_3 = 5/4$.

4.1. 1) Следствием тригонометрического тождества

$$\cos((n+m)\eta) + \cos((n-m)\eta) = 2 \cos(n\eta) \cos(m\eta)$$

является полиномиальное тождество

$$2T_n(x)T_m(x) = T_{n+m}(x) + T_{n-m}(x), \quad n \geq m \geq 0,$$

из которого при $n = m$ получаем искомое выражение.

2) Положим $x = \cos \eta$, тогда $dx = -\sin \eta d\eta$ и

$$I_{mn} = \int_0^\pi \cos(n\eta) \cos(m\eta) d\eta = \frac{\pi}{2} (\delta_{n-m}^0 + \delta_{n+m}^0).$$

3) Так как $\frac{T'_n(x)}{n} = \frac{\sin(n \arccos x)}{\sqrt{1-x^2}}$, то при $x = \cos \eta$ имеем

$$\begin{aligned} \frac{1}{2} \left(\frac{1}{n+1} T'_{n+1}(x) - \frac{1}{n-1} T'_{n-1}(x) \right) &= \frac{\sin((n+1)\eta) - \sin((n-1)\eta)}{2 \sin \eta} = \\ &= \frac{2 \cos(n\eta) \sin \eta}{2 \sin \eta} = T_n(x); \end{aligned}$$

теперь искомое равенство справедливо с точностью до постоянной, которую легко определить, поскольку $T_n(-1) = (-1)^n$.

4) Непосредственно дифференцированием вычисляется $T''_n(x)$; напомним, что $(\arccos x)' = -(1-x^2)^{-1/2}$.

4.3. $x_m = \cos \frac{\pi(2m-1)}{2n}$, где $m = 1, \dots, n$ (все нули лежат внутри отрезка $[-1, 1]$, их ровно n).

4.4. $x_{(m)} = \cos \frac{\pi m}{n}$, $m = 0, \dots, n$ (на $[-1, 1]$ имеется $n+1$ экстремум и $T_n(x_{(m)}) = (-1)^m$).

4.5. Пусть $\|P_n(x)\| < 2^{1-n}$. Тогда в точках экстремума многочлена Чебышёва знак разности $\bar{T}_n(x) - P_n(x)$ определяется знаком $\bar{T}_n(x)$: $\text{sign}(\bar{T}_n(x_{(m)}) - P_n(x_{(m)})) = \text{sign}((-1)^m 2^{1-n} - P_n(x_{(m)})) = (-1)^m$.

При этом указанная разность является многочленом степени не выше $n-1$, но имеет n нулей, поскольку $n+1$ раз меняет знак в точках экстремума, т. е. $P_n(x) \equiv \bar{T}_n(x)$, что невозможно в силу $\|P_n\| < \|\bar{T}_n\|$. Полученное противоречие завершает доказательство.

4.7. Сделаем линейную замену переменных $x = \frac{a+b}{2} + \frac{b-a}{2}y$ для отображения отрезка $[-1, 1]$ в заданный отрезок $[a, b]$. Многочлен $\bar{T}_n(y)$ при этом преобразуется в многочлен $\bar{T}_n\left(\frac{2x-(b+a)}{b-a}\right)$ со старшим коэффициентом $\left(\frac{2}{b-a}\right)^n$. После нормировки и использования схемы доказательства из **4.5** имеем

$$\bar{T}_n^{[a,b]}(x) = (b-a)^n 2^{1-2n} T_n\left(\frac{2x-(b+a)}{b-a}\right).$$

4.8. Использовать решение **4.7**.

$$4.10. P_n^*(x) = c \left(\frac{b-a}{2} \right)^k \frac{T_n \left(\frac{2x - (a+b)}{b-a} \right)}{T_n^{(k)} \left(\frac{a+b}{a-b} \right)}.$$

$$4.11. P_n^*(x) = 2^{1-n} \left(\cos \frac{\pi}{2n} \right)^{-n} T_n \left(x \cos \frac{\pi}{2n} \right).$$

4.12. Предположив противное, т. е. допустив существование такого ξ , $|\xi| \geq 1$, что $|P_n(\xi)| > M |T_n(\xi)|$, получить противоречие, доказав, что у многочлена $Q_n(x) = \frac{P_n(\xi)}{T_n(\xi)} T_n(x) - P_n(x)$ степени не выше n как минимум $n+1$ нуль.

4.13. Воспользоваться третьим свойством из 4.1 в виде $\frac{T'_n}{n} = 2T_{n-1} + \frac{T'_{n-2}}{n-2}$, $n > 2$.

$$4.15. 1) T_{3k} \left(\frac{1}{2} \right) = (-1)^k, T_{3k \pm 1} \left(\frac{1}{2} \right) = \frac{1}{2} (-1)^k;$$

$$2) T_{3k} \left(-\frac{1}{2} \right) = 1, T_{3k \pm 1} \left(-\frac{1}{2} \right) = -\frac{1}{2}.$$

$$4.16. 1) T'_n(1) = n^2; \quad 2) T'_n(-1) = (-1)^{n+1} n^2.$$

$$4.17. p = 2.$$

$$4.18. p = 3.$$

$$4.20. P(x) = 4 \frac{T_3(x-4)}{T_3^{(2)}(-4)} = -\frac{x^3}{6} + 2x^2 - \frac{63x}{8} + \frac{61}{6}.$$

$$4.21. a_2 = -a_0 \text{ при любом } |a_0| \leq 1/2.$$

$$4.22. P(x) = \frac{5}{32} T_3(2x-3) = \frac{5}{32} (32x^3 - 144x^2 + 210x - 99).$$

$$4.23. P(x) = 4 \frac{T_3(x-2)}{T_3(-2)} = - \left(\frac{8x^3}{13} - \frac{48x^2}{13} + \frac{90x}{13} - 4 \right).$$

$$4.24. P(x) = 3 \frac{T_3(x-3)}{T_3'(-3)} = \frac{4x^3}{35} - \frac{36x^2}{35} + 3x - \frac{99}{35}.$$

$$4.27. \lambda_3 = 5/3.$$

4.29. Вычислить определитель

$$\begin{vmatrix} T_0(x) & T_1(x) \\ U_{-1}(x) & U_0(x) \end{vmatrix} = \begin{vmatrix} 1 & x \\ 0 & 1 \end{vmatrix} = 1;$$

откуда следует, что $T_n(x)$ и $U_{n-1}(x)$ — линейно независимы.

5.1. Использовать явное представление погрешности для производной многочлена Лагранжа.

5.2. Использовать разложение в ряд Тейлора.

5.3.
$$\bar{\partial}\partial f(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2},$$

$$\bar{\partial}\bar{\partial}\partial f(x) = \frac{f(x+2h) - 2f(x+h) + 2f(x-h) - f(x-2h)}{2h^3},$$

$$\bar{\partial}^2\partial^2 f(x) = \frac{f(x+2h) - 4f(x+h) + 6f(x) - 4f(x-h) + f(x-2h)}{h^4}.$$

5.4. 1) $K_2 = h^2/12$; 2) $K_3 = h^2/4$; 3) $K_4 = h^2/6$.

5.5. Решение провести по аналогии с 1.5 и получить:

1) $h_0 = 2(3\Delta/M_4)^{1/4}$ для $R_2(h, \Delta) \leq 4\Delta/h^2 + h^2M_4/12$;

2) $h_0 = 2(3\Delta/M_6)^{1/6}$ для $R_4(h, \Delta) \leq 16\Delta/h^4 + h^2M_6/6$.

5.6. 1) $a = -1/6, b = -1/2, c = 2/3$; 2) $a = 1/2, b = -2, c = 2, d = -1/2$.

5.7. Представить интеграл в виде суммы двух слагаемых, раскрывая модуль, и интегрировать по частям.

5.13. 1) $a = -3/2, b = 2, c = -1/2; h_0 = (2\Delta/A_3)^{1/3}$;

2) $a = 1/2, b = -2/3, c = 1/6; h_0 = (2\Delta/A_3)^{1/3}$;

3) $a = 3/2, b = -2, c = 1/2; h_0 = (2\Delta/A_3)^{1/3}$;

4) $a = -5/6, b = 3/2, c = -2/3; h_0 = (3\Delta/10A_3)^{1/3}$.

6.1. $Q_{50}(x) = 0$.

6.2. Пусть $g(x) = f(x) - Q_1(x)$, тогда эта функция непрерывна и выпукла вместе с $f(x)$. Пусть $M = \|g\| = \max_{a \leq x \leq b} |g(x)|$, $\{\xi_i\}$ — множество точек, в которых $g(x)$ достигает максимального значения M , а $\theta = \inf_i \{\xi_i : g(\xi_i) = M\}$. Из непрерывности $g(x)$ следует, что $g(\theta) = M$. Допустим, что $\theta > a$, тогда при достаточно малых ε точки $\theta - \varepsilon$ и $\theta + \varepsilon$ принадлежат отрезку $[a, b]$ и из выпуклости $g(x)$ следует

$$M = g(\theta) \leq \frac{g(\theta + \varepsilon) + g(\theta - \varepsilon)}{2} \leq \frac{M + g(\theta - \varepsilon)}{2} < \frac{M + M}{2} = M.$$

Полученное противоречие доказывает, что $\theta = a$, т. е. a является точкой альтернанса. Аналогично устанавливается принадлежность множеству точек альтернанса другого конца отрезка.

6.3. Введем обозначения: $Q_1(x) = c_0 + c_1x$, $L = \|f(x) - Q_1(x)\|$ и, воспользовавшись выпуклостью $f(x)$, запишем соотношения из теоремы Чебышёва:

$$f(a) - (c_0 + c_1 a) = \alpha L,$$

$$f(d) - (c_0 + c_1 d) = -\alpha L,$$

$$f(b) - (c_0 + c_1 b) = \alpha L.$$

Кроме того, поскольку d — внутренняя точка альтернанса и $f(x)$ — дифференцируема, отсюда получаем недостающее уравнение:

$$(f(x) - (c_0 + c_1 x))' \Big|_{x=d} = 0.$$

Решая систему, имеем: $Q_1(x) = 7x - 3 - \frac{7}{3} \sqrt{\frac{7}{3}}$.

6.4. $Q_1(x) = 2x/3 + 5/6$.

6.5. По определению многочлена наилучшего равномерного приближения, разность $P_{n+1}(x)$ и $Q_n(x)$ представляет собой наименее уклоняющийся на отрезке $[a, b]$ многочлен степени $(n + 1)$ со старшим коэффициентом a_{n+1} . Следовательно, $P_{n+1}(x) - Q_n(x) = a_{n+1} \bar{T}_{n+1}^{[a,b]}(x)$, где $\bar{T}_{n+1}^{[a,b]}(x)$ — приведенный многочлен Чебышёва. Отсюда имеем $Q_n(x) = P_{n+1}(x) - a_{n+1} \bar{T}_{n+1}^{[a,b]}(x)$. Отметим, что точки альтернанса определяются экстремумами многочлена $\bar{T}_{n+1}^{[a,b]}(x)$.

6.6. По определению многочлена наилучшего равномерного приближения, величина $L = \|f(x) - Q_n(x)\|$ не превосходит нормы погрешности приближения $f(x)$ интерполяционным многочленом $L_{n+1}(x)$ по узлам, являющимся нулями многочлена Чебышёва, т. е.

$$L \leq \max_{[a,b]} |f^{(n+1)}(x)| \frac{(b-a)^{n+1}}{2^{2n+1}(n+1)!} \equiv C_2.$$

С другой стороны, по теореме Чебышёва разность $f(x) - Q_n(x)$ обращается в нуль в $(n + 1)$ -й точке, которые можно рассматривать как узлы интерполяции y_1, \dots, y_{n+1} . Поэтому для погрешности справедливо представление

$$f(x) - Q_n(x) = f^{(n+1)}(\xi) \frac{\omega_{n+1}(x)}{(n+1)!},$$

где $\omega_{n+1}(x) = (x - y_1) \dots (x - y_{n+1})$ и $\xi = \xi(x) \in [a, b]$. Пусть точка x_0 такова, что $|\omega_{n+1}(x_0)| = \|\omega_{n+1}(x)\|$. Тогда

$$L \geq |f(x_0) - Q_n(x_0)| = |f^{(n+1)}(\xi(x_0))| \frac{|\omega_{n+1}(x_0)|}{(n+1)!}.$$

Поскольку $\|\omega_{n+1}(x)\| \geq (b-a)^{n+1}/2^{2n+1}$, окончательно имеем

$$L \geq \min_{[a,b]} |f^{(n+1)}(x)| \frac{(b-a)^{n+1}}{2^{2n+1}(n+1)!} \equiv C_1.$$

Таким образом, если $f^{(n+1)}(x)$ сохраняет знак и меняется не очень сильно, то разница между погрешностями приближения функции $f(x)$ многочленом наилучшего равномерного приближения и интерполяционным многочленом по нулям многочленов Чебышёва несущественна.

6.7. Пусть $Q_n(x)$ — многочлен наилучшего равномерного приближения $f(x)$ на $[-1, 1]$. Тогда $|f(x) - Q_n(x)| \leq L = \|f(x) - Q_n(x)\|$. Заменяя x на $-x$ и умножая выражение под знаком модуля на -1 , получим $|-f(-x) - (-Q_n(-x))| \leq L$, или $|f(x) - (-Q_n(-x))| \leq L$. Следовательно, $-Q_n(-x)$ также многочлен наилучшего равномерного приближения $f(x)$ на $[-1, 1]$. По теореме единственности (для непрерывных функций) имеем $Q_n(x) = -Q_n(-x)$, что и требовалось показать.

6.8. $\|\sin x - Q_{2n-1}(x)\| = \|\sin x - Q_{2n}(x)\|$; $C_{2n-1} = C_{2n} = \left(\frac{\pi}{6}\right)^{2n+1} \frac{1}{(2n+1)!}$.

6.9. $f(x) = \operatorname{sign} x$ на $[-1, 1]$, $Q_1(x) = \alpha x$, $\alpha \in [0, 2]$.

6.10. 1) $Q_1(x) = \frac{3}{4}x$; 2) $Q_3(x) = (e-1)x^2 + \frac{e}{2} - \frac{1}{2}(e-1)\ln(e-1)$;
3) $Q_3(x) = -x^2 + 7x - \frac{17}{2}$.

6.22. $Q_1(x) = x + 9/8$.

6.23. $Q_3(x) = 1/2$.

6.24. $Q_1(x) = (x+2)/3$.

6.25. $Q_1(x) = 3x/4$.

6.32. Наилучшее приближение ищется в виде $\sum_{j=1}^n a_j x^{j-1}$ с неизвестными коэффициентами a_j , которые определяются из условия минимума функционала $\int_0^1 \left(f(x) - \sum_{j=1}^n a_j x^{j-1}\right)^2 dx$. Дифференцируя функционал по a_i и приравнивания производные к нулю, получим уравнения

$$\int_0^1 \left(f(x) - \sum_{j=1}^n a_j x^{j-1}\right) x^{i-1} dx = 0, \quad i = 1, 2, \dots, n,$$

или

$$\sum_{j=1}^n \frac{a_j}{i+j-1} = \int_0^1 f(x) x^{i-1} dx, \quad i = 1, 2, \dots, n.$$

7.1. $S_1(x) = f(0)(1-x) + f(1)x$.

7.2. Пусть $x_i = ih$, $h = 1/n$, $i = 0, 1, \dots, n$; тогда линейный интерполяционный сплайн на отрезке $[x_{i-1}, x_i]$ имеет вид

$$S_1(x) = f(x_{i-1}) \frac{x_i - x}{h} + f(x_i) \frac{x - x_{i-1}}{h}.$$

Если $f(x) \in C^{(2)}[0, 1]$, то из оценки погрешности для интерполяционного многочлена Лагранжа первой степени следует

$$\max_{[x_{i-1}, x_i]} |f(x) - S_1(x)| \leq \max_{[x_{i-1}, x_i]} |f''(x)| \frac{h^2}{8}.$$

Это неравенство справедливо на каждом отрезке $[x_{i-1}, x_i]$, значит, оно справедливо на всем отрезке $[0, 1]$.

7.3. По определению, $S_3''(x)$ — линейная на каждом отрезке $[x_{i-1}, x_i]$ функция. В силу ее непрерывности в точках x_i , имеем представление:

$$S_3''(x) = M_{i-1} \frac{x_i - x}{h_i} + M_i \frac{x - x_{i-1}}{h_i}.$$

Двукратно интегрируя и учитывая условия $S_3(x_i) = f_i, S_3(x_{i-1}) = f_{i-1}$, получим аналитическое представление кубического интерполяционного сплайна на отрезке $[x_{i-1}, x_i]$:

$$S_3(x) = M_{i-1} \frac{(x_i - x)^3}{6 h_i} + M_i \frac{(x - x_{i-1})^3}{6 h_i} + \left(f_{i-1} - \frac{M_{i-1} h_i^2}{6} \right) \frac{x_i - x}{h_i} + \left(f_i - \frac{M_i h_i^2}{6} \right) \frac{x - x_{i-1}}{h_i}.$$

Теперь вычислим производную сплайна $S_3'(x)$ слева в точке x_i , воспользовавшись представлением на отрезке $[x_{i-1}, x_i]$:

$$S_3'(x_i - 0) = M_{i-1} \frac{h_i}{6} + M_i \frac{h_i}{3} + \frac{f_i - f_{i-1}}{h_i},$$

и аналогично найдем производную сплайна $S_3'(x)$ справа в точке x_i , воспользовавшись представлением на отрезке $[x_i, x_{i+1}]$:

$$S_3'(x_i + 0) = -M_i \frac{h_{i+1}}{3} - M_{i+1} \frac{h_{i+1}}{6} + \frac{f_{i+1} - f_i}{h_{i+1}}.$$

Непрерывность $S_3'(x)$ в точках $x_i, i = 1, \dots, n - 1$, т. е. $S_3'(x_i - 0) = S_3'(x_i + 0)$, порождает искомую систему из $(n - 1)$ уравнения относительно $(n + 1)$ неизвестных.

7.4. Из решения **7.3** следует, что неизвестные величины M_0, M_1, M_2 удовлетворяют уравнению

$$\frac{1}{6} M_0 + \frac{2}{3} M_1 + \frac{1}{6} M_2 = f(2) - 2f(1) + f(0).$$

При этом искомый сплайн имеет следующий вид:
на отрезке $[0, 1]$

$$S_3(x) = M_0 \frac{(1-x)^3}{6} + M_1 \frac{x^3}{6} + \left(f(0) - \frac{M_0}{6} \right) (1-x) + \left(f(1) - \frac{M_1}{6} \right) x;$$

на отрезке $[1, 2]$

$$S_3(x) = M_1 \frac{(2-x)^3}{6} + M_2 \frac{(x-1)^3}{6} + \left(f(1) - \frac{M_1}{6}\right)(2-x) + \left(f(2) - \frac{M_2}{6}\right)(x-1).$$

У построенного сплайна две степени свободы, которые фиксируются заданием M_0 и M_2 или уравнениями для них. Естественному сплайну соответствуют значения $M_0 = M_2 = 0$.

7.5. Пусть $\max_i |M_i| = |M_j|$, $1 \leq j \leq n-1$. Рассмотрим j -е уравнение системы

$$d_j = M_{j-1} \frac{h_j}{6} + M_j \frac{h_j + h_{j+1}}{3} + M_{j+1} \frac{h_{j+1}}{6},$$

из которого следует неравенство

$$|d_j| \geq |M_j| \frac{h_j + h_{j+1}}{3} - \left(|M_{j-1}| \frac{h_j}{6} + |M_{j+1}| \frac{h_{j+1}}{6} \right) \geq |M_j| \frac{h_j + h_{j+1}}{6},$$

так как $|M_{j\pm 1}| \leq |M_j|$. Оценивая левую часть неравенства сверху через $\max_i |d_i|$ и множитель в его правой части снизу как

$$\min_i \frac{h_i + h_{i+1}}{6} \geq \frac{1}{3} \min_i h_i,$$

приходим к искомому неравенству.

7.7. Пусть $x \in [x_{i-1}, x_i]$; тогда, используя аналитическое представление сплайна из **7.3**, получим

$$\begin{aligned} \max_{[x_{i-1}, x_i]} |S_3(x) - S_3^*(x)| &\leq \\ &\leq |M_{i-1} - M_{i-1}^*| \frac{h^2}{3} + |M_i - M_i^*| \frac{h^2}{3} + |f_{i-1} - f_{i-1}^*| + |f_i - f_i^*|. \end{aligned}$$

Разность $M_i - M_i^*$ удовлетворяет уравнению

$$A(M_i - M_i^*) = d_i - d_i^* = \frac{1}{h} [f_{i+1} - 2f_i + f_{i-1} - (f_{i+1}^* - 2f_i^* + f_{i-1}^*)].$$

Отсюда на основании решения **7.5** для коэффициентов естественного сплайна имеем оценку $\max_i |M_i - M_i^*| \leq \frac{12}{h^2} \varepsilon$. Поэтому справедливо неравенство

$$\max_{[x_{i-1}, x_i]} |S_3(x) - S_3^*(x)| \leq \frac{12}{h^2} \varepsilon \frac{2h^2}{3} + 2\varepsilon = 10\varepsilon.$$

Правая часть неравенства не зависит от рассматриваемого отрезка $[x_{i-1}, x_i]$, значит, оно справедливо для $x_0 \leq x \leq x_n$.

7.8. Справедливость первого утверждения следует из свойств стандартного сплайна $B(x)$: он является кусочно-кубической функцией, имеющей в точках $\pm 1, \pm 2$ непрерывные производные до второго порядка включительно (проверяется непосредственно). Линейная комбинация таких функций удовлетворяет определению кубического сплайна.

Далее рассмотрим в формуле $B_2^{(k)}(x)$ множитель $B\left(\frac{x+h}{h}\right)$ при α_{-1} . Эта функция обращается в нуль при $\left|\frac{x+h}{h}\right| \geq 2$, т. е. при $x \leq -3h$ и $x \geq h$. Аналогично множитель $B\left(\frac{x-(n+1)h}{h}\right)$ при α_{n+1} обращается в нуль при $x \leq x_n - h = 1 - h$ и $x \geq 1 + 3h$. Эти слагаемые являются крайними в сумме (первым и последним), поэтому определяют область, где $B_2^{(k)}(x) \neq 0$, а именно — отрезок $[-3h, 1 + 3h]$. Напомним, что областью определения приближаемой функции $f(x)$ является отрезок $[0, 1]$.

7.9. Построим многочлен Лагранжа первой степени для $f(x)$ по значениям f_0, f_1 :

$$L_2(x) = f_0 \frac{x_1 - x}{h} + f_1 \frac{x - x_0}{h}, \quad x_1 = x_0 + h,$$

и вычислим его значение в точке $x = x_0 - h$. Имеем $f_{-1} = L_2(x_0 - h) = 2f_0 - f_1$. Аналогично по значениям f_n, f_{n-1} строится формула $f_{n+1} = 2f_n - f_{n-1}$.

7.10. Построим многочлен Лагранжа третьей степени для $f(x)$ по значениям $f_n, f_{n-1}, f_{n-2}, f_{n-3}$. Имеем

$$L_4(x) = \sum_{i=1}^4 f_{n+1-i} \prod_{\substack{j=1 \\ j \neq i}}^4 \frac{x - x_{n+1-j}}{x_{n+1-i} - x_{n+1-j}},$$

где $x_i = x_0 + ih$; $i = n, n-1, n-2, n-3$; $h = (x_n - x_0)/n$, и вычислим значения многочлена в точках $x_n + h$ и $x_n + 2h$. Получим

$$f_{n+1} = L_4(x_n + h) = 4f_n - 6f_{n-1} + 4f_{n-2} - f_{n-3},$$

$$f_{n+2} = 10f_n - 20f_{n-1} + 15f_{n-2} - 4f_{n-3}.$$

Аналогично по значениям f_0, f_1, f_2, f_3 строятся формулы

$$f_{-2} = 10f_0 - 20f_1 + 15f_2 - 4f_3, \quad f_{-1} = 4f_0 - 6f_1 + 4f_2 - f_3.$$

7.11. Пусть $x \in [x_{k-1}, x_k]$, тогда $x = \theta h$, $k - 1 \leq \theta \leq k$. Неравенство $\left|\frac{x - ih}{h}\right| \equiv |\theta - i| < 2$ выполняется только для значений

$i = k - 2, k - 1, k, k + 1$. Так как стандартный сплайн $B(x)$ равен нулю при $|x| \geq 2$, то $B_2^{(1)}(x)$ зависит только от значений f_i , $i = k - 2, k - 1, k, k + 1$. Напомним, что для $B_2^{(1)}(x)$ коэффициенты определяются как $\alpha_i = f_i$.

Анализ для $B_2^{(2)}(x)$ проводится аналогично, только зависимость от значений в шести точках связана с другой формулой для коэффициентов: $\alpha_i = (8f_i - f_{i+1} - f_{i-1})/6$.

7.12. Покажем в качестве примера равенство $B_2^{(2)}(x_n) = f_n$. Так как $x_n = nh$, то

$$\begin{aligned} B_2^{(2)}(x_n) &= \sum_{i=-1}^{n+1} \alpha_i B\left(\frac{nh-ih}{h}\right) = \alpha_{n+1}B(-1) + \alpha_n B(0) + \alpha_{n-1}B(1) = \\ &= \frac{\alpha_{n+1} + 4\alpha_n + \alpha_{n-1}}{6} = \frac{1}{6} \left(\frac{-f_{n+2} + 8f_{n+1} - f_n}{6} + 4 \frac{-f_{n+1} + 8f_n - f_{n-1}}{6} + \right. \\ &\quad \left. + \frac{-f_n + 8f_{n-1} - f_{n-2}}{6} \right) = \frac{1}{36} (-f_{n+2} + 4f_{n+1} + 30f_n + 4f_{n-1} - f_{n-2}) = f_n. \end{aligned}$$

Для получения последнего равенства использованы выражения для f_{n+2} и f_{n+1} из решения **7.10**.

К главе III. Численное интегрирование

8.1. При вычислении интегралов использовать замену переменной

$$x = x(t) = \frac{b+a}{2} + \frac{b-a}{2}t. \text{ В частности,}$$

$$\int_a^b |\omega_n(x)| dx = \left(\frac{b-a}{2}\right)^{n+1} \int_{-1}^1 |\omega_n^0(t)| dt,$$

где $\omega_n^0(t) = \prod_{i=1}^n (t - t_i)$, а t_i являются образами узлов x_i на отрезке $[-1, 1]$.

$n=1$ — формула прямоугольников

$$S_1(f) = (b-a)f\left(\frac{a+b}{2}\right), \quad R_1 = \|f'(x)\| \frac{(b-a)^2}{4};$$

$n=2$ — формула трапеций

$$S_2(f) = \frac{b-a}{2} (f(a) + f(b)), \quad R_2 = \|f''(x)\| \frac{(b-a)^3}{12};$$

$n=3$ — формула парабол (Симпсона)

$$S_3(f) = \frac{b-a}{6} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right), \quad R_3 = \|f^{(3)}(x)\| \frac{(b-a)^4}{192}.$$

8.2. Можно сравнивать точность квадратурных формул только для функций из одного класса, поэтому необходимо получить для формулы прямоугольников другую оценку погрешности. Воспользуемся в качестве приближения к функции $f(x)$ отрезком ряда Тейлора в точке $(a+b)/2$:

$$f(x) = f\left(\frac{a+b}{2}\right) + f'\left(\frac{a+b}{2}\right)\left(x - \frac{a+b}{2}\right) + \frac{f''(\xi)}{2}\left(x - \frac{a+b}{2}\right)^2, \quad \xi \in [a, b].$$

Тогда для квадратурной формулы $\tilde{S}_1(f)$, полученной с помощью интегрирования двух первых слагаемых, справедливо равенство

$$\tilde{S}_1(f) = \int_a^b \left[f\left(\frac{a+b}{2}\right) + f'\left(\frac{a+b}{2}\right)\left(x - \frac{a+b}{2}\right) \right] dx = S_1(f).$$

Оценка погрешности для $\tilde{S}_1(f)$ принимает вид

$$\tilde{R}_1 = \frac{\|f''(x)\|}{2} \int_a^b \left(x - \frac{a+b}{2}\right)^2 dx = \|f''(x)\| \frac{(b-a)^3}{24}.$$

Следовательно, на классе функций с непрерывной второй производной формула прямоугольников имеет оценку погрешности в два раза меньше, чем формула трапеций.

Аналогичным образом можно получить известную оценку погрешности для формулы Симпсона: $\tilde{R}_3 = \|f^{(4)}(x)\| \frac{(b-a)^5}{2880}$.

8.3. В формуле для коэффициента квадратуры

$$c_{n+1-i} = a \int_{-1}^1 p(at) \prod_{j \neq n+1-i} \frac{t-t_j}{t_i-t_j} dt$$

заменить узлы на симметричные $t_{n+1-i} = -t_i$, $t_j = -t_{n+1-j}$, формально поменять индекс в произведении и использовать свойство определенного интеграла $\int_{-1}^1 g(t) dt = \int_{-1}^1 g(-t) dt$.

8.4. Проинтегрировать правую часть равенства по частям два раза или использовать формулу Тейлора с остаточным членом в интегральной форме.

8.6. 1) $N=13$; 2) $N=16$ (для этого случая полезно применить **8.4**).

8.7. Использовать **8.4**. Для второго случая дополнительно ввести функцию

$$\varphi_k(x) = \begin{cases} (x_k - x)(x_{k+1} - x) & \text{при } x \in [x_k, x_{k+1}], \\ 0 & \text{при } x \notin [x_k, x_{k+1}]. \end{cases}$$

Тогда имеют место соотношения

$$\sum_{k=0}^{N-1} \int_{x_k}^{x_{k+1}} (x_k - \xi)(x_{k+1} - \xi) f''(\xi) d\xi = \sum_{k=0}^{N-1} \int_a^b \varphi_k(\xi) f''(\xi) d\xi = \int_a^b f''(\xi) \sum_{k=0}^{N-1} \varphi_k(\xi) d\xi,$$

к последнему из которых следует применить неравенство Коши — Буняковского.

8.8. Покажем, что $\|f^{(n)}(x)\| = n!$. Для этого введем функцию $y = \operatorname{arctg} x$. Тогда $y' = f(x)$. Используя обратную функцию $x = \operatorname{tg} y$, получим

$$y' = \cos^2 y, \quad y'' = -2y' \cos y \sin y, \quad \dots$$

Эти выражения можно преобразовать к виду

$$\begin{aligned} y' &= \cos y \sin \left(y + \frac{\pi}{2} \right), \\ y'' &= \cos^2 y \sin 2 \left(y + \frac{\pi}{2} \right), \\ &\dots \\ y^{(n)} &= (n-1)! \cos^n y \sin n \left(y + \frac{\pi}{2} \right). \end{aligned}$$

Отсюда следует $\|f^{(n)}(x)\| = \|y^{(n+1)}(x)\| = n!$. Используя оценку погрешности из решения **8.2**, окончательно имеем:

$$\|f^{(4)}(x)\| \frac{1}{2880 \cdot 5^4} = \frac{1}{75000}.$$

8.9. $\|f''(x)\| \frac{1}{12 \cdot 10^2} = \frac{1}{600}$ (см. решение **8.8**).

8.10. $N \geq 45 > \left[10^2 \sqrt{\theta/12} \right] + 1, \quad \theta = \max(2, 4 \sin 1 - 2 \cos 1) < 2,29.$

8.11. $N \geq [50\sqrt{e}] + 1 = 83.$

8.16. Рассмотрим непрерывную на отрезке $[a, b]$ функцию $f(x)$ такую, что

$$\max_{[a,b]} |f(x)| = 1, \quad \int_a^b p(x) f(x) dx \geq \int_a^b p(x) dx - \varepsilon$$

и $f(x_i) = -\operatorname{sign} c_i, i = 1, 2, \dots, n$, где ε — произвольно малое число. Она строится конструктивно по заданным узлам x_i и весам c_i : $f(x) \equiv 1$ вне малых окрестностей точек x_i , внутри них — непрерывная функция $|f(x)| \leq 1$ и $f(x_i) = -\operatorname{sign} c_i$. Для такой функции справедливо неравенство

$$R_n(f) = I(f) - S_n(f) \geq \int_a^b p(x) dx + \sum_{i=1}^n |c_i| - \varepsilon,$$

и, так как $|R_n(f)| \leq \|R_n(C)\| \max_{[a,b]} |f(x)| = \|R_n(C)\|$, то имеет место оценка

$$\|R_n(C)\| \geq \int_a^b p(x) dx + \sum_{i=1}^n |c_i| - \varepsilon,$$

откуда в силу произвольности ε следует, что

$$\|R_n(C)\| \geq \int_a^b p(x) dx + \sum_{i=1}^n |c_i|.$$

Неравенство противоположного знака устанавливается просто, поэтому искомое утверждение доказано.

9.1. Сначала построить формулу на отрезке $[-1, 1]$, а затем отобразить ее на $[a, b]$.

9.2. Для формулы трапеций $|R_2(f)| \leq \frac{(b-a)^3}{8} \|f''(x)\|$, так как $m=1$.

Для формулы Симпсона $|R_3(f)| \leq \frac{(b-a)^5}{1536} \|f^{(4)}(x)\|$, так как $m=3$.

9.3. 1) $S(f) = \frac{11}{15} f(0) + \frac{49}{15} f\left(\frac{10}{7}\right)$; 2) $S(f) = \frac{1}{48} f(0) + \frac{5}{16} f\left(-\frac{4}{5}\right)$;
3) $S(f) = \frac{2}{3} f(0)$.

9.4. Представим производную $P_n(x)$ в виде

$$P_n'(x) = P_n(x) \frac{d}{dx} \ln P_n(x) = \sum_{k=1}^n \frac{P_n(x)}{x-x_k},$$

где введено обозначение

$$\begin{aligned} \frac{P_n(x)}{x-x_k} &= x^{n-1} + (a_1+x_k)x^{n-2} + (a_2+a_1x_k+x_k^2)x^{n-3} + \dots \\ &\dots + (a_{n-1}+a_{n-2}x_k+\dots+a_1x_k^{n-2}+x_k^{n-1}). \end{aligned}$$

Положим $a_0 = 1$. Тогда соотношение для производной можно записать в виде

$$\begin{aligned} \sum_{k=0}^{n-1} (n-k)a_k x^{n-k-1} &= n x^{n-1} + (na_1 + B_1) x^{n-2} + \\ &+ (na_2 + a_1 B_1 + B_2) x^{n-3} + \dots + (na_{n-1} + a_{n-2} B_1 + \dots + a_1 B_{n-2} + B_{n-1}). \end{aligned}$$

Из равенства коэффициентов при одинаковых степенях x следуют соотношения для a_1, \dots, a_{n-1} . Последнее соотношение (для a_n)

получается в результате сложения равенств

$$P_n(x_k) = \sum_{j=0}^n a_j x_k^{n-j}, \quad k=1, 2, \dots, n.$$

9.5. Имеем следующие соотношения для $f(x) = x^j$, $j=1, 2, \dots, n$:

$$I(x^j) = S_n(x^j), \quad \text{или} \quad \frac{1 - (-1)^{j+1}}{j+1} = \frac{2}{n} \sum_{k=1}^n x_k^j = \frac{2}{n} B_j,$$

где B_j определены в **9.4**. Решив эти системы, получим

$$P_2(x) = x^2 - \frac{1}{3}, \quad P_3(x) = x^3 - \frac{1}{2}x, \quad P_4(x) = x^4 - \frac{2}{3}x^2 + \frac{1}{45}.$$

9.6. Представим произвольный многочлен $P_{2n-1}(x)$ степени $2n-1$ в виде суммы многочленов Чебышёва: $P_{2n-1}(x) = \sum_{m=0}^{2n-1} a_m T_m(x)$, где $T_m(x) = \cos(m \arccos x)$, и проверим утверждение.

При $m=0$ имеем

$$I(T_0) = \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} dx = \pi, \quad S_n(T_0) = \pi.$$

При $m > 0$ выполняется свойство ортогональности $I(T_m T_0) = 0$. Для квадратурной формулы проведем преобразования

$$\begin{aligned} S_n(T_m) &= \frac{\pi}{n} \sum_{j=1}^n \cos(m \arccos x_j) = \frac{\pi}{n} \sum_{j=1}^n \cos m \frac{(2j-1)\pi}{2n} = \\ &= \frac{\pi}{2n} \sum_{j=1-n}^n \exp\left(\frac{m(2j-1)\pi i}{2n}\right). \end{aligned}$$

Далее используем формулу суммы членов геометрической прогрессии $\sum_{j=1}^n aq^{j-1} = \frac{a(q^n - 1)}{q - 1}$, $q = \exp\left(\frac{m\pi i}{n}\right)$, и окончательно для $m=1, \dots, 2n-1$ получаем

$$S_n(T_m) = \frac{\pi}{2n} \frac{\exp\left(\frac{m(2n+1)\pi i}{2n}\right) - \exp\left(\frac{m(1-2n)\pi i}{2n}\right)}{\exp\left(\frac{m\pi i}{n}\right) - 1} = 0.$$

9.8. Рассмотрим величины $I(f)$ и $S_n(f)$ для функций вида $f(x) = \exp\left(2\pi m i \frac{x}{\omega}\right)$, $m=0, 1, \dots, n$. При этом для интегралов

имеем $I(f) = \begin{cases} \omega & \text{при } m=0, \\ 0 & \text{при } m \neq 0. \end{cases}$ Используя квадратурную формулу, получаем

$$S_n(f) = \frac{\omega}{n} \sum_{j=0}^{n-1} \exp\left(2\pi m i \frac{j}{n}\right) = \begin{cases} \frac{\omega}{n} \sum_{j=0}^{n-1} 1 = \omega, & \frac{m}{n} - \text{целое,} \\ \frac{\exp(2\pi m i) - 1}{\exp(2\pi m i/n) - 1} = 0, & \frac{m}{n} - \text{не целое.} \end{cases}$$

Приведенное выражение означает, что квадратурная формула точна для всех $\sin \frac{2\pi m x}{\omega}$ и $\cos \frac{2\pi m x}{\omega}$, если $m = 0$ или $\frac{m}{n}$ не целое, т. е. точна для всех тригонометрических многочленов степени не выше $n - 1$. Из явного выражения для $S_n(f)$ следует, что эта формула также точна для функции $\sin \frac{2\pi n x}{\omega}$.

- 9.9.** Линейным невырожденным преобразованием, якобиан которого постоянен и отличен от нуля, произвольный треугольник отобразить в равнобедренный прямоугольный с вершинами $(0, 0)$, $(0, 1)$, $(1, 0)$.
- 9.10.** Линейным невырожденным преобразованием, якобиан которого постоянен и отличен от нуля, произвольный прямоугольник отобразить в квадрат с вершинами $(\pm 1, \pm 1)$, симметричный относительно начала координат.
- 9.11.** 1) $P_3(x) = x^3 - 3x^2 + \frac{5}{2}x - \frac{1}{2}$, $x_1 = 1$, $x_{2,3} = 1 \pm \frac{1}{\sqrt{2}}$, $c = \frac{2}{3}$;
 2) $P_3(x) = x^3 - \frac{3}{2}x^2 + \frac{5}{8}x - \frac{1}{16}$, $x_1 = \frac{1}{2}$, $x_{2,3} = \frac{1}{2} \pm \frac{1}{2\sqrt{2}}$, $c = \frac{1}{3}$;
 3) $x_1 = -\frac{1}{2}$, $x_{2,3} = -\frac{1}{2} \pm \frac{1}{2\sqrt{2}}$, $c = \frac{1}{3}$;
 4) $x_1 = -1$, $x_{2,3} = -1 \pm \frac{1}{\sqrt{2}}$, $c = \frac{2}{3}$.
- 9.12.** 1) $S_2(f) = \frac{1}{6}f(0) + \frac{5}{2}f\left(-\frac{8}{5}\right)$;
 2) $S_2(f) = \frac{1}{18}f(0) + \frac{4}{9}f\left(\frac{3}{4}\right)$;
 3) $S_2(f) = \frac{4(\pi-3)}{\pi^2-8}f(0) + \frac{(\pi-2)^2}{\pi^2-8}f\left(\frac{\pi^2-8}{2(\pi-2)}\right)$;
 4) $S_2(f) = \frac{26}{21}f(0) + \frac{100}{21}f\left(\frac{7}{5}\right)$.
- 9.18.** Получить систему уравнений для коэффициентов c_1 и c_2 :

$$c_1 + c_2 = \frac{1}{\alpha} \left(e^{\alpha b} - e^{\alpha a} \right), \quad c_1 a + c_2 b = \frac{1}{\alpha} \left[e^{\alpha b} \left(b - \frac{1}{\alpha} \right) - e^{\alpha a} \left(a - \frac{1}{\alpha} \right) \right].$$

10.1. $\psi_0=1, \psi_1=x, \psi_2=x^2-\frac{1}{3}, \psi_3=x^3-\frac{3}{5}x, \dots$

10.2. Если $\psi_n(x)$ имеет на $[a, b]$ только $r < n$ нулей x_1, \dots, x_r нечетной кратности, то многочлен

$$Q_{n+r}(x) = \psi_n(x) \prod_{l=1}^r (x - x_l)$$

не меняет знака на этом отрезке. Следовательно, $\int_a^b p(x)Q_{n+r}(x) dx$

отличен от нуля, что противоречит свойству ортогональности $\psi_n(x)$ любому многочлену низшей степени.

10.3. 1) $\frac{1}{2}f\left(\frac{2}{3}\right)$; 2) $(e-1)f\left(\frac{1}{e-1}\right)$.

10.4. 1) $\frac{1}{3}\left(f\left(\sqrt{\frac{3}{5}}\right) + f\left(-\sqrt{\frac{3}{5}}\right)\right)$; 2) $f\left(\sqrt{\frac{\pi^2}{4}-2}\right) + f\left(-\sqrt{\frac{\pi^2}{4}-2}\right)$.

10.5. $\frac{5}{9}f\left(-\sqrt{\frac{3}{5}}\right) + \frac{8}{9}f(0) + \frac{5}{9}f\left(\sqrt{\frac{3}{5}}\right)$.

10.6. Пусть x_1, \dots, x_n — узлы квадратуры Гаусса. Рассмотрим многочлен степени $2n-2$ вида $P_k(x) = \left(\prod_{\substack{i=1 \\ i \neq k}}^n (x - x_i)\right)^2$. Для интеграла от этого многочлена формула Гаусса дает точный результат

$$\int_a^b p(x)P_k(x) dx = \sum_{j=1}^n c_j P_k(x_j) = \sum_{\substack{j=1 \\ j \neq k}}^n c_j P_k(x_j) + c_k P_k(x_k).$$

Так как справедливо выражение $\sum_{\substack{j=1 \\ j \neq k}}^n c_j P_k(x_j) = 0$, то имеет место равенство

$$c_k = \frac{\int_a^b p(x)P_k(x) dx}{P_k(x_k)} > 0.$$

10.7. Симметрия узлов квадратуры следует из решения 10.29, а равенство коэффициентов — следствие симметрии узлов (см. 8.3).

10.10. Квадратурная формула и вычисляемый по ней интеграл определяют линейные функционалы на пространстве непрерывных функций. Поэтому здесь применима теорема Банаха о сходимости последовательности линейных операторов (необходимым и достаточным условием сходимости является выполнение следующих двух требований: 1) сходимость на множестве элементов, всюду плотном

в пространстве, где определены операторы; 2) ограниченность в совокупности норм операторов.)

Для квадратур Гаусса положительность коэффициентов гарантирует выполнение второго требования. Проверить, что оценка погрешности дает сходимость по n для произвольного алгебраического многочлена, откуда следует выполнение первого требования.

$$10.18. \quad -x_{-1} = x_1 = \sqrt{\frac{15 - 2\sqrt{30}}{35}}, \quad c_{-1} = c_1 = \frac{18 + \sqrt{30}}{36},$$

$$-x_{-2} = x_2 = \sqrt{\frac{15 + 2\sqrt{30}}{35}}, \quad c_{-2} = c_2 = \frac{18 - \sqrt{30}}{36}.$$

$$10.19. \quad \psi_3(x) = x^3 - 3x/2.$$

$$10.20. \quad \psi_3(x) = x^3 - 3x/4.$$

$$10.21. \quad \psi_3(x) = x^3 - x/2.$$

$$10.22. \quad \psi_1(x) = x - 1, \quad \psi_2(x) = x^2 - 4x + 2, \quad \psi_3(x) = x^3 - 9x^2 + 18x - 6.$$

$$10.23. \quad S_2(f) = f\left(\frac{\pi + \sqrt{\pi^2 - 8}}{2}\right) + f\left(\frac{\pi - \sqrt{\pi^2 - 8}}{2}\right).$$

$$10.24. \quad S_2(f) = \frac{2 + \sqrt{2}}{4} f(2 - \sqrt{2}) + \frac{2 - \sqrt{2}}{4} f(2 + \sqrt{2}).$$

$$10.25. \quad S_2(f) = \frac{1}{24} \left[f\left(\frac{1}{2} + \sqrt{\frac{3}{20}}\right) + f\left(\frac{1}{2} - \sqrt{\frac{3}{20}}\right) \right].$$

10.27. Пусть $P_n(x)$ — произвольный многочлен степени n со старшим коэффициентом, равным единице. Тогда $P_n(x) = \psi_n(x) + r_{n-1}(x)$, и из ортогональности $\psi_n(x)$ любому многочлену низшей степени следует

$$\|P_n(x)\|^2 = \|\psi_n(x)\|^2 + \|r_{n-1}(x)\|^2.$$

10.28. Представим многочлен $x\psi_{n-1}$ в виде $\sum_{k=0}^n \alpha_k \psi_k$, где коэффициенты α_j определяются из условий ортогональности

$$(x\psi_{n-1}, \psi_j) = \alpha_j (\psi_j, \psi_j), \quad j = 0, \dots, n.$$

При $j < n - 2$ имеем $(x\psi_{n-1}, \psi_j) = (\psi_{n-1}, x\psi_j) = (\psi_{n-1}, Q_{j+1}(x)) = 0$, т. е. все $\alpha_j = 0$ при $j < n - 2$ (здесь $Q_{j+1}(x) = x\psi_j$ — многочлен степени $j + 1$). Таким образом, $x\psi_{n-1} = \alpha_n \psi_n + \alpha_{n-1} \psi_{n-1} + \alpha_{n-2} \psi_{n-2}$, при этом $\alpha_n = 1$ в силу равенства коэффициентов при старшей степени x . Отсюда следует, что $\psi_n(x) = (x - \alpha_{n-1})\psi_{n-1} - \alpha_{n-2}\psi_{n-2}$, $b_n = -\alpha_{n-1}$. Так как $(x\psi_{n-1}, \psi_{n-2}) = (\psi_{n-1}, \psi_{n-1})$, то

$$c_n = \alpha_{n-2} = \frac{(\psi_{n-1}, \psi_{n-1})}{(\psi_{n-2}, \psi_{n-2})} > 0.$$

- 10.29.** $\psi_0(x) = 1$, $\psi_1(x) = x$. Продолжить решение по индукции, используя рекуррентное соотношение из **10.28** при $b_n = 0$, $n \geq 2$.
- 10.30.** Все корни x_k многочлена $\psi_n(x)$ положительны, а его коэффициенты выражаются через величины $B_j = \sum_{k=1}^n x_k^j$ (см. **9.4**). Доказательство также можно построить на основе теоремы Виета.
- 10.31.** Подставим $x = x_i^{(n)}$ в рекуррентное соотношение (см. **10.28**)

$$\psi_{n+1}(x) = (x - \alpha_n)\psi_n(x) - \alpha_{n-1}\psi_{n-1}(x).$$

Напомним, что здесь $\alpha_{n-1} > 0$. Тогда имеем

$$\psi_{n+1}(x_i^{(n)}) + \alpha_{n-1}\psi_{n-1}(x_i^{(n)}) = 0.$$

Пусть утверждение верно для некоторого n . Отсюда и из равенств

$$\text{sign } \psi_{n-1}(b) = 1, \quad \text{sign } \psi_{n-1}(a) = (-1)^{n-1}$$

следует, что

$$\psi_{n-1}(x_i^{(n)}) = (-1)^{n-i},$$

а знаки $\text{sign } \psi_{n+1}(x_i^{(n)}) = -\text{sign } \psi_{n-1}(x_i^{(n)})$ противоположны. Так как

$$\text{sign } \psi_{n+1}(b) = 1 \quad \text{и} \quad \text{sign } \psi_{n+1}(a) = (-1)^{n+1},$$

то $\psi_{n+1}(x)$ имеет чередующиеся знаки в последовательно расположенных точках $a, x_1^{(n)}, \dots, x_n^{(n)}, b$, что и завершает доказательство.

- 10.34.** Представить исходный интеграл в виде

$$I(f) = \int_{-1}^1 Q_1(x) dx + \int_{-1}^1 \frac{f(x) - Q_1(x)}{1 - x^2} p(x) dx$$

где $Q_1(x) = f(-1)(1-x)/2 + f(1)(1+x)/2$, $p(x) = 1 - x^2$.

Построить формулу Гаусса с $(n-2)$ узлами, соответствующую весовой функции $p(x) = 1 - x^2$:

$$\int_{-1}^1 f(x)p(x) dx \approx \sum_{j=2}^{n-1} \bar{c}_j f(x_j).$$

Показать, что квадратурная формула

$$\int_{-1}^1 Q_1(x) dx + \sum_{j=2}^{n-1} \bar{c}_j \frac{f(x_j) - Q_1(x_j)}{1 - x_j^2}$$

имеет требуемый в условии вид и $c_j = \bar{c}_j / (1 - x_j^2)$, $j = 2, \dots, n-1$.

Показать, что квадратура точна для произвольного многочлена $P_k(x)$ степени $k = 2n - 3$ учитывая, что $P_{2n-3}(x)/(1 - x^2)$ есть многочлен степени $2n - 5$.

11.1. Пусть $[x_k, x_{k+1}]$ — один из подотрезков длины h , на которые разбит отрезок $[a, b]$, и пусть $\bar{x} = (x_k + x_{k+1})/2$. Используя тейлоровское разложение подынтегральной функции в точке \bar{x} , получить следующие представления:

$$\int_{x_k}^{x_{k+1}} f(x) dx = hf(\bar{x}) + \frac{h^3}{24} f''(\bar{x}) + \frac{h^5}{1920} f^{(4)}(\bar{x}) + \dots,$$

$$\int_{x_k}^{x_{k+1}} f(x) dx = \frac{f(x_k) + f(x_{k+1})}{2} h - \frac{h^3}{12} f''(\bar{x}) - \frac{h^5}{480} f^{(4)}(\bar{x}) - \dots$$

11.2. Применить **11.1**, учитывая, что $x_{k+1} - x_k = h \varphi'(kh) + o(1)$.

11.5. $S_{h,h/2} = S_{h/2} + \frac{S_{h/2} - S_h}{2^m - 1}$.

11.6. В обозначениях **11.5** имеем $m = 2$, $S_{h,h/2} = S_{h/2} + \frac{1}{3} (S_{h/2} - S_h)$ при $S_h = \frac{b-a}{2} (f(a) + f(b))$, $b - a = h$. Порядок главного члена погрешности увеличится на 2.

11.7. Действительно, если $S_{h/2} > S_h$, то $S_{h,h/2} > S_{h/2} > S_h$. Если $S_{h/2} < S_h$, то $S_{h,h/2} < S_{h/2} < S_h$.

11.8. Возможен следующий способ (*процесс Эйткена*), являющийся обобщением правила Рунге. Пусть I — точное значение интеграла. Запишем его приближенные значения с шагами h , $h/2$ и $h/4$. Если учитывать только главный член погрешности, то получаем систему трех уравнений:

$$I = S_h + ch^p, \quad I = S_{h/2} + \frac{1}{2^p} ch^p, \quad I = S_{h/4} + \frac{1}{4^p} ch^p,$$

в которой значения I , c и p неизвестны. Из первого и второго уравнений имеем $ch^p \left(1 - \frac{1}{2^p}\right) = S_{h/2} - S_h$. Из второго и третьего уравнений находим

$$\frac{1}{2^p} ch^p \left(1 - \frac{1}{2^p}\right) = S_{h/4} - S_{h/2}.$$

Из последних двух равенств получаем уравнение для определения p

$$2^p = \frac{S_{h/2} - S_h}{S_{h/4} - S_{h/2}}.$$

Выражение для главного члена погрешности имеет вид

$$ch^p = \frac{(S_{h/2} - S_h)^2}{2S_{h/2} - S_h - S_{h/4}}.$$

12.1. $c_1 = -c_2 = 1/(100\pi)$.

12.2. $c_1 = c_2 = 0$.

12.3. Обозначим $p = \omega(b - a)/2$ и вычислим коэффициенты:

$$D_1(p) = \int_{-1}^1 \frac{1 - \xi}{2} \exp\{ip\xi\} d\xi = \frac{\sin p}{p} + \frac{p \cos p - \sin p}{p^2} i,$$

$$D_2(p) = \int_{-1}^1 \frac{1 + \xi}{2} \exp\{ip\xi\} d\xi = \frac{\sin p}{p} - \frac{p \cos p - \sin p}{p^2} i.$$

12.4. При малых ω величина p мала. Функции $\cos p$ и $\sin p$ вычисляются с погрешностями $O(2^{-t})$ и $O(p2^{-t})$ соответственно, где t — длина мантиссы. Как следствие коэффициенты $D_1(p)$ и $D_2(p)$ из **12.3** приобретают погрешность $O(2^{-t})/p$.

12.7. Выделить функцию x^α в качестве весовой, а $f(x)$ на каждом отрезке разбиения заменить многочленом Лагранжа первой степени.

12.10. Представить подынтегральную функцию в виде $f(x) = G(x) + g(x)$, где $G(x) = \ln x$, $g(x) = -\frac{x^2 \ln x}{1 + x^2}$, при этом вычислить $\int_0^1 G(x) dx$ в явном виде.

12.11. См. указание к **12.10** при $G(x) = (1 - x^2) \ln x$.

12.12. Разбить отрезок интегрирования на $[0, \varepsilon]$ и $[\varepsilon, 1]$ с $\varepsilon \approx 1/\omega$. На первом отрезке функция $\sin(\omega x)$ не является осциллирующей, поэтому в качестве весовой функции можно взять $x^{-\alpha}$, а на втором отрезке использовать неравномерные узлы и весовую функцию $\sin(\omega x)$.

К главе IV. Матричные вычисления

13.1. Нет, поскольку неравенство треугольника не выполнено, например, для векторов $(1, 0)^T$ и $(0, 1)^T$.

13.2. Да.

13.3. Из неравенств $\max_{1 \leq i \leq n} |x_i| \leq \sum_{i=1}^n |x_i| \leq n \max_{1 \leq i \leq n} |x_i|$ следует

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1 \leq n \|\mathbf{x}\|_\infty. \text{ Так как } \sum_{i=1}^n x_i^2 \leq \left(\sum_{i=1}^n |x_i| \right)^2, \text{ то}$$

$\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1$. Из неравенства Коши—Буняковского $\sum_{i=1}^n |x_i| \leq \left(\sum_{i=1}^n 1\right)^{1/2} \left(\sum_{i=1}^n x_i^2\right)^{1/2} = n^{1/2} \left(\sum_{i=1}^n x_i^2\right)^{1/2}$, поэтому $n^{-1/2} \|\mathbf{x}\|_1 \leq \|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1$. Наконец, из неравенств $\max_{1 \leq i \leq n} x_i^2 \leq \sum_{i=1}^n x_i^2 \leq n \max_{1 \leq i \leq n} x_i^2$ имеем $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq n^{1/2} \|\mathbf{x}\|_\infty$.

13.4. 1) Достаточно проверить неравенство треугольника.

2) Матрица $A = A^T > 0$ обладает ортонормированной системой собственных векторов $\mathbf{q}_1, \dots, \mathbf{q}_n$: $A\mathbf{q}_i = \lambda_i \mathbf{q}_i$, $(\mathbf{q}_i, \mathbf{q}_j) = \delta_i^j$, $\lambda_i > 0$ — вещественные собственные значения A . Любой вектор \mathbf{x} представим в виде $\mathbf{x} = \sum_{i=1}^n c_i \mathbf{q}_i$, $\|\mathbf{x}\|_2^2 = \sum_{i=1}^n c_i^2$, поэтому

$$(A\mathbf{x}, \mathbf{x}) = \left(\sum_{i=1}^n \lambda_i c_i \mathbf{q}_i, \sum_{i=1}^n c_i \mathbf{q}_i \right) = \sum_{i=1}^n \lambda_i c_i^2. \text{ Отсюда}$$

$$\min_i \lambda_i \|\mathbf{x}\|_2^2 \leq (A\mathbf{x}, \mathbf{x}) \leq \max_i \lambda_i \|\mathbf{x}\|_2^2.$$

13.5. Для любого вектора $\mathbf{x} \in \mathbb{R}^n$ справедливо

$$\|A\mathbf{x}\|_\infty = \max_i \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \max_i \left(\sum_{j=1}^n |a_{ij}| \max_j |x_j| \right) \leq \max_i \left(\sum_{j=1}^n |a_{ij}| \right) \|\mathbf{x}\|_\infty.$$

Покажем, что эта оценка достигается. Пусть максимум по i имеет место при $i = l$. Тогда для $\mathbf{x} = (\text{sign}(a_{l1}), \text{sign}(a_{l2}), \dots, \text{sign}(a_{ln}))^T$ имеем $\|\mathbf{x}\|_\infty = 1$ и точные равенства во всей цепочке выше. Таким образом, $\|A\|_\infty = \max_i \left(\sum_{j=1}^n |a_{ij}| \right)$. Аналогично показывается, что

$$\|A\|_1 = \max_j \left(\sum_{i=1}^n |a_{ij}| \right).$$

По определению

$$\|A\|_2 = \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sup_{\mathbf{x} \neq 0} \sqrt{\frac{(A\mathbf{x}, A\mathbf{x})}{(\mathbf{x}, \mathbf{x})}} = \sup_{\mathbf{x} \neq 0} \sqrt{\frac{(A^T A \mathbf{x}, \mathbf{x})}{(\mathbf{x}, \mathbf{x})}}.$$

Матрица $B = A^T A$ — симметричная и $(B\mathbf{x}, \mathbf{x}) = (A\mathbf{x}, A\mathbf{x}) \geq 0$, следовательно, все ее собственные значения $\lambda_i(B) \geq 0$. Рассуждая далее, как в **13.4**, получим $\sup_{\mathbf{x} \neq 0} \frac{(B\mathbf{x}, \mathbf{x})}{(\mathbf{x}, \mathbf{x})} = \max_i \lambda_i(B)$, а равенство достигается на соответствующем собственном векторе. Поэтому

$$\|A\|_2 = \sqrt{\max_i \lambda_i(A^T A)}.$$

Отметим, что в случае симметричной матрицы $\|A\|_2 = \max_i |\lambda_i(A)|$.

13.6. Зафиксируем произвольный собственный вектор \mathbf{x} матрицы A и построим квадратную матрицу X , столбцами которой являются векторы \mathbf{x} . Тогда $\lambda X = AX$, откуда $|\lambda| \|X\| \leq \|A\| \|X\|$, т. е. $|\lambda| \leq \|A\|$.

13.7. 1) В силу неравенства Коши — Буняковского, и учитывая, что спектральная норма матрицы подчинена евклидовой векторной норме, получим

$$|\mathbf{y}^T A \mathbf{x}| = |(\mathbf{y}, A \mathbf{x})| \leq \|\mathbf{y}\|_2 \|A \mathbf{x}\|_2 \leq \|\mathbf{y}\|_2 \|\mathbf{x}\|_2 \|A\|_2 = \|A\|_2.$$

Осталось показать, что существуют векторы \mathbf{x} и \mathbf{y} единичной длины, на которых $|\mathbf{y}^T A \mathbf{x}| = \|A\|_2$. Пусть вектор \mathbf{x} такой, что $\|A \mathbf{x}\|_2 = \|A\|_2$, а $\mathbf{y} = A \mathbf{x} / \|A \mathbf{x}\|_2$. Тогда $\|\mathbf{y}\|_2 = 1$ и

$$|\mathbf{y}^T A \mathbf{x}| = \frac{(A \mathbf{x})^T}{\|A \mathbf{x}\|_2} A \mathbf{x} = \frac{(A \mathbf{x}, A \mathbf{x})}{\|A \mathbf{x}\|_2} = \|A \mathbf{x}\|_2 = \|A\|_2.$$

2) Используя свойство 1), получим

$$\|A^T\|_2 = \sup_{\substack{\|\mathbf{x}\|_2=1 \\ \|\mathbf{y}\|_2=1}} |\mathbf{y}^T A^T \mathbf{x}| = \sup_{\substack{\|\mathbf{x}\|_2=1 \\ \|\mathbf{y}\|_2=1}} (\mathbf{y}, A^T \mathbf{x}) = \sup_{\substack{\|\mathbf{x}\|_2=1 \\ \|\mathbf{y}\|_2=1}} (A \mathbf{y}, \mathbf{x}) = \|A\|_2.$$

3) Из свойства 2) следует неравенство $\|A^T A\|_2 \leq \|A^T\|_2 \|A\|_2 = \|A\|_2^2$. Возьмем теперь такой вектор \mathbf{x} , что $\|\mathbf{x}\|_2 = 1$ и $\|A \mathbf{x}\|_2 = \|A\|_2$, и применим свойство 1) к матрице $A^T A$, положив $\mathbf{y} = \mathbf{x}$. Получим неравенство

$$\|A^T A\|_2 \geq |\mathbf{x}^T A^T A \mathbf{x}| = (A \mathbf{x}, A \mathbf{x}) = \|A \mathbf{x}\|_2^2 = \|A\|_2^2.$$

Из этих двух неравенств следует $\|A^T A\|_2 = \|A\|_2^2$. Аналогично — для $\|A A^T\|_2$.

13.8. Из свойства 3) спектральной нормы (см. **13.7**) следует

$$\|QA\|_2^2 = \|(QA)^T QA\|_2 = \|A^T Q^T QA\|_2 = \|A^T A\|_2 = \|A\|_2^2,$$

а из свойства 2) и полученного равенства имеем

$$\|AQ\|_2 = \|(AQ)^T\|_2 = \|Q^T A^T\|_2 = \|A^T\|_2 = \|A\|_2.$$

13.9. Модуль любого собственного значения матрицы не больше любой ее нормы, поэтому (см. **13.5**)

$$\|A\|_2^2 = \max \lambda(A^T A) \leq \|A^T A\|_1 \leq \|A\|_1 \|A^T\|_1 = \|A\|_1 \|A\|_\infty.$$

13.10. Для любой матричной нормы должно выполняться неравенство $\|AC\| \leq \|A\| \|C\|$. Однако для матриц $A = C$, $a_{ij} = c_{ij} = 1 \forall i, j$, для которых имеют место соотношения $\eta(AC) = n$, $\eta(A) = \eta(C) = 1$, оно не выполняется.

13.11. Проверим четвертое условие в определении нормы:

$$\begin{aligned} M(AC) &= n \max_{i,j} \left| \sum_{k=1}^n a_{ik} c_{kj} \right| \leq n \max_{i,j} \sum_{k=1}^n |a_{ik} c_{kj}| \leq \\ &\leq n \sum_{k=1}^n \eta(A) \eta(C) = n \eta(A) n \eta(C) = M(A)M(C). \end{aligned}$$

13.12. Заметим, что $\|\mathbf{x}\|_h = \|\mathbf{y}\|_\infty$, где $\mathbf{y} = S\mathbf{x}$, $S = \begin{pmatrix} 1 & 0 \\ -1/h & 1/h \end{pmatrix}$.

Поэтому

$$\begin{aligned} \|A\|_h &= \sup_{\mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_h}{\|\mathbf{x}\|_h} = \sup_{\mathbf{x} \neq 0} \frac{\|S A \mathbf{x}\|_\infty}{\|S \mathbf{x}\|_\infty} = \sup_{\mathbf{y} \neq 0} \frac{\|S A S^{-1} \mathbf{y}\|_\infty}{\|\mathbf{y}\|_\infty} = \\ &= \|S A S^{-1}\|_\infty = \left\| \begin{pmatrix} a_{11} + a_{12} & a_{12} h \\ \frac{a_{21} + a_{22} - a_{11} - a_{12}}{h} & a_{22} - a_{12} \end{pmatrix} \right\|_\infty = \\ &= \max\{|a_{11} + a_{12}| + h|a_{12}|; |a_{22} - a_{12}| + \frac{1}{h}|a_{21} + a_{22} - a_{11} - a_{12}|\}. \end{aligned}$$

13.13. $n^{-1/2}N(A) \leq \|A\|_1 \leq n^{1/2}N(A)$, $n^{-1/2}N(A) \leq \|A\|_2 \leq N(A)$, $n^{-1/2}N(A) \leq \|A\|_\infty \leq n^{1/2}N(A)$.

13.14. $\|DAD^{-1}\|_\infty$, где $D = \text{diag}(d_1, \dots, d_n)$.

13.15. $\|DAD^{-1}\|_1$, где $D = \text{diag}(d_1, \dots, d_n)$.

13.16. $\|DAD^{-1}\|_2$, где $D = \text{diag}(\sqrt{d_1}, \dots, \sqrt{d_n})$.

13.17. $\|LAL^{-1}\|_\infty$, где $l_{ij} = 1$ при $i \leq j$ и $l_{ij} = 0$ при $i > j$.

13.18. $c_1 = 1/n$, $c_2 = 1$.

13.19. $c_1 = 1/n$, $c_2 = 1$.

13.20. Пусть для определенности $|x_1| = \max_i |x_i| \neq 0$. Тогда

$$\|\mathbf{x}\|_p = |x_1| \left(1 + \sum_{i=2}^n \alpha_i^p\right)^{1/p}, \quad \|\mathbf{x}\|_q = |x_1| \left(1 + \sum_{i=2}^n \alpha_i^q\right)^{1/q}$$

с $\alpha_i = |x_1|^{-1}|x_i| \leq 1$. Так как $q \leq p$, то $\alpha_i^p \leq \alpha_i^q$, а $t^{1/p} \leq t^{1/q}$ для $t \geq 1$. Таким образом, справедливо

$$\left(1 + \sum_{i=2}^n \alpha_i^p\right)^{1/p} \leq \left(1 + \sum_{i=2}^n \alpha_i^q\right)^{1/q} \quad \text{и} \quad \|\mathbf{x}\|_p \leq \|\mathbf{x}\|_q.$$

13.21. Следует из неравенств $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_p \leq n^{1/p} \|\mathbf{x}\|_\infty$ в пределе при $p \rightarrow \infty$.

13.24. Обозначить $x_k = a_k + i b_k$ и воспользоваться неравенством треугольника для векторов, координатами которых являются $|a_k|$ и $|b_k|$. В результате получить $c_0 = 1/\sqrt{2}$.

13.25. $\|\cdot\|_1$.

13.26. Пусть $A = \begin{pmatrix} B & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$, где B — подматрица размерности $m \times n$.

По определению, $\|A\|_p = \sup_{\|\mathbf{x}\|_p=1} \|A\mathbf{x}\|_p$. Возьмем вектор \mathbf{x}^* , $\|\mathbf{x}^*\|_p = 1$ такой, что $\|B\|_p = \|B\mathbf{x}^*\|_p$, и положим $\tilde{\mathbf{x}} = (x_1^*, \dots, x_n^*, 0, \dots, 0)^T$. Тогда $\|\tilde{\mathbf{x}}\|_p = 1$ и

$$\|A\|_p \geq \|A\tilde{\mathbf{x}}\|_p \geq \|B\mathbf{x}^*\|_p = \|B\|_p.$$

13.29. $\|A^{-1}\| = \sup_{\mathbf{y} \neq 0} \|A^{-1}\mathbf{y}\|/\|\mathbf{y}\| = \sup_{\mathbf{x} \neq 0} \|\mathbf{x}\|/\|A\mathbf{x}\| = \left(\inf_{\mathbf{x} \neq 0} \|A\mathbf{x}\|/\|\mathbf{x}\| \right)^{-1}$.

13.30. Воспользоваться решениями **13.6** и **13.9**.

13.32. Пусть Q — ортогональная матрица, i -й столбец которой есть i -й собственный вектор из полной ортонормированной системы собственных векторов матрицы A , а D — диагональная матрица с i -м собственным значением матрицы A в i -й строке. Тогда $A = QDQ^T$ и $A^{1/2} = QD^{1/2}Q^T$. Матрица $A^{1/2}$ коммутирует с A и любой ее степенью, значит, — с $p_m(A)$. Используя этот факт, а также определения норм $\|\cdot\|_2$, $\|\cdot\|_A$, получим цепочку равенств

$$\begin{aligned} \|p_m(A)\|_A &= \sup_{\mathbf{x} \neq 0} \frac{(p_m(A)\mathbf{x}, p_m(A)\mathbf{x})^{1/2}}{(A\mathbf{x}, \mathbf{x})^{1/2}} = \\ &= \sup_{\mathbf{x} \neq 0} \frac{(p_m(A)A^{1/2}\mathbf{x}, p_m(A)A^{1/2}\mathbf{x})^{1/2}}{\|A^{1/2}\mathbf{x}\|_2} = \sup_{\mathbf{y} \neq 0} \frac{\|p_m(A)\mathbf{y}\|_2}{\|\mathbf{y}\|_2} = \|p_m(A)\|_2. \end{aligned}$$

13.33. Всякая вещественная матрица A представима в виде $A = S + K$, где $S = (A + A^T)/2$ — симметричная, а $K = (A - A^T)/2$ — кососимметричная матрицы. При этом для любого вещественного $\mathbf{x} \neq 0$ имеем $(A\mathbf{x}, \mathbf{x}) = (S\mathbf{x}, \mathbf{x}) \geq \delta(\mathbf{x}, \mathbf{x})$, где $\delta \geq 0$ — минимальное собственное значение матрицы S . Из неравенства $(A\mathbf{x}, \mathbf{x}) > 0$ следует, что $\delta > 0$.

13.34. Матрица

$$A = \begin{pmatrix} A_{11} & 0 \\ 0 & B \end{pmatrix}, \text{ где } A_{11} = \begin{pmatrix} a & -1 \\ 1 & a \end{pmatrix}$$

с положительной константой a и симметричной положительно определенной подматрицей B , положительно определена, но имеет пару комплексно сопряженных собственных значений $\lambda = a \pm i$.

- 13.35.** Воспользоваться тем фактом, что для любой подчиненной нормы справедливо $\|I\| = 1$, где I — единичная матрица.
- 13.36.** Воспользоваться решением **13.6**.
- 13.37.** Воспользоваться решением **13.6**.
- 13.38.** Воспользоваться тем фактом, что нормальная матрица имеет полную ортонормированную систему собственных векторов.
- 13.39.** Воспользоваться решением **13.33**.

14.1. Из равенства $A^{-1}\mathbf{r} = A^{-1}\mathbf{b} - A^{-1}A\tilde{\mathbf{x}} = \mathbf{x} - \tilde{\mathbf{x}}$ следует $\|\mathbf{x} - \tilde{\mathbf{x}}\| \leq \|A^{-1}\| \|\mathbf{r}\|$, а из равенства $\mathbf{b} = A\mathbf{x}$ имеем $\|\mathbf{b}\| = \|A\mathbf{x}\| \leq \|A\| \|\mathbf{x}\|$, поэтому

$$\frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} \leq \|A\| \|A^{-1}\| \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|} = \text{cond}(A) \frac{\|\mathbf{r}\|}{\|\mathbf{b}\|} = \text{cond}(A) \frac{\|\mathbf{b} - A\tilde{\mathbf{x}}\|}{\|\mathbf{b}\|}.$$

Отсюда видно, что если матрица A плохо обусловлена, то даже очень маленькая невязка не может гарантировать малость относительной ошибки в $\tilde{\mathbf{x}}$. С другой стороны, может оказаться так, что достаточно точное решение имеет большую невязку.

Рассмотрим пример с матрицей $A = \begin{pmatrix} 1,000 & 1,001 \\ 1,000 & 1,000 \end{pmatrix}$. Пусть

$\mathbf{b} = (2,001, 2,000)^T$. Тогда $\mathbf{x} = (1, 1)^T$ — точное решение системы $A\mathbf{x} = \mathbf{b}$. Однако вектор $\tilde{\mathbf{x}} = (2, 0)^T$, который никак нельзя назвать близким к \mathbf{x} , дает маленькую невязку $\mathbf{r} = (10^{-3}, 0)^T$. Возьмем теперь $\mathbf{b} = (1, 0)^T$. Тогда $\mathbf{x} = (-1000, 1000)^T$ — точное решение, вектор $\tilde{\mathbf{x}} = (-1001, 1000)^T$ близок к \mathbf{x} в смысле относительной погрешности, однако $\tilde{\mathbf{x}}$ дает большую невязку $\mathbf{r} = (1, 1)^T$, близкую по норме к вектору \mathbf{b} .

- 14.2.** $(1, 1)^T$ и $(7, -1)^T$. Обе матрицы получены малыми возмущениями одной вырожденной матрицы. В данном случае это приводит к большой разнице в решениях систем.
- 14.3.** $\text{cond}(A) \geq 1$, так как $1 = \|A A^{-1}\| \leq \|A\| \|A^{-1}\| = \text{cond}(A)$.

Умножение матрицы на ортогональную не меняет ее спектральную норму, поэтому $\|Q\|_2 = \|QI\|_2 = \|I\|_2 = 1$ и $\|Q^T\|_2 = \|Q^T I\|_2 = \|I\|_2 = 1$. Отсюда $\text{cond}_2(Q) = \|Q\|_2 \|Q^{-1}\|_2 = \|Q\|_2 \|Q^T\|_2 = 1$.

- 14.4.** Пусть $D = \varepsilon I$, где $\varepsilon > 0$ — малое число и I — единичная матрица. Определитель $\det(D) = \varepsilon^n$ весьма мал, тогда как матрица D хорошо обусловлена, поскольку $\text{cond}(D) = \|D\| \|D^{-1}\| = 1$.

Рассмотрим теперь матрицу

$$A = \begin{pmatrix} 1 & -1 & -1 & \dots & -1 \\ 0 & 1 & -1 & \dots & -1 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix},$$

у которой определитель равен 1, и вычислим ее число обусловленности. Для этого построим в явном виде обратную матрицу. Возьмем произвольный вектор $\mathbf{b} \neq 0$ и решим систему $A\mathbf{x} = \mathbf{b}$ с помощью обратной подстановки:

$$x_n = b_n, \quad x_{n-1} = b_{n-1} + b_n, \quad \dots, \quad x_1 = b_1 + b_2 + 2b_3 + \dots + 2^{n-3}b_{n-1} + 2^{n-2}b_n.$$

Отсюда следует, что обратная матрица имеет вид

$$A^{-1} = \begin{pmatrix} 1 & 1 & 2 & 4 & \dots & 2^{n-3} & 2^{n-2} \\ 0 & 1 & 1 & 2 & \dots & 2^{n-4} & 2^{n-3} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

$\|A^{-1}\|_\infty = 1 + 1 + 2 + 2^2 + \dots + 2^{n-2} = 2^{n-1}$, $\|A\|_\infty = n$ и $\text{cond}_\infty(A) = n2^{n-1}$, т. е. матрица A плохо обусловлена, хотя $\det(A) = 1$.

Рассмотренные примеры показывают, что обусловленность матрицы зависит не только от величины определителя.

14.5. Как в **14.4**, методом обратной подстановки получим обратную матрицу

$$A^{-1} = \begin{pmatrix} 1 & -a & (-a)^2 & \dots & (-a)^{n-2} & (-a)^{n-1} \\ 0 & 1 & -a & \dots & (-a)^{n-3} & (-a)^{n-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & -a \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

$$\|A^{-1}\|_\infty = 1 + |a| + a^2 + \dots + |a|^{n-1} = \frac{|a|^n - 1}{|a| - 1},$$

$$\|A\|_\infty = 1 + |a|, \quad \text{cond}_\infty(A) = \frac{(|a| + 1)(|a|^n - 1)}{|a| - 1}.$$

Отсюда видно, что матрица A плохо обусловлена при $|a| > 1$ и хорошо обусловлена при $|a| < 1$.

14.6. Воспользоваться решением **14.5**.

14.7. Воспользоваться решением **14.5**.

14.8. Вычислим в явном виде

$$A^{-1} = \begin{pmatrix} 0 & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1-\varepsilon}{4\varepsilon} & \frac{1+\varepsilon}{4\varepsilon} \\ \frac{1}{2} & \frac{1+\varepsilon}{4\varepsilon} & \frac{1-\varepsilon}{4\varepsilon} \end{pmatrix}, \quad (A')^{-1} = \begin{pmatrix} 0 & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1-\varepsilon}{4} & \frac{1+\varepsilon}{4} \\ \frac{1}{2} & \frac{1+\varepsilon}{4} & \frac{1-\varepsilon}{4} \end{pmatrix},$$

поэтому число обусловленности исходной матрицы стремится к бесконечности при $\varepsilon \rightarrow 0$, а число обусловленности матрицы A' остается ограниченным.

14.9. Получить $\text{cond}_2(A + \alpha I) = \frac{M + \alpha}{m + \alpha} = 1 + \frac{M - m}{m + \alpha}$.

14.10. Да. Примером такой матрицы является

$$A = \begin{pmatrix} 10^3 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 10^{-3} \end{pmatrix}, \quad \begin{aligned} \lambda(A^T A) &\in \{10^6, 10^{-6}, 4, 5 \pm \sqrt{4,25}\}, \\ \text{cond}(A^2) &= \|A^2\| \|A^{-2}\| = 10^{12}, \\ \text{cond}(A) &= \|A\| \|A^{-1}\| = 10^6. \end{aligned}$$

14.11. Воспользовавшись неравенством для векторных норм $\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq \sqrt{n} \|\mathbf{x}\|_2$, получим $\frac{1}{\sqrt{n}} \|A\|_2 \leq \|A\|_1 \leq \sqrt{n} \|A\|_2$, откуда следует требуемый результат.

14.12. Из неравенства для чисел обусловленности в матричных нормах $\|\cdot\|_\infty$ и $\|\cdot\|_2$ и равенства $\|A\|_2^2 = \lambda_{\max}(A^T A)$ следует

$$n^{-1} \sqrt{\beta/\alpha} \leq \text{cond}_\infty(A) \leq n \sqrt{\beta/\alpha}.$$

14.13. Воспользоваться решением 13.6.

14.15. Воспользоваться явной формулой для собственных векторов матрицы A (см. 22.12): $\mathbf{y}_k^{(j)} = \sin \frac{\pi j k}{n+1}$, $j = 1, \dots, n$. Собственные значения равны $\lambda^{(j)} = 4 \sin^2 \frac{\pi j}{2(n+1)}$, так что

$$\text{cond}_2(A) = \text{ctg}^2 \frac{\pi}{2(n+1)} \approx \frac{4(n+1)^2}{\pi^2}.$$

14.16. Воспользоваться явной формулой для собственных векторов матрицы A : $\mathbf{y}_k^{(j)} = \sin \frac{\pi j k}{n+1}$, $j = 1, \dots, n$. Собственные значения равны $\lambda^{(j)} = \frac{2}{3} + \frac{1}{3} \cos \frac{\pi j}{n+1}$, так что $\text{cond}_2(A) \leq 3$.

14.17. Так как $1 - \|\delta I\| > 0$ и $\|\mathbf{x}\| \leq \|\mathbf{x} - \delta I \mathbf{x}\| + \|\delta I \mathbf{x}\|$, то

$$\|(I - \delta I) \mathbf{x}\| \geq \|\mathbf{x}\| - \|\delta I \mathbf{x}\| \geq \|\mathbf{x}\| - \|\delta I\| \|\mathbf{x}\| = (1 - \|\delta I\|) \|\mathbf{x}\| > 0 \quad \forall \mathbf{x} \neq 0.$$

Следовательно, матрица $I - \delta I$ невырожденная. Из тождества $(I - \delta I)(I - \delta I)^{-1} = I$ получим $(I - \delta I)^{-1} = I + \delta I(I - \delta I)^{-1}$ и

$$\|(I - \delta I)^{-1}\| \leq \|I\| + \|\delta I\| \|(I - \delta I)^{-1}\| = 1 + \|(I - \delta I)^{-1}\| \|\delta I\|.$$

14.18. Из равенства $(I - \delta I)^{-1} = I + \delta I(I - \delta I)^{-1}$ (см. **14.17**) следует $I - (I - \delta I)^{-1} = -\delta I(I - \delta I)^{-1}$. Отсюда, в силу неравенства из **14.17**, имеем

$$\|I - (I - \delta I)^{-1}\| \leq \|\delta I\| \|(I - \delta I)^{-1}\| \leq \frac{\|\delta I\|}{1 - \|\delta I\|}.$$

14.19. Имеем $A + \delta A = A(I + A^{-1}\delta A)$. Поскольку $\|A^{-1}\delta A\| < 1$, из **14.17** следует, что матрица $I + A^{-1}\delta A$ невырожденная. Это означает, что матрица $A + \delta A$ также не вырождена. Из равенства $(A + \delta A)^{-1} = (I + A^{-1}\delta A)^{-1}A^{-1}$, в силу неравенства из **14.17**, следует

$$\|(A + \delta A)^{-1}\| \leq \|(I + A^{-1}\delta A)^{-1}\| \|A^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\delta A\|}.$$

14.20. Из равенства $(A + \delta A)^{-1} = (I + A^{-1}\delta A)^{-1}A^{-1}$ получим $A^{-1} - (A + \delta A)^{-1} = (I - (I + A^{-1}\delta A)^{-1})A^{-1}$. Тогда, в силу неравенства из **14.18**,

$$\|A^{-1} - (A + \delta A)^{-1}\| \leq \|I - (I + A^{-1}\delta A)^{-1}\| \|A^{-1}\| \leq \frac{\|A^{-1}\delta A\|}{1 - \|A^{-1}\delta A\|} \|A^{-1}\|.$$

Относительная ошибка в матрице $(A + \delta A)^{-1}$ оценивается неравенством

$$\frac{\|A^{-1} - (A + \delta A)^{-1}\|}{\|A^{-1}\|} \leq \frac{\|A^{-1}\| \|\delta A\|}{1 - \|A^{-1}\delta A\|} = \frac{\text{cond}(A)}{1 - \|A^{-1}\delta A\|} \frac{\|\delta A\|}{\|A\|}.$$

14.24. Вывести неравенства:

$$\begin{aligned} \max_i |a_{ii}| &\leq \|A\|_\infty \leq (1 + q) \max_i |a_{ii}|, \\ \frac{1}{(1 + q) \min_i |a_{ii}|} &\leq \|A^{-1}\|_\infty \leq \frac{1}{1 - q} \frac{1}{\min_i |a_{ii}|}. \end{aligned}$$

14.25. $\max_R \text{cond}_\infty(R) = n 2^{n-1}$.

14.26. Вычтем первый столбец определителя последовательно из второго, третьего, ..., n -го столбцов, а затем вынесем за знак определителя $b_1 - b_2$ из второго столбца, $b_1 - b_3$ — из третьего и т. д. Затем вынесем $(a_1 + b_1)^{-1}$ из первой строки, $(a_2 + b_1)^{-1}$ — из второй строки

и т. д. Далее вычтем первую строку последовательно из второй, третьей, ..., n -й строки, а затем вынесем за знак определителя $(a_1 - a_2) \dots (a_1 - a_n)(a_1 + b_2)^{-1} \dots (a_1 + b_n)^{-1}$. В результате останется определитель матрицы Коши $(n - 1)$ -го порядка. Поэтому искомая формула получается по индукции.

14.27. Рассмотрим матрицу Коши K_n с элементами $k_{ij} = 1/(a_i + b_j)$, $1 \leq i, j \leq n$. Ее определитель вычислен в **14.26**. Элементы матрицы K_n^{-1} являются отношениями алгебраических дополнений к определителю исходной матрицы. Миноры матрицы Коши снова являются матрицами Коши. Поэтому можно получить явные выражения для элементов K_n^{-1} :

$$\tilde{k}_{ij} = \prod_{k=1}^n (a_j + b_k)(a_k + b_i) \left[(a_j + b_i) \prod_{k \neq j} (a_j - a_k) \prod_{k \neq i} (b_i - b_k) \right]^{-1}.$$

Полагая $a_i = i$, $b_i = i - 1$, получим частный случай матрицы Коши — матрицу Гильберта и искомую формулу для элементов H_n^{-1} .

14.28. Величина $\frac{(i + n - 1)!}{((i - 1)!)^2(n - i)!}$ принимает максимальное значение при $i = [n/\sqrt{2}]$, поэтому для элементов \tilde{h}_{ij} матрицы H_n^{-1} (см. **14.27**) по формуле Стирлинга $n! = \sqrt{2\pi n} n^n e^{-n} e^{\theta(n)}$, $|\theta(n)| \leq \frac{1}{12n}$, имеем асимптотику

$$\max_{i,j} |\tilde{h}_{ij}| = \frac{1}{4\sqrt{2}\pi^2 n} (\sqrt{2} + 1)^{4n} \left(1 + O\left(\frac{1}{n}\right)\right).$$

Отсюда следует равенство

$$\|H_n^{-1}\|_\infty = \frac{1}{(2\pi)^{3/2} 2^{7/4} \sqrt{n}} (\sqrt{2} + 1)^{4n} \left(1 + O\left(\frac{1}{n}\right)\right).$$

Так как для $n \geq 2$ справедливо

$$\ln n \leq \|H_n\|_\infty = \sum_{j=1}^n \frac{1}{j} \leq 3 \ln n,$$

то главный член асимптотики $\text{cond}_\infty(H_n)$ имеет вид $\text{const} \cdot 4^n \ln n / \sqrt{n}$.

15.1. Проверить достаточное условие $\|B\|_1 = \|B\|_\infty < 1$.

15.2. Собственные числа матрицы B равны α и $\alpha \pm \sqrt{2}\beta$, откуда следуют условия $|\alpha| < 1$, $|\alpha \pm \sqrt{2}\beta| < 1$.

15.3. $B = \begin{pmatrix} 1 & 0,5 \\ 0,5 & 1 \end{pmatrix}$, $\mathbf{x} = (1, 1)^T$, $\mathbf{c} = -(0,5, 0,5)^T$, $\lambda(B) \in \{0,5, 1,5\}$,

$B\mathbf{q} = 0,5\mathbf{q}$, $\mathbf{q} = (1, -1)^T$, $\mathbf{x}^0 = \mathbf{x} + \gamma\mathbf{q}$ при $\gamma \neq 0$.

15.4. $N \approx 1/(1 - \alpha)$.

15.5. Собственные значения оператора перехода $B = I - \tau A$ имеют вид $\lambda(B) = 1 - \tau\lambda(A)$. Так как $0 < \lambda(A) \leq \|A\|$, то $0 \leq \lambda(B) < 1$.

15.6. Пусть λ — произвольное собственное значение матрицы A и \mathbf{x} — соответствующий ему собственный вектор. Обозначим через x_i максимальную по модулю компоненту вектора \mathbf{x} . Если таких компонент несколько, то x_i — любая из них. Из равенства $A\mathbf{x} = \lambda\mathbf{x}$ следует $(\lambda - a_{ii})x_i = \sum_{j \neq i} a_{ij}x_j$. Отсюда имеем

$$|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}| \frac{|x_j|}{|x_i|} \leq \sum_{j \neq i} |a_{ij}| = R_i(A).$$

Это утверждение называется *теоремой Гершгорина*. Имеется обобщение этого факта:

Теорема. Если указанное объединение кругов $G(A)$ распадается на несколько связанных частей, то каждая такая часть содержит столько собственных значений, сколько кругов ее составляют.

15.7. Собственные значения матриц A и A^T совпадают.

15.8. $1,6 \leq \lambda_1 \leq 2,4$, $3,5 \leq \lambda_2 \leq 4,5$, $4,8 \leq \lambda_3 \leq 5,2$.

15.9. У матрицы $\begin{pmatrix} 0 & 0,1 \\ -40 & 5 \end{pmatrix}$ оба собственных значения $\lambda_1 = 1$ и $\lambda_2 = 4$ не принадлежат системе кругов: $|z| \leq 0,1$, $|z - 5| \leq 0,1$.

15.13. Для вектора ошибки \mathbf{z}^k справедливо равенство $\mathbf{z}^{k+1} = (I - \tau A)\mathbf{z}^k = (QQ^{-1} - \tau QDQ^{-1})\mathbf{z}^k$, откуда для $Q^{-1}\mathbf{z}^k = \tilde{\mathbf{z}}^k$ получаем $\tilde{\mathbf{z}}^{k+1} = (I - \tau D)\tilde{\mathbf{z}}^k$ с диагональной матрицей $B = I - \tau D$. Собственные значения матрицы B равны $\lambda(B) = 1 - \tau\lambda(A)$. Поэтому необходимым и достаточным условием сходимости метода является неравенство $|1 - \tau\lambda(A)| < 1 \forall \lambda(A) \in [m, M]$.

15.14. Воспользоваться теоремой Гершгорина и решением 15.13.

15.15. Воспользоваться видом оператора перехода в базисе из собственных векторов матрицы A и получить $\tau = 1/\lambda$.

15.16. Разложить ошибку $\mathbf{x} - \mathbf{x}^k$ по базису $\{\mathbf{q}_i\}_{i=1}^n$ из собственных векторов матрицы A . Выбор $\tau_k = \lambda_k^{-1}$, $k = 1, \dots, n$, обеспечивает на каждом шаге обнуление коэффициента при векторе \mathbf{q}_k в разложении ошибки (см. 15.15).

15.18. Имеем

$$\mathbf{x}^{k+1} = (\alpha I + (1 - \alpha)B)\mathbf{x}^k + (1 - \alpha)\mathbf{c},$$

$$\min_{\alpha} \|\alpha I + (1 - \alpha)B\| = \min_{\alpha} \max_{\lambda} |\alpha + (1 - \alpha)\lambda|, \quad \alpha = \frac{m + M}{m + M - 2}.$$

15.19. $a_{ij} = \{1 \text{ при } i = j; 1 \text{ при } i + 1 = j; 0 \text{ иначе}\}.$

15.23. Следует из равенства $\mathbf{x} - \mathbf{x}^k = (I - B)^{-1}(\mathbf{x}^{k+1} - \mathbf{x}^k).$

15.26. Собственные значения λ оператора перехода $I - \tau A$ имеют вид

$$\lambda(I - \tau A) = 1 - \tau u - i\tau v, \quad 0 < \delta \leq u \leq 1, \quad |v| \leq 1.$$

Из условия сходимости $|\lambda|^2 = (1 - \tau u)^2 + \tau^2 v^2 < 1$ следует $\tau < 2u/(u^2 + v^2).$ $\min_{u,v} \frac{u}{u^2 + v^2} = \min_u \frac{u}{u^2 + 1} = \frac{\delta}{\delta^2 + 1},$ поэтому $0 < \tau < \frac{2\delta}{\delta^2 + 1}.$

15.27. Метод сходится с произвольного начального приближения, так как $\|B\|_1 < 1.$

15.28. Матрица положительно определена и имеет два кратных собственных значения $\lambda_1 = 1$ и $\lambda_2 = 2.$ Условие сходимости: $0 < \tau < 1/2.$

15.29. Метод простой итерации $\mathbf{x}^{k+1} = B\mathbf{x}^k + \mathbf{c}$ сходится при $|\lambda_i(B)| < 1.$ Собственные значения матрицы перехода в методе $\mathbf{x}^{k+1} = (2B^2 - I)\mathbf{x}^k + 2(B + I)\mathbf{c}$ равны $\lambda_i(2B^2 - I) = 2\lambda_i^2(B) - 1.$ Этот итерационный метод сходится быстрее метода простой итерации, если спектр матрицы B расположен в подмножестве единичного круга комплексной плоскости, где функция $|2z^2 - 1|$ меньше функции $|z|.$ В частности, если спектр матрицы B вещественный, то он должен принадлежать объединению интервалов $(-1, -1/2)$ и $(1/2, 1).$

16.1. Оператор перехода в методе Якоби $B_J = -D^{-1}(L + R).$ Задача на собственные значения $B_J \mathbf{x} = \lambda \mathbf{x}$ преобразуется к виду $(L + \lambda D + R)\mathbf{x} = 0.$ Уравнение $\det(L + \lambda D + R) = \alpha\lambda(\alpha^2\lambda^2 - 2\beta^2) = 0$ имеет решения $\lambda_1 = 0, \lambda_{2,3}^2 = 2\beta^2/\alpha^2.$ Отсюда следует ответ $|\beta/\alpha| < 1/\sqrt{2}.$

Оператор перехода в методе Гаусса—Зейделя имеет вид $B_{GZ} = -(D + L)^{-1}R$ и задача на собственные значения $B_{GZ}\mathbf{x} = \lambda \mathbf{x}$ приводит к характеристическому уравнению $\det(\lambda L + \lambda D + R) = \alpha\lambda^2(\alpha^2\lambda - 2\beta^2) = 0.$ Следовательно, $\lambda_{1,2} = 0, \lambda_3 = 2\beta^2/\alpha^2.$ Отсюда следует ответ $|\beta/\alpha| < 1/\sqrt{2}.$

Как видно, в данном случае области сходимости методов совпадают.

16.2. Операторы перехода методов Якоби и Гаусса–Зейделя

$$B_J = \begin{pmatrix} 0 & -a_{12} a_{11}^{-1} \\ -a_{21} a_{22}^{-1} & 0 \end{pmatrix}, \quad B_{GZ} = \begin{pmatrix} 0 & -a_{12} a_{11}^{-1} \\ 0 & a_{12} a_{21} (a_{11} a_{22})^{-1} \end{pmatrix}$$

имеют собственные значения

$$\lambda_{1,2}^J = \pm \sqrt{\frac{a_{12} a_{21}}{a_{11} a_{22}}}, \quad \lambda_1^{GZ} = 0, \quad \lambda_2^{GZ} = \frac{a_{12} a_{21}}{a_{11} a_{22}},$$

откуда следует искомое утверждение.

16.3. Вектор ошибки \mathbf{z}^k удовлетворяет уравнению $(D+L)\mathbf{z}^{k+1} + R\mathbf{z}^k = 0$.

Пусть $\|\mathbf{z}^{k+1}\|_\infty = |z_l^{k+1}|$. Из l -го уравнения получаем

$$z_l^{k+1} = - \sum_{j=1}^{l-1} \frac{a_{lj}}{a_{ll}} z_j^{k+1} - \sum_{j=l+1}^n \frac{a_{lj}}{a_{ll}} z_j^k,$$

откуда $\|\mathbf{z}^{k+1}\|_\infty = |z_l^{k+1}| \leq \alpha \|\mathbf{z}^{k+1}\|_\infty + \beta \|\mathbf{z}^k\|_\infty$, или

$$\|\mathbf{z}^{k+1}\|_\infty \leq \frac{\beta}{1-\alpha} \|\mathbf{z}^k\|_\infty, \quad \text{где } \alpha = \sum_{j=1}^{l-1} \left| \frac{a_{lj}}{a_{ll}} \right|, \quad \beta = \sum_{j=l+1}^n \left| \frac{a_{lj}}{a_{ll}} \right|.$$

По условию $\alpha + \beta \leq q < 1$, следовательно,

$$\frac{\beta}{1-\alpha} \leq \frac{q-\alpha}{1-\alpha} \leq \frac{q-q\alpha}{1-\alpha} \leq q.$$

16.4. Метод сходится в обоих случаях.

16.5. Если формулу метода релаксации

$$(D + \tau L)\mathbf{x}^{k+1} + [\tau R + (\tau - 1)D]\mathbf{x}^k = \tau \mathbf{b}$$

умножить слева на матрицу D^{-1} , то оператор перехода можно записать в следующем виде:

$$B = (I + \tau M)^{-1}((1 - \tau)I + \tau N).$$

Здесь I — единичная, $M = D^{-1}L$ и $N = D^{-1}R$ — строго нижняя и верхняя треугольные матрицы, соответственно. Для характеристического многочлена $d(\lambda) = \det(B - \lambda I)$ согласно теореме

Виета справедливо равенство $(-1)^n d(0) = \prod_{i=1}^n \lambda_i(B)$. Так как у треугольных матриц M и N на главной диагонали расположены

нули, то $d(0) = \det(B) = (1 - \tau)^n$. Отсюда получим оценку для спектрального радиуса оператора перехода

$$\rho(B) = \max_i |\lambda_i(B)| \geq \left| \prod_{i=1}^n \lambda_i(B) \right|^{1/n} = |\det(B)|^{1/n} = |1 - \tau|,$$

которая, в силу необходимого неравенства $\rho(B) < 1$, дает требуемое.

16.6. Используя эквивалентную форму метода

$$\left(I + \frac{\tau}{2}A\right) \mathbf{x}^{k+1} = \left(I - \frac{\tau}{2}A\right) \mathbf{x}^k + \tau \mathbf{b}$$

и общность системы собственных векторов матриц слева и справа, выразим собственные значения оператора перехода B через собственные значения исходной матрицы:

$$\lambda(B) = \frac{1 - \tau \lambda(A)/2}{1 + \tau \lambda(A)/2}, \text{ поэтому } |\lambda(B)| < 1 \text{ для любого } \tau > 0.$$

Для определения τ_{opt} следует решить минимаксную задачу

$$\min_{\tau > 0} \max_{\lambda \in [m/2, M/2]} \frac{|1 - \tau \lambda|}{1 + \tau \lambda}.$$

Функция $f(\lambda) = (1 - \tau \lambda)/(1 + \tau \lambda)$ при $\lambda > 0$ и фиксированном $\tau > 0$ является убывающей, поэтому максимальное значение функции $|f(\lambda)|$ принимает на границе отрезка. Можно убедиться, что минимум по τ имеет место в случае равенства $|f(m/2)| = |f(M/2)|$, которое приводит к уравнению для оптимального параметра

$$\frac{1 - \tau_{\text{opt}} m/2}{1 + \tau_{\text{opt}} m/2} = -\frac{1 - \tau_{\text{opt}} M/2}{1 + \tau_{\text{opt}} M/2} \text{ с решением } \tau_{\text{opt}} = \frac{2}{\sqrt{mM}}.$$

16.7. Используя идею решения **16.6**, запишем условие сходимости метода

$$\max_{\lambda \in [m/2, M/2]} \left| \frac{1 - \tau(1 - \alpha)\lambda}{1 + \tau\alpha\lambda} \right| < 1 \quad \forall \tau > 0.$$

Сделав замену $t = \tau \lambda > 0$, получим неравенство $|1 - t(1 - \alpha)| < 1 + t\alpha$. В силу условия $t > 0$, решением этого неравенства является $\alpha \geq 1/2$.

16.8. Спектральный радиус матрицы перехода в методе Гаусса—Зейделя равен $|a|$. Если начальное приближение таково, что начальная погрешность \mathbf{z}^0 имеет ненулевую вторую координату z_2^0 , то $z_1^k = -a^{2k-1} z_2^0$, $z_2^k = a^{2k} z_2^0$ и метод не сходится при $|a| > 1$.

16.11. Из уравнения для вектора ошибки $\mathbf{z}^k = \mathbf{x} - \mathbf{x}^k$ следует

$$\mathbf{z}^{k+1} = (I - \tau B^{-1}A)\mathbf{z}^k, \quad A\mathbf{z}^{k+1} = (A - \tau AB^{-1}A)\mathbf{z}^k. \quad (2)$$

Вычислим скалярное произведение, используя симметрию A ,

$$(A\mathbf{z}^{k+1}, \mathbf{z}^{k+1}) = (A\mathbf{z}^k, \mathbf{z}^k) - 2\tau \left(\left[B - \frac{\tau}{2} A \right] B^{-1} A\mathbf{z}^k, B^{-1} A\mathbf{z}^k \right). \quad (3)$$

Из первого соотношения (2) имеем $B^{-1}A\mathbf{z}^k = -(\mathbf{z}^{k+1} - \mathbf{z}^k)/\tau$, это позволяет переписать (3) в виде

$$\|\mathbf{z}^{k+1}\|_A^2 - \|\mathbf{z}^k\|_A^2 + \frac{2}{\tau} \left(\left[B - \frac{\tau}{2} A \right] (\mathbf{z}^{k+1} - \mathbf{z}^k), \mathbf{z}^{k+1} - \mathbf{z}^k \right) = 0,$$

где $\|\mathbf{u}\|_A = (A\mathbf{u}, \mathbf{u})^{1/2}$. В силу конечномерности векторного пространства условие $B - \frac{\tau}{2}A > 0$ равносильно условию $B - \frac{\tau}{2}A \geq \varepsilon I$ с некоторым $\varepsilon > 0$, поэтому

$$\|\mathbf{z}^{k+1}\|_A^2 - \|\mathbf{z}^k\|_A^2 + 2\varepsilon\tau^{-1}\|\mathbf{z}^{k+1} - \mathbf{z}^k\|_2^2 \leq 0 \quad \forall k \geq 0.$$

Из этого неравенства следует монотонное убывание ограниченной снизу нулем последовательности $\{\|\mathbf{z}^k\|_A^2\}$, следовательно, сходимость $\|\mathbf{z}^k\|_A$ к некоторой неотрицательной величине d . Переходя к пределу в последнем неравенстве, получаем $\|\mathbf{z}^{k+1} - \mathbf{z}^k\|_2 \rightarrow 0$, поэтому $\lim_{k \rightarrow \infty} \|\mathbf{z}^{k+1} - \mathbf{z}^k\|_2^2 = \lim_{k \rightarrow \infty} \|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2^2 = 0$. Таким образом, последовательность $\{\mathbf{x}^k\}$ сходится к некоторому \mathbf{x}^∞ . Из вида итерационного процесса следует неравенство

$$\|\mathbf{b} - A\mathbf{x}^k\|_2 \leq \frac{1}{\tau} \|B\|_2 \|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2,$$

переходя в котором к пределу убеждаемся, что \mathbf{x}^∞ — решение уравнения $A\mathbf{x} = \mathbf{b}$.

16.12. Использовать утверждение 16.11 при $B = D + \tau L$.

16.13. Пусть $\|B\|_\infty = \max_i \sum_j |b_{ij}| = q < 1$, $q_{1i} = \sum_{j < i} |b_{ij}|$, $q_{2i} = \sum_{j \geq i} |b_{ij}|$. Доказать, что для погрешности итерационного метода $\mathbf{x}^{k+1} = L\mathbf{x}^{k+1} + R\mathbf{x}^k + \mathbf{c}$ справедлива оценка

$$\|\mathbf{z}^{k+1}\|_\infty \leq \max_i \frac{q_{2i}}{1 - q_{1i}} \|\mathbf{z}^k\|_\infty \leq \max_i \frac{q - q_{1i}}{1 - q_{1i}} \|\mathbf{z}^k\|_\infty \leq q \|\mathbf{z}^k\|_\infty.$$

16.14. Спектральный радиус оператора перехода, асимптотическая скорость сходимости и оптимальный параметр таковы:

- 1) $\rho(B) = \cos \pi h$, $R_\infty(B) = \frac{1}{2} \pi^2 h^2$;
- 2) $\rho(B) = \cos^2 \pi h$, $R_\infty(B) = \pi^2 h^2$;

$$3) \rho(B) = \frac{1 - \sin \pi h}{1 + \sin \pi h}, R_\infty(B) = 2\pi h, \omega_{\text{opt}} = \frac{2}{1 + \sin \pi h};$$

$$4) \rho(B) = \frac{1 - \sin \frac{\pi h}{2}}{1 + \sin \frac{\pi h}{2}}, R_\infty(B) = \pi h, \omega_{\text{opt}} = \frac{2}{1 + \sin \frac{\pi h}{2}}.$$

16.15. Матрица имеет строгое диагональное преобладание.

16.16. См. решение **16.1**. Для случаев 1) и 2) имеем условие $|\beta| < |\alpha|$. В случае 3) таких α и β нет, так как существует собственное значение оператора перехода $\lambda = \frac{\alpha^2 + \beta^2}{\alpha\beta}$, модуль которого больше единицы.

16.17. Обозначим $\mathbf{z}^k = \mathbf{x} - \mathbf{x}^k$, $\mathbf{z}^{k+1/2} = \mathbf{x} - \mathbf{x}^{k+1/2}$, где \mathbf{x} — решение системы $A\mathbf{x} = \mathbf{b}$. Тогда

$$\mathbf{z}^{k+1} = (I + \tau A_2)^{-1}(I - \tau A_1)(I + \tau A_1)^{-1}(I - \tau A_2)\mathbf{z}^k \equiv P\mathbf{z}^k.$$

Матрица перехода P подобна матрице

$$B = (I - \tau A_1)(I + \tau A_1)^{-1}(I - \tau A_2)(I + \tau A_2)^{-1}.$$

Коммутирующие матрицы простой структуры A_1 и A_2 имеют общую полную систему собственных векторов. Это дает представления $A_i = QD_iQ^{-1}$, $i = 1, 2$, с диагональными матрицами D_i и совпадение собственных значений матриц A_i и D_i . Отсюда получаем оценку для спектрального радиуса матрицы B

$$\begin{aligned} \rho(B) &= \rho((I - \tau D_1)(I + \tau D_1)^{-1}(I - \tau D_2)(I + \tau D_2)^{-1}) = \\ &= \max_i \left| \frac{1 - \tau \lambda_i(A_1)}{1 + \tau \lambda_i(A_1)} \frac{1 - \tau \lambda_i(A_2)}{1 + \tau \lambda_i(A_2)} \right| \leq \max_{m \leq t \leq M} \left(\frac{1 - \tau t}{1 + \tau t} \right)^2. \end{aligned}$$

Оптимальное значение $\tau_{\text{opt}} = (mM)^{-1/2}$, при этом $\rho(B) \leq \left(\frac{\sqrt{M} - \sqrt{m}}{\sqrt{M} + \sqrt{m}} \right)^2$.

16.19. $\tau_{\text{opt}} = \frac{1}{\sqrt{mM}}$, при этом $\rho(B) \leq \frac{M + m}{(\sqrt{M} + \sqrt{m})^2}$.

16.21. Имеем $A = M - N = M(I - M^{-1}N)$. Если $\rho(M^{-1}N) < 1$, то существует $(I - M^{-1}N)^{-1}$, как следствие, существует $A^{-1} = (I - M^{-1}N)^{-1}M^{-1}$.

16.22. Предположив, что $\rho(M^{-1}N) \geq 1$, выбрать такое начальное приближение \mathbf{x}^0 , что погрешность $\mathbf{z}^0 = \mathbf{x} - \mathbf{x}^0$ пропорциональна собственному вектору матрицы $M^{-1}N$, соответствующему собственному значению λ такому, что $|\lambda| \geq 1$.

16.23. $\rho(B_1) = 1/2$, $\rho(B_2) = 1/4$.

17.1. Собственные значения матриц A и B совпадают, так как равны их характеристические многочлены:

$$\begin{aligned} \det(B - \lambda I) &= \det(S^{-1}AS - \lambda S^{-1}S) = \\ &= \det(S^{-1})\det(A - \lambda I)\det(S) = \det(A - \lambda I). \end{aligned}$$

Условие $A\mathbf{x} = \lambda\mathbf{x}$ равносильно $BS^{-1}\mathbf{x} = \lambda S^{-1}\mathbf{x}$, т. е. собственные векторы \mathbf{x}_A и \mathbf{x}_B связаны соотношением $\mathbf{x}_B = S^{-1}\mathbf{x}_A$.

17.2. Пусть $\{\mathbf{q}_k\}$, $k = 1, \dots, n$, — полная ортонормированная система собственных векторов матрицы A , $\mathbf{x} = \sum_{k=1}^n c_k \mathbf{q}_k$. Тогда $\|A\mathbf{x} - \lambda\mathbf{x}\|_2^2 =$
 $= \sum_{k=1}^n (\lambda_k - \lambda)^2 c_k^2 \geq \min_k (\lambda_k - \lambda)^2 \sum_{k=1}^n c_k^2 = \min_k (\lambda_k - \lambda)^2$.

17.3. Справедливы неравенства

$$\begin{aligned} \lambda_{\min}(A) &= \min_{\|\mathbf{x}\|_2=1} (A\mathbf{x}, \mathbf{x}) \leq (A\mathbf{e}_i, \mathbf{e}_i) = a_{ii}, \\ \lambda_{\max}(A) &= \max_{\|\mathbf{x}\|_2=1} (A\mathbf{x}, \mathbf{x}) \geq (A\mathbf{e}_i, \mathbf{e}_i) = a_{ii}, \end{aligned}$$

где \mathbf{e}_i — вектор с i -й компонентой 1 и остальными компонентами 0.

17.4. Пусть $D = \text{diag}(d_1, \dots, d_n)$, $d_1 = 1$, $d_i = d_{i-1} \sqrt{c_{i-1}/a_i}$, $2 \leq i \leq n$. Тогда $B = DAD^{-1}$ — симметричная матрица, вещественный спектр которой совпадает со спектром подобной матрицы A .

17.10. Использовать то, что AB подобна симметричной матрице $A^{1/2}BA^{1/2}$.

17.11. Воспользоваться указанием к 17.10.

17.12. Воспользоваться тем, что если $SAS^{-1}\mathbf{x} = \lambda\mathbf{x}$, то $S^{-1}\mathbf{x}$ — собственный вектор матрицы A , соответствующий тому же собственному значению λ . Доказать, что из полноты системы векторов $\{\mathbf{x}_i\}$ следует полнота системы $\{S^{-1}\mathbf{x}_i\}$.

17.13. Воспользоваться указанием к 17.10 и 17.12.

17.14. Все собственные значения AB положительны в силу 17.11. Из перестановочности A и B следует симметрия AB , а критерием положительной определенности симметричной матрицы является положительность ее собственных значений.

17.16. Симметричная часть $S = (A + A^T)/2$ матрицы A имеет положительную и строго доминирующую диагональ, поэтому ее собственные значения положительны и S положительно определена, а вместе с ней положительно определена и матрица A .

17.17. $A = \begin{pmatrix} 1 & q & 1/2 \\ q & 1 & q \\ 1/2 & q & 1 \end{pmatrix}$ при $q = \sqrt{(1 + \alpha)/2}$, $\alpha \in (0, 1/2)$.

17.22. Характеристическое уравнение для возмущенной матрицы Уилкинсона имеет вид

$$\det(A - \lambda I) \equiv (20 - \lambda)(19 - \lambda) \dots (1 - \lambda) - 20^{19} \cdot \varepsilon = 0.$$

Свободный член в этом уравнении равен нулю, следовательно, наименьшее собственное значение также равно нулю.

17.23. 1) $\lambda_k = 0,5 \exp \frac{\pi i(k-1)}{10}$, $k = 1, \dots, 20$. 2) $\lambda_k = 0$, $k = 1, \dots, 20$.

17.26. К максимальному по модулю λ из λ_{i_k} , $1 \leq k \leq t$.

17.27. К λ_s , для которого справедливо $|\lambda_s - c| = \max_i |\lambda_i - c|$. Скорость сходимости равна $O(q^{2n})$, где $q = \max_{i \neq s} |\lambda_i - c| / |\lambda_s - c|$. Отсюда следует, что процесс сходится к λ_1 при $c > (\lambda_1 + \lambda_n)/2$ или к λ_n при $c < (\lambda_1 + \lambda_n)/2$.

17.28. Как следует из **17.27**, оптимальное значение c_s для $s = 1, n$ является решением минимаксной задачи $\min_c \max_{i \neq s} |\lambda_i - c| / |\lambda_s - c|$. Так как рассматриваемая функция линейна по λ_i , то ее модуль принимает максимальное значение в граничных точках. Например, при $s = 1$ это соответствует $\min_c \max\{|\lambda_2 - c| / |\lambda_1 - c|, |\lambda_n - c| / |\lambda_1 - c|\}$. Можно показать (например, графически), что оптимальное значение $c_1 = (\lambda_2 + \lambda_n)/2$. Аналогично $c_n = (\lambda_1 + \lambda_{n-1})/2$. Скорость сходимости степенного метода при оптимальном сдвиге зависит от $\lambda_1, \dots, \lambda_n$ и не может стать сколь угодно высокой за счет параметра сдвига.

17.29. К λ_t , для которого справедливо $|\lambda_t - c| = \min_i |\lambda_i - c|$. Скорость сходимости равна $O(q^{2n})$, где $q = |\lambda_t - c| / \min_{i \neq t} |\lambda_i - c|$. Отсюда следует, что процесс в зависимости от значения c может сходиться к любому λ_t . При этом $q \rightarrow 0$, если $c \rightarrow \lambda_t$.

17.30. Пусть $\lambda_i - i$ -е собственное значение и $A\mathbf{q}_i = \lambda_i \mathbf{q}_i$. Из условия $A = A^T$ следует, что собственные векторы образуют базис и можно считать, что $(\mathbf{q}_i, \mathbf{q}_j) = \delta_i^j$. При этом $\lambda_i = \frac{(A\mathbf{q}_i, \mathbf{q}_i)}{(\mathbf{q}_i, \mathbf{q}_i)}$. Представив произвольный вектор \mathbf{x} в виде разложения по собственным векторам, получим требуемый результат.

17.31. Рассмотрим квадратичную по μ функцию $\|(A - \mu I)\mathbf{x}\|_2^2 = (A\mathbf{x}, A\mathbf{x}) - 2\mu(A\mathbf{x}, \mathbf{x}) + \mu^2(\mathbf{x}, \mathbf{x})$, минимум которой достигается в точке $\mu = R_A(\mathbf{x})$.

17.32. Для приближений $\mathbf{x}^k, \mathbf{x}^{k+1}$ метода обратной итерации справедливо

$$R_A(\mathbf{x}^{k+1}) = \frac{(A\mathbf{x}^{k+1}, \mathbf{x}^{k+1})}{(\mathbf{x}^{k+1}, \mathbf{x}^{k+1})} = c + \frac{(\mathbf{x}^k, \mathbf{x}^{k+1})}{(\mathbf{x}^{k+1}, \mathbf{x}^{k+1})} = \lambda^{k+1}.$$

Отсюда и из **17.31**, **17.2** получаем

$$\begin{aligned} 1 &= \|\mathbf{x}^k\|_2 = \|(A - cI)\mathbf{x}^{k+1}\|_2 \geq \|(A - R_A(\mathbf{x}^{k+1})I)\mathbf{x}^{k+1}\|_2 = \\ &= \|(A - \lambda^{k+1}I)\mathbf{x}^{k+1}\|_2 \geq \min_i |\lambda_i - \lambda^{k+1}| \|\mathbf{x}^{k+1}\|_2. \end{aligned}$$

Так как метод сходится к λ_c , то, начиная с некоторого k , имеем $\min_i |\lambda_i - \lambda^{k+1}| = |\lambda_c - \lambda^{k+1}|$. Отсюда следует искомая оценка.

К главе V. Решение нелинейных уравнений

- 18.1.** Функция $\varphi'(x)$ непрерывна и $|\varphi'(z)| < 1$, следовательно, $|\varphi'(x)| < 1$ в некоторой окрестности $Q_\delta = [z - \delta, z + \delta]$. Возьмем произвольную точку $x_0 \in Q_\delta$. По теореме Лагранжа $\varphi(x_0) = \varphi(z) + \varphi'(\xi)(x_0 - z)$, $\xi \in Q_\delta$. Так как $z = \varphi(z)$, то $|\varphi(x_0) - z| = |\varphi'(\xi)(x_0 - z)| < |x_0 - z|$, т. е. функция φ отображает отрезок Q_δ в себя и все итерации будут принадлежать Q_δ , если начальное приближение $x_0 \in Q_\delta$. Тогда для любого n справедливо

$$|x_{n+1} - z| = |\varphi(x_n) - \varphi(z)| = |\varphi'(\xi_n)(x_n - z)| \leq q|x_n - z|,$$

так как $\xi_n \in Q_\delta$, и $|x_{n+1} - z| \leq q^{n+1}|x_0 - z| \rightarrow 0$ при $n \rightarrow \infty$.

- 18.2.** Оценка скорости сходимости итерационного метода для начального приближения $x_0 \in Q_\delta$ (см. решение **18.1**) вытекает из следующего представления погрешности:

$$\begin{aligned} x_{n+1} - z &= \varphi(x_n) - \varphi(z) = \varphi(z + (x_n - z)) - \varphi(z) = \varphi(z) + \\ &+ (x_n - z)\varphi'(z) + \frac{1}{2}(x_n - z)^2 \varphi''(\xi_n) - \varphi(z) = \frac{\varphi''(\xi_n)}{2}(x_n - z)^2, \quad \xi_n \in Q_\delta. \end{aligned}$$

- 18.3.** Пусть $x_0 \in Q_\delta$, тогда

$$|\varphi(x_0) - a| \leq |\varphi(x_0) - \varphi(a)| + |\varphi(a) - a| \leq q|x_0 - a| + (1 - q)\delta \leq \delta.$$

Таким образом, функция $\varphi(x)$ отображает Q_δ в себя и, по условию Липшица, осуществляет сжимающее отображение с константой q .

- 18.4.** Для определенности будем считать, что $f'(x) > 0$. Тогда $0 < m \leq f'(x) \leq M$ на $[a, b]$ с некоторыми постоянными m и M . Заменим исходное уравнение $f(x) = 0$ равносильным $x = \varphi(x)$, $\varphi(x) = x - \lambda f(x)$. При любом фиксированном $\lambda \in (0, 1/M)$ имеем $0 \leq \varphi'(x) = 1 - \lambda f'(x) \leq q < 1$.

- 18.5.** Табличным способом выделим отрезки, на концах которых функция $f(x)$ имеет разные знаки

x	-3	-2	-1	0	1	2	3
$\text{sign} f(x)$	-	+	+	-	+	+	+

Таким образом, корни исходного уравнения лежат на отрезках $[-3, -2]$, $[-1, 0]$ и $[0, 1]$, для каждого из которых построим свой итерационный процесс.

Для $x \in [-3, -2]$ разделим исходное уравнение на x^2 . В результате получим равносильное уравнение $x = \varphi(x)$, $\varphi(x) = \frac{1}{x^2} - 3$. Итерационный процесс для нахождения первого корня: $x_{n+1} = \frac{1}{x_n^2} - 3$.

Поскольку $|\varphi'(x)| = \left| -\frac{2}{x^3} \right| \leq \frac{1}{4} < 1$ для $x \in [-3, -2]$, то сходимость имеет место для всех начальных приближений $x_0 \in [-3, -2]$.

Для двух других отрезков исходное уравнение перепишем в виде $x^2(x + 3) - 1 = 0$. Если $x_0 \in [-1, 0]$, то определим итерационный процесс $x_{n+1} = -\frac{1}{\sqrt{x_n + 3}}$; если $x_0 \in [0, 1]$, то

$x_{n+1} = \frac{1}{\sqrt{x_n + 3}}$. Можно показать, что в процессе итераций соответствующие отрезки отображаются в себя, поэтому (см. 18.1)

сходимость построенных итерационных процессов следует из оценки

$$|\varphi'(x)| = \frac{1}{2} \left| \frac{1}{\sqrt{x+3}} \right|^3 < 1.$$

На каждом рассмотренном отрезке допустим также метод бисекции.

- 18.6.** Уравнение $x^3 - 20x + 1 = 0$ имеет три различных вещественных корня: $z_1 < z_2 < z_3$. В зависимости от выбора начального приближения x_0 итерационный процесс либо расходится, либо сойдется к одному из корней z_i , $i = 1, 2, 3$.

Перепишем формулу итерационного процесса в виде

$$x_{n+1} - x_n = \frac{x_n^3 - 20x_n + 1}{20} = \frac{(x_n - z_1)(x_n - z_2)(x_n - z_3)}{20}.$$

Если $x_n < z_1$, то $x_{n+1} - x_n < 0$ и последовательность x_n монотонно убывает. Это означает расходимость итерационного процесса при $x_0 < z_1$, так как $x_n < x_0 < z_i$, $i = 1, 2, 3$. Аналогично показывается, что при $z_3 < x_0$ выполняются неравенства $z_i < x_n < x_{n+1}$, и метод расходится.

Точки $x_0 = z_1$, $x_0 = z_2$ и $x_0 = z_3$ являются неподвижными, а отображение $x_{n+1} = (x_n^3 + 1)/20$ монотонно. Поэтому для $z_1 < x_0 < z_2$ имеем $z_1 < x_n < x_{n+1} < z_2$. Таким образом, последовательность x_n монотонно возрастает, ограничена сверху и сходится к точке z_2 . Аналогично доказывается, что для $x_0 \in (z_2, z_3)$ последовательность x_n , монотонно убывая, сходится к z_2 .

18.8. Представим метод хорд как частный случай метода простой итерации для уравнения

$$x = \varphi(x), \quad \varphi(x) = x - \frac{f(x)}{f(x) - f(x_0)} (x - x_0).$$

Вблизи корня z уравнения $f(x) = 0$ справедливо представление

$$x_{n+1} - z = (x_n - z) \varphi'(z) + \frac{1}{2} (x_n - z)^2 \varphi''(\xi), \quad \xi \in [x_n, z],$$

где

$$\begin{aligned} \varphi'(z) &= 1 + \frac{f'(z)}{f(x_0)} (z - x_0) = \\ &= \frac{1}{f(x_0)} \left(f(z) - f'(z)(z - x_0) + \frac{f''(\eta)}{2} (z - x_0)^2 + f'(z)(z - x_0) \right) = \\ &= \frac{(z - x_0)^2}{2} \frac{f''(\eta)}{f(x_0)} = \frac{(z - x_0)^2}{2} \frac{f''(\eta)}{f(z) + f'(\xi)(z - x_0)}, \quad \eta, \xi \in [x_0, z]. \end{aligned}$$

Если взять начальное приближение x_1 из окрестности корня, где $|\varphi'(x)| \leq q < 1$, то учитывая **18.1**, получим, что метод хорд имеет линейную скорость сходимости.

18.9. Преобразуем расчетную формулу метода секущих

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n)$$

к виду

$$x_{n+1} - z = x_n - z - \frac{((x_n - z) - (x_{n-1} - z))f(z + (x_n - z))}{f(z + (x_n - z)) - f(z + (x_{n-1} - z))}.$$

Разложим $f(z + (x_n - z))$ и $f(z + (x_{n-1} - z))$ в ряды Тейлора в точке z и подставим в последнюю формулу, учитывая, что $f(z) = 0$. Получим

$$\begin{aligned} x_{n+1} - z &= x_n - z - \frac{(x_n - z)f'(z) + 0,5(x_n - z)^2 f''(z) + \dots}{f'(z) + 0,5((x_n - z) + (x_{n-1} - z))f''(z) + \dots} = \\ &= (x_n - z) \left(1 - \frac{1 + 0,5(x_n - z) \frac{f''(z)}{f'(z)} + \dots}{1 + 0,5(x_n - z) \frac{f''(z)}{f'(z)} + 0,5(x_{n-1} - z) \frac{f''(z)}{f'(z)} + \dots} \right) = \\ &= \frac{1}{2} (x_n - z) (x_{n-1} - z) \frac{f''(z)}{f'(z)} + O((x_n - z)^3). \end{aligned}$$

Опустив члены более высокого порядка малости, для ошибки будем иметь уравнение

$$x_{n+1} - z = C(x_n - z)(x_{n-1} - z), \quad C = \frac{1}{2} \frac{f''(z)}{f'(z)}, \quad f'(z) \neq 0.$$

Предположим, что скорость сходимости определяется соотношением $x_{n+1} - z = A(x_n - z)^m$, в котором значения A и m пока неизвестны. Тогда $x_n - z = A(x_{n-1} - z)^m$, откуда $x_{n-1} - z = A^{-1/m}(x_n - z)^{1/m}$. Подставим эти соотношения в уравнение для ошибки $A(x_n - z)^m = C(x_n - z)A^{-1/m}(x_n - z)^{1/m}$. Отсюда $m = 1 + 1/m$, $1 = CA^{-(1+1/m)}$. Из первого уравнения найдем показатель скорости сходимости метода секущих $m = 0,5(1 + \sqrt{5}) \approx 1,618$. При этом константа A равна $\left(\frac{1}{2} \frac{f''(z)}{f'(z)}\right)^{\frac{1}{m}}$.

18.10. Для корней $|z| > 1$ имеем

$$a_n z^n = -(a_{n-1} z^{n-1} + \dots + a_1 z + a_0),$$

$$|z| \leq \sum_{i=1}^n \frac{|a_{n-i}|}{|a_n|} \frac{1}{|z|^{i-1}} \leq \frac{c}{|a_n|} \frac{|z|}{|z| - 1},$$

откуда $|z| - 1 \leq \frac{c}{|a_n|}$ и $|z| \leq 1 + \frac{c}{|a_n|}$.

Если теперь $|a_0| > 0$, то все корни уравнения отличны от нуля. Делая замену $u = 1/z$, приходим к уравнению $a_0 u^n + a_1 u^{n-1} + \dots + a_n = 0$. Из предыдущей оценки следует $|u| \leq 1 + \frac{b}{|a_0|}$, или $|z| \geq \frac{|a_0|}{b + |a_0|}$.

18.11. Использовать формулу Тейлора для многочлена $f(x)$ степени n

$$f(x) = f(a) + \sum_{k=1}^n \frac{1}{k!} f^{(k)}(a)(x - a)^k.$$

18.12. Применяя идею решения **18.11**, получить, что положительные корни расположены на отрезке $[0, 74, 22]$, а отрицательных корней нет.

18.13. Пусть $\varphi(x) = \sqrt{x+2}$ и $x_0 \geq -2$. Так как $\varphi(x_0) \geq 0$, то сходимость метода достаточно доказать для $x_0 \geq 0$. Рассмотрим отрезок $Q = [0, 2 + x_0]$. Несложно проверить, что $\varphi(x_n) \in Q$ для всех n и $|\varphi'(x)| < 1$ для всех $x \in Q$. Следовательно, приближения $x_{n+1} = \varphi(x_n)$ сходятся к $x = 2$ — единственной неподвижной точке отображения $\varphi(x)$.

- 18.14.** При любом $x_0 \in \mathbb{R}^1$ имеем $x_1 \in [-1, 1]$. Для этого отрезка выполнены условия сходимости метода простой итерации $x_{n+1} = \cos x_n$.
- 18.15.** Уравнение $x = x^2 - 2x + 2$ имеет два корня $z_1 = 1$ и $z_2 = 2$. Пусть $\varphi(x) = x^2 - 2x + 2$, тогда при $x \in (1/2, 3/2)$ имеем $\varphi(x) \in (1/2, 3/2)$, $|\varphi'(x)| < 1$. Поэтому при $x_0 \in (1/2, 3/2)$ приближения сходятся к z_1 . Дальнейший анализ проводим аналогично решению **18.6**, используя, в частности, эквивалентную запись итерационного метода в виде $x_{n+1} - x_n = (x_n - 2)(x_n - 1)$.

Метод сходится к $z_1 = 1$ при $x_0 \in (0, 2)$. Если $x_0 = 0$ или $x_0 = 2$, то метод сходится к $z_2 = 2$. Для остальных начальных приближений метод расходится.

- 18.16.** Пусть $\varphi(x) = 2^{x-1}$, тогда $\varphi'(x) < 1$ при $x \in (-\infty, x^*)$, где $x^* \in (1, 2)$ — решение уравнения $2^{x-1} \ln 2 = 1$. Функция $\varphi(x)$ отображает промежуток $(-\infty, x^*)$ в себя. Таким образом, при $x_0 \in (-\infty, x^*)$ метод сходится к $z = 1$. При $x_0 < 2$ приближения x_n монотонно убывают и при некотором n попадают в промежуток $(-\infty, x^*)$ (ср. с решением **18.13**), поэтому сходятся к $z = 1$.

При $x_0 > 2$ приближения стремятся к ∞ , метод расходится.

- 18.17.** 2) Найти максимальное значение $|\varphi'(x)|$, $\varphi(x) = ae^{-bx^2} + c$, и убедиться, что оно при указанных условиях меньше 1.
- 18.19.** Рассмотрим два примера. 1) $\varphi(x) = \sin x$, $z = 0$, $|\varphi'(z)| = 1$, метод сходится с любого начального приближения. 2) $\varphi(x) = x^2 + x$, $z = 0$, $|\varphi'(z)| = 1$, метод расходится, если $x_0 > 0$.
- 18.20.** Заменой $y = x + a$ преобразуем уравнение к виду $y + \varepsilon \sin(y - a) = 0$ и обозначим $\varphi(y) = -\varepsilon \sin(y - a)$. Функция $\varphi(y)$ преобразует множество \mathbb{R}^1 в отрезок $Q_\varepsilon = [-|\varepsilon|, |\varepsilon|]$, поэтому можно считать, что $y \in Q_\varepsilon$. При $|\varepsilon| < 1$ имеем $|\varphi'(y)| \leq |\varepsilon| < 1$, поэтому $\varphi(y)$ является сжимающим отображением на отрезке Q_ε и имеет единственную неподвижную точку $y^* = z(a, \varepsilon)$.

Пусть теперь $|\varepsilon| = 1$. Тогда для произвольных y_1, y_2 имеем

$$\varphi(y_2) - \varphi(y_1) = \int_{y_1}^{y_2} \varphi'(t) dt = -\varepsilon \int_{y_1}^{y_2} \cos(t - a) dt.$$

Так как $y_1, y_2 \in Q_\varepsilon$, то $\cos(t - a)$ может обратиться в единицу лишь

в конечном числе точек, откуда $\left| \int_{y_1}^{y_2} \cos(t - a) dt \right| \leq q|y_2 - y_1|$, $q < 1$.

Таким образом, $|\varphi(y_2) - \varphi(y_1)| \leq q|y_2 - y_1|$, т. е. $\varphi(y)$ задает сжимающее отображение, поэтому имеет единственную неподвижную точку.

19.1. Значение $\sqrt[p]{a}$ является корнем уравнения $f(x) \equiv x^p - a = 0$. Для этого уравнения метод Ньютона имеет вид

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^p - a}{px_n^{p-1}} = \frac{p-1}{p}x_n + \frac{a}{px_n^{p-1}}.$$

Для $p = 2$ получаем $x_{n+1} = \frac{1}{2} \left(x_n + \frac{a}{x_n} \right)$.

19.2. Рассмотрим метод Ньютона для уравнения $f(x) = 0$ как частный случай метода простой итерации для уравнения $x = \varphi(x)$, $\varphi(x) = x - \frac{f(x)}{f'(x)}$. Производная $\varphi'(x) = \frac{f(x)f''(x)}{(f'(x))^2}$, поэтому, если z — искомый корень, то $\varphi'(z) = 0$. Согласно **18.1** найдется такая окрестность корня Q_δ , что $\varphi(x) \in Q_\delta \forall x \in Q_\delta$. В окрестности точки z справедливо разложение в ряд Тейлора:

$$x_{n+1} - z = \varphi(x_n) - \varphi(z) = \frac{1}{2} (x_n - z)^2 \varphi''(\xi), \quad \xi \in [x_n, z].$$

Следовательно, вблизи корня метод Ньютона имеет квадратичную скорость сходимости.

19.3. Поступая так же, как в **19.2**, получим $x_{n+1} - z = (x_n - z)\varphi'(z) + \frac{1}{2}(x_n - z)^2\varphi''(\xi)$, где $\xi \in [x_n, z]$. Однако в случае $p > 1$ в выражении $\varphi'(x) = \frac{f(x)f''(x)}{(f'(x))^2}$ при $x = z$ содержится неопределенность «нуль на нуль», так как z является также корнем уравнения $f'(x) = 0$. Оценим $\varphi'(x)$. Функция $f(x)$ в окрестности корня z кратности p ведет себя как $a(x-z)^p + o(|x-z|^p)$, где a — константа. Тогда в малой окрестности корня

$$\varphi'(x) = \frac{a(x-z)^p ap(p-1)(x-z)^{p-2}}{a^2 p^2 (x-z)^{2p-2}} + o(1) = \frac{p-1}{p} + o(1) \approx \frac{p-1}{p} < 1.$$

Отсюда следует, что чем выше кратность корня, тем медленнее сходимость.

19.4. Требуемую модификацию будем искать в виде

$$x_{n+1} = x_n - \alpha \frac{f(x_n)}{f'(x_n)},$$

и подберем параметр α так, чтобы имела место квадратичная сходимость. Рассмотрим данную модификацию как специальный случай метода простой итерации $x_{n+1} = \varphi(x_n)$, для которого

выполнено $z = \varphi(z)$, причем вблизи корня

$$\varphi'(x) = 1 - \alpha + \alpha \frac{f(x)f''(x)}{(f'(x))^2} = 1 - \alpha + \alpha \frac{p-1}{p} + o(1) \approx \frac{p-\alpha}{p}.$$

Для обеспечения квадратичной сходимости параметр α надо подобрать таким, чтобы $\varphi'(z) = 0$, т. е. $\alpha = p$.

- 19.5.** Искомое число является корнем уравнения $(ax)^{-1} - 1 = 0$. Для этого уравнения метод Ньютона имеет вид: $x_{n+1} = 2x_n - ax_n^2$, или $x_{n+1} = x_n(2 - ax_n)$.

При $x_0 \geq 2/a$ и $x_0 \leq 0$ сходимости нет, поскольку все x_n остаются неположительными. Если $x_n \in (0, 2/a]$, то $x_{n+1} \in (0, 1/a]$. Пусть $x_n \in (0, 1/a]$. Тогда из равенства $x_{n+1} - x_n = x_n(1 - ax_n)$ получим $x_{n+1} > x_n$, а из условия $x_{n+1} = x_n(2 - ax_n)$ — неравенство $x_{n+1} \leq 1/a$. Так как итерационный процесс имеет две неподвижные точки 0 и $1/a$, то приближения сходятся к $1/a$. Таким образом, сходимость к корню имеет место, если начальное приближение берется из интервала $(0, 2/a)$.

- 19.6.** Для уравнения $g(x) \equiv \frac{f(x)}{f'(x)} = 0$ корень z — простой, следовательно для уравнения $g(x) = 0$ метод Ньютона

$$x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)} = x_n - \frac{f(x_n)f'(x_n)}{(f'(x_n))^2 - f(x_n)f''(x_n)}$$

имеет квадратичный порядок сходимости. В окрестности z функция $f(x) \approx a(x-z)^p$, поэтому

$$g(x) = \frac{f(x)}{f'(x)} \approx \frac{a(x-z)^p}{ap(x-z)^{p-1}} = \frac{1}{p}(x-z).$$

Из двух последовательных итераций для x_1 и x_2 имеем систему приближенных уравнений

$$g(x_1) \approx \frac{1}{p}(x_1 - z), \quad g(x_2) \approx \frac{1}{p}(x_2 - z),$$

откуда следует оценка для кратности p корня z :

$$p \approx \frac{x_2 - x_1}{g(x_2) - g(x_1)}.$$

- 19.11.** Обратная к f функция имеет вид $F(y) = (a+y)^{1/p}$, а производные F выражаются формулой

$$F^{(k)}(y) = x^{1-kp} \prod_{j=0}^{k-1} \left(\frac{1}{p} - j \right).$$

Таким образом, имеем

$$\varphi_m(x) = x + x \sum_{k=1}^{m-1} \frac{1}{k!} \left(\frac{a - x^p}{px^p} \right)^k \prod_{j=0}^{k-1} (1 - jp).$$

В частности, $\varphi_2(x) = (x/p)(p - 1 + a/x^p)$. При $p = 2$ получаем формулу Ньютона—Херона $x_{n+1} = \varphi_2(x_n)$ для приближенного вычисления квадратных корней.

Если $p = -1$, то $\varphi_m(x) = x \sum_{k=0}^{m-1} (1 - ax)^k$. В этом случае итерационный процесс $x_{n+1} = \varphi_m(x_n)$ при $|1 - ax| < 1$ сходится к решению уравнения $x - 1/a = 0$. Данный метод позволяет находить значение $1/a$ с произвольной точностью без использования операции деления.

19.13. Порядок сходимости $m = 3$.

19.15. Порядок сходимости $m = 3$.

19.22. Обозначим области сходимости метода Ньютона

$$x_{n+1} = \varphi(x_n), \quad \varphi(x) = \frac{2x^3}{3x^2 - 1}$$

к корням $z = -1, 0, +1$ через X_-, X_0, X_+ соответственно. Кроме того, определим последовательности точек $\{x_n^\pm\}$ для $n \geq 0$ следующими условиями:

$$\varphi(x_{n+1}^\pm) = x_n^\pm, \quad x_0^\pm = \pm \frac{1}{\sqrt{3}},$$

для элементов которых справедливы неравенства

$$\begin{aligned} -\frac{1}{\sqrt{3}} = x_0^- < x_1^+ < x_2^- < \dots < -\frac{1}{\sqrt{5}} < 0 < \\ < \frac{1}{\sqrt{5}} < \dots < x_2^+ < x_1^- < x_0^+ = \frac{1}{\sqrt{3}} \end{aligned}$$

и существуют пределы

$$\lim_{k \rightarrow \infty} x_{2k}^- = \lim_{k \rightarrow \infty} x_{2k-1}^+ = -\frac{1}{\sqrt{5}}, \quad \lim_{k \rightarrow \infty} x_{2k}^+ = \lim_{k \rightarrow \infty} x_{2k-1}^- = \frac{1}{\sqrt{5}}.$$

Тогда

$$X_- = (-\infty, x_0^-) \cup \bigcup_{k=1}^{\infty} \left[(x_{2k-1}^+, x_{2k}^-) \cup (x_{2k-1}^-, x_{2(k-1)}^+) \right],$$

$$X_0 = \left(-\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right),$$

$$X_+ = (x_0^+, \infty) \cup \bigcup_{k=1}^{\infty} \left[(x_{2(k-1)}^-, x_{2k-1}^+) \cup (x_{2k}^+, x_{2k-1}^-) \right].$$

Кроме того, если $x_0 = x_n^\pm$, $n \geq 0$, то метод не определен, а при $x_0 = \pm \frac{1}{\sqrt{5}}$ имеем $x_1 = \mp \frac{1}{\sqrt{5}}$, т. е. метод «заикливается».

Таким образом, области сходимости к корням $z = \pm 1$ являются объединениями перемежающихся открытых интервалов, разделенных точками заикливания метода.

К главе VI. Разностные уравнения

20.1. Корнями характеристического уравнения $b\mu^2 - c\mu + a = 0$ являются числа

$$\mu_{1,2} = \frac{c \pm \sqrt{D}}{2b}, \quad D = c^2 - 4ab.$$

Рассмотрим следующие три случая:

а) $D > 0$, $\mu_1 \neq \mu_2$ — вещественные корни. Тогда $y_k = C_1\mu_1^k + C_2\mu_2^k$.

б) $D < 0$, $\mu_{1,2} = \rho e^{\pm i\varphi}$ — комплексно сопряженные корни, $\rho = \sqrt{\frac{a}{b}}$,

$$\varphi = \left\{ \begin{array}{l} \operatorname{arctg} \frac{\sqrt{|D|}}{c} \quad \text{при } c/b > 0; \\ \pi - \operatorname{arctg} \frac{\sqrt{|D|}}{c} \quad \text{при } c/b < 0; \end{array} \right.$$

$$\frac{\pi}{2} \quad \text{при } c = 0 \}.$$

Тогда $y_k = \rho^k (C_1 \cos k\varphi + C_2 \sin k\varphi)$. Так записывают общее действительное решение, а для комплексного решения можно использовать формулу из пункта а).

в) $D = 0$, $\mu_1 = \mu_2 = \mu$ — кратные корни. Тогда $y_k = C_1\mu^k + C_2k\mu^k$.

20.2. $y_k = (\sqrt{2})^k (C_1 \sin k\varphi + C_2 \cos k\varphi)$, $\varphi = \operatorname{arctg} \sqrt{7}$.

20.3. Да, так как характеристический многочлен второго уравнения делится на характеристический многочлен первого без остатка.

20.4. Для определителя $D_k = \det A_k$ справедливо разностное уравнение $D_k = \frac{a-1}{a_1} D_{k-1}$. Соответствующее утверждение можно обобщить на случай разностных уравнений более высокого порядка. Равенство нулю определителя означает линейную зависимость соответствующих частных решений.

20.5. Характеристическое уравнение имеет вид $(\mu^2 + \mu + 1)^2 = 0$. Отсюда

$$y_k = \frac{2(k-1)}{\sqrt{3}} \sin \frac{2\pi k}{3}.$$

20.6. Формула для чисел Фибоначчи имеет вид

$$f_k = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^k - \left(\frac{1-\sqrt{5}}{2} \right)^k \right].$$

20.7. Разлагая определитель Δ_k по первой строке, получим следующую разностную задачу: $\Delta_k = b\Delta_{k-1} - ac\Delta_{k-2}$, $\Delta_1 = b$, $\Delta_2 = b^2 - ac$. В случае ненулевого дискриминанта $D = b^2 - 4ac$ имеем

$$\Delta_k = \frac{(b + \sqrt{D})^{k+1} - (b - \sqrt{D})^{k+1}}{2^{k+1}\sqrt{D}},$$

а в случае нулевого — $\Delta_k = \left(\frac{b}{2}\right)^k (1 + k)$. Данное решение можно получить из вида Δ_k для $D \neq 0$ предельным переходом при $D \rightarrow 0$.

20.8. Можно показать, что $I_{k-1} + I_{k+1} = 2I_k \cos \alpha$, $I_0 = 0$, $I_1 = 1$, откуда для $0 < \alpha < \pi$ следует формула $I_k(\alpha) = \frac{\sin k\alpha}{\sin \alpha}$. Для оставшихся значений корни характеристического уравнения кратные, поэтому формула имеет другой вид.

20.9. При $p < 1$ положим $p = \cos \alpha$ ($\alpha \neq 0$), тогда $y_k = C_1 \cos k\alpha + C_2 \sin k\alpha$. При $p > 1$ положим $p = \cosh \alpha$, тогда $y_k = C_1 \cosh k\alpha + C_2 \sinh k\alpha$. При $p = 1$ имеем $y_k = C_1 + C_2 k$.

20.10. Если μ — корень характеристического уравнения $\mu^2 - 2\lambda\mu + 1 = 0$, то $1/\mu$ — другой его корень. Ограниченность решений разностного уравнения равносильна следующему условию: корни характеристического уравнения различны и лежат на единичной окружности. Поэтому (см. **20.9**) решение ограничено, если только $\mu_{1,2} = \cos \alpha \pm i \sin \alpha$, $\alpha \neq 0, \pi$. В этом случае $\lambda = \cos \alpha$.

20.12. Если N не кратное 3, то $y_k = \sin \frac{(N-k)\pi}{3} / \sin \frac{N\pi}{3}$. В противном случае решения не существует.

20.13. $y_k = (\sqrt{20})^k (C_1 \sin k\varphi + C_2 \cos k\varphi)$, $\varphi = \arctg(1/2)$.

20.14. $y_k = (\sqrt{2})^k (C_1 \sin(k\pi/4) + C_2 \cos(k\pi/4))$.

20.15. $y_k = (\sqrt{26})^k (C_1 \sin k\varphi + C_2 \cos k\varphi)$, $\varphi = \pi + \arctg(1/5)$.

20.16. $y_k = (\sqrt{13})^k (C_1 \sin k\varphi + C_2 \cos k\varphi)$, $\varphi = \pi + \arctg(3/2)$.

20.17. $y_k = (-2)^k (1 - 3k)$.

20.18. $y_k = (-1)^k (5 - 3 \cdot 2^k)$.

20.19. $y_k = 2 \cos(\pi k/2) + \sin(\pi k/2)$.

20.20. $y_k = (-1)^k + 2^{k+1} + 3^{k+1}$.

21.1. Корень характеристического уравнения $\mu = 2$, поэтому частное решение ищем в виде $y_k^1 = bk^2 + ck + d$. Подставив его в уравнение, получим

$$2bk^2 + 2ck + 2d - [b(k+1)^2 + c(k+1) + d] = 1 + 2k - k^2 \quad \forall k.$$

Сравнение коэффициентов при линейно независимых функциях $\varphi_1(k) = 1$, $\varphi_2(k) = k$ и $\varphi_3(k) = k^2$ дает $b = -1$, $c = 0$ и $d = 0$. Следовательно, $y_k^1 = -k^2$.

- 21.2.** Корень характеристического уравнения $\mu = 2$, поэтому частное решение ищем в виде $y_k^1 = 2^k k(bk + c)$. Подставив его в уравнение, получим

$$2^{k+1}(bk^2 + ck) - 2^{k+1}(b(k+1)^2 + c(k+1)) = k2^k \quad \forall k.$$

Сравнение коэффициентов при линейно независимых функциях $\varphi_1(k) = 2^k$, $\varphi_2(k) = k2^k$ и $\varphi_3(k) = k^2 2^k$ дает $b = -1/4$, $c = 1/4$. Следовательно, $y_k^1 = 2^{k-2}(k - k^2)$.

- 21.3.** Корень характеристического уравнения $\mu = 2$, поэтому частное решение ищем в виде $y_k^1 = c \sin k + d \cos k$. Подставим его в уравнение:

$$2(c \sin k + d \cos k) - (c \sin(k+1) + d \cos(k+1)) = \sin k \quad \forall k.$$

Так как $\sin(k+1) = \sin k \cos 1 + \cos k \sin 1$ и $\cos(k+1) = \cos k \cos 1 - \sin k \sin 1$, то совпадение коэффициентов при линейно независимых функциях $\varphi_1(k) = \sin k$ и $\varphi_2(k) = \cos k$ приводит к равенствам

$$(2 - \cos 1)c + d \sin 1 = 1, \quad (2 - \cos 1)d - c \sin 1 = 0,$$

откуда $c = \frac{2 - \cos 1}{5 - 4 \cos 1}$, $d = \frac{\sin 1}{5 - 4 \cos 1}$.

- 21.4.** Корень характеристического уравнения $\mu = b$, поэтому возможны два случая: $y_k = \frac{a-b-1}{a-b} b^k + \frac{1}{a-b} a^k$ при $b \neq a$ и $y_k = a^{k-1}(a+k)$ при $b = a$.

- 21.5.** Преобразуем уравнение к виду $y_{k+1} - y_{k-1} = -\frac{1}{2} \left(\frac{1}{k+1} - \frac{1}{k-1} \right)$.

Отсюда находим частное решение $y_k^1 = -\frac{1}{2k}$. Окончательно имеем $y_k = \frac{1}{8} (3 - (-1)^k) - \frac{1}{2k}$.

- 21.6.** При $k = 0$ из уравнения получим $y_1 = 1$. Запишем исходное уравнение в виде $y_{k+1} = (2^k(k-1)! + y_k)k$. Замена $y_k = z_k(k-1)!$ дает разностную задачу для z_k :

$$z_{k+1} - z_k = 2^k, \quad z_1 = 1.$$

Ее решение $z_k = 2^k - 1$, откуда $y_k = (k-1)!(2^k - 1)$.

- 21.7.** Исходное уравнение эквивалентно уравнению $y_{k+1} = \frac{1}{1/y_k + 1}$.

Заменяя $y_k = 1/z_k$, получаем $z_k = k+1$, откуда $y_k = \frac{1}{k+1}$.

21.8. Положим $y_k = \frac{u_k}{v_k}$, тогда $\frac{u_{k+1}}{v_{k+1}} = \frac{au_k + bv_k}{cu_k + dv_k}$. Рассмотрим систему

$$u_{k+1} = a u_k + b v_k, \quad v_{k+1} = c u_k + d v_k,$$

из которой следует уравнение второго порядка $v_{k+2} = (a + d)v_{k+1} - (ad - bc)v_k$. Его характеристическое уравнение $\mu^2 - \mu(a + d) + ad - bc = 0$ имеет корни

$$\mu_{1,2} = \frac{a + d}{2} \pm \sqrt{\frac{(a + d)^2}{4} - (ad - bc)}.$$

Из условия на коэффициенты следует, что дискриминант больше нуля, значит вещественные корни различны и $v_k = A \mu_1^k + B \mu_2^k$. Из второго уравнения системы получим

$$u_k = \frac{v_{k+1} - d v_k}{c} = \frac{1}{c} \left[A \mu_1^k (\mu_1 - d) + B \mu_2^k (\mu_2 - d) \right].$$

В итоге

$$y_k = \frac{u_k}{v_k} = \left(\mu_1 - d + K \left(\frac{\mu_2}{\mu_1} \right)^k (\mu_2 - d) \right) c^{-1} \left(1 + K \left(\frac{\mu_2}{\mu_1} \right)^k \right)^{-1},$$

$$K = B/A.$$

Константа K определяется из начального условия $y_0 = 1$:

$$K = -(\mu_1 - (c + d))(\mu_2 - (c + d))^{-1}.$$

21.9. $y_k = -2^k$.

21.13. $\mu_1 = 1, \mu_2 = -1, \mu_3 = 2, \mu_4 = 1/2, y_k = 8 \cdot (-1)^{k-1} + 8 \cdot 2^{-k} - k$.

21.14. Решение удовлетворяет разностному уравнению $S_{k+1} - S_k = a_{k+1}$ и начальному условию $S_0 = a_0$.

21.15. Преобразуем исходное уравнение к виду $y_{k+1}(1 - y_k) = 1 - y_{k+1}$. Сделаем замену $y_k = 1 - 1/z_k$, получим $z_{k+1} - z_k = 1$, откуда $y_k = \frac{y_0 + k(1 - y_0)}{1 + k(1 - y_0)} = \frac{1 + k}{2 + k}$.

21.16. Преобразуем исходное уравнение к виду $y_{k+1} - 1 = \frac{y_k - 1}{y_k}$. Сделаем замену $y_k = 1 + \frac{1}{z_k}$, получим $z_{k+1} = z_k + 1$, откуда $y_k = \frac{k + 2}{k + 1}$.

21.17. $y_k = \sqrt{k + C}, C \geq 0$.

21.18. Прологарифмируем обе части уравнения и выполним замену $z_k = \log y_k$. Получим уравнение $2z_{k+1} - z_k = \log 2$, общее решение которого $z_k = C_1 2^{-k} + \log 2$, следовательно, $y_k = 2C^{(1/2)^k}$.

21.19. $y_k = e^{-k^2/2}$ (см. решение 21.18).

21.21. $y_k = 4k2^{-k}$.

21.22. $y_k = \frac{k}{7}4^k$.

21.23. $y_k = \frac{k}{19}3^{-k}$.

21.24. $y_k = -k2^k$.

21.25. $y_k = C_1 2^k + C_2 2^{-k} + \frac{\cos k}{2 \cos 1 - 5/2}$.

21.26. $y_k = C_1 + C_2 \left(\frac{1 + \sqrt{17}}{2}\right)^k + C_3 \left(\frac{1 - \sqrt{17}}{2}\right)^k - \frac{3}{16}k - \frac{1}{8}k^2$.

21.27. $y_k = C_1 + C_2 k + C_3 (-3)^k + \frac{1}{8}k^2$.

21.28. $y_k = C_1 (-2)^k + C_2 4^k - \frac{7 \cos 1 + 2}{D} \sin k - \frac{9 \sin 1}{D} \cos k$, где $D = (2 + 7 \cos 1)^2 + (9 \sin 1)^2$.

22.1. Обозначим искомое фундаментальное решение через G_k . Для определения G_k имеем три группы уравнений:

$$\begin{cases} a G_k + b G_{k+1} = 0 & \text{при } k \leq -1, \\ a G_0 + b G_1 = 1 & \text{при } k = 0, \\ a G_k + b G_{k+1} = 0 & \text{при } k \geq 1. \end{cases}$$

При $k \leq 0$ возьмем $G_k = 0$. Тогда все уравнения первой группы выполнены, из уравнений второй группы следует $G_1 = 1/b$, а общее решение третьей группы уравнений имеет вид $G_k = C \mu^k$, где $\mu = -a/b$. Определив константу C из G_1 , получим частное решение неоднородного уравнения

$$G_k^1 = \left\{ 0 \text{ при } k \leq 0; -\frac{1}{a} \left(-\frac{a}{b}\right)^k \text{ при } k \geq 1 \right\}.$$

Сложив полученное частное решение с общим решением $A \left(-\frac{a}{b}\right)^k$ однородного уравнения, получим фундаментальное решение

$$G_k = \left\{ A \left(-\frac{a}{b}\right)^k \text{ при } k \leq 0; \left(A - \frac{1}{a}\right) \left(-\frac{a}{b}\right)^k \text{ при } k \geq 1 \right\}.$$

Условие ограниченности выражается в виде зависимости постоянной A от величины $|a/b|$: $A = 0$ при $|a/b| < 1$, $A = \frac{1}{a}$ при $|a/b| > 1$ и A любое при $|a/b| = 1$.

22.2. Рассмотрим случай $|a/b| > 1$. Из **22.1** следует, что

$$G_{k-n} = \left\{ \frac{1}{a} \left(-\frac{a}{b}\right)^{k-n} \text{ при } k \leq n; \ 0 \text{ при } k \geq n+1 \right\}.$$

Каждый член ряда может быть оценен сверху членом сходящейся геометрической прогрессии

$$\left| \frac{1}{a} \left(-\frac{a}{b}\right)^{k-n} f_n \right| < \frac{F}{|a|} \left| \frac{b}{a} \right|^{n-k},$$

поэтому ряд абсолютно сходится. Кроме того, сумма ряда является частным решением заданного уравнения, так как

$$a \sum_{n=-\infty}^{\infty} G_{k-n} f_n + b \sum_{n=-\infty}^{\infty} G_{k+1-n} f_n = \sum_{n=-\infty}^{\infty} (a G_{k-n} + b G_{k+1-n}) f_n = f_k.$$

Для этого решения верна оценка

$$|y_k^1| \leq \frac{F}{|a|} \sum_{n=k}^{\infty} \left| \frac{b}{a} \right|^{n-k} = \frac{F}{|a| - |b|},$$

т. е. полученное частное решение является ограниченным.

22.3. Для определения G_k имеем три группы уравнений:

$$\begin{cases} G_{k-1} - 2G_k + G_{k+1} = 0 & \text{при } k \leq -1, \\ G_{-1} - 2G_0 + G_1 = 1 & \text{при } k = 0, \\ G_{k-1} - 2G_k + G_{k+1} = 0 & \text{при } k \geq 1. \end{cases}$$

Общие решения первой и третьей групп имеют одинаковый вид, отличающийся только постоянными: $C_1^- + C_2^- k$ при $k \leq 0$ и $C_1^+ + C_2^+ k$ при $k \geq 0$. Так как G_0 входит во все три группы уравнений, то из полученных соотношений имеем $G_0 = C_1^- = C_1^+ = A$. Теперь воспользуемся уравнением при $k = 0$ для установления связи между C_2^- и C_2^+ : $(A - C_2^-) - 2A + (A + C_2^+) = 1$. Отсюда $C_2^- = B$, $C_2^+ = 1 + B$. Окончательное выражение для фундаментального решения

$$G_k = \{A + Bk \text{ при } k \leq 0; \ A + (B + 1)k \text{ при } k \geq 0\}.$$

Ограниченное решение не существует, поскольку B не может одновременно быть равным 0 и -1 .

$$\mathbf{22.4.} \quad G_k = \begin{cases} A \cos \frac{k\pi}{3} + \left[B + \frac{2\sqrt{3}}{3}(1 - 2A) \right] \sin \frac{k\pi}{3} & \text{при } k \geq 0, \\ A \cos \frac{k\pi}{3} + B \sin \frac{k\pi}{3} & \text{при } k \leq 0. \end{cases}$$

$$22.5. G_k = \begin{cases} A 2^{-k} & \text{при } k \geq 0, \\ A 2^k & \text{при } k \leq 0, \end{cases} \quad A = -\frac{2}{3}.$$

$$22.6. G_k = \begin{cases} 2^k - 3^k & \text{при } k \leq 0, \\ 0 & \text{при } k \geq 0. \end{cases}$$

$$22.7. G_k = \begin{cases} C 6^k & \text{при } k \leq 0, \\ C 2^{-k} & \text{при } k \geq 0, \end{cases} \quad C = -\frac{2}{11}.$$

$$22.8. G_k = \begin{cases} 0 & \text{при } k \leq 0, \\ 4(2^{-k} - 4^{-k}) & \text{при } k \geq 0. \end{cases}$$

$$22.9. G_k = \begin{cases} 0 & \text{при } k \geq 0, \\ C((-3)^k - 4^k) & \text{при } k \leq 0, \end{cases} \quad C = \frac{1}{7}.$$

22.10. Решение удобно записать в виде

$$y_k = \alpha + k \frac{\beta - \alpha}{N} + \sum_{i=1}^{N-1} G_{k,i} f_i,$$

где $G_{k,i}$ (функция Грина) представима формулой

$$G_{k,i} = \begin{cases} \frac{i}{N}(k - N) & \text{при } k \geq i, \\ \frac{k}{N}(i - N) & \text{при } k \leq i, \end{cases} \quad i = 1, \dots, N-1, \quad k = 0, \dots, N.$$

22.11. Перепишем разностное уравнение в виде

$$y_{k+1} + 2h\lambda y_k - y_{k-1} = 0.$$

Его характеристическое уравнение $\mu^2 + 2h\lambda\mu - 1 = 0$ имеет корни: $\mu_1 = -h\lambda + \sqrt{1 + h^2\lambda^2}$ и $\mu_2 = -h\lambda - \sqrt{1 + h^2\lambda^2}$. Можно показать, что при $\mu_1 = \mu_2$ существует только тривиальное решение $y_k \equiv 0$, поэтому общее решение разностного уравнения

$$y_k = C_1 \mu_1^k + C_2 \mu_2^k.$$

Константы C_1 и C_2 определяются из системы $C_1 + C_2 = 0$, $C_1 \mu_1^N + C_2 \mu_2^N = 0$, откуда $C_2 = -C_1$ и $C_1(\mu_1^N - \mu_2^N) = 0$, т. е. нетривиальное решение разностной задачи существует тогда и только тогда, когда $\mu_1^N = \mu_2^N$. Следовательно,

$$\frac{\mu_1}{\mu_2} = \exp\left(i \frac{2\pi m}{N}\right), \quad m = 0, 1, \dots, N-1.$$

Так как $\mu_1 \mu_2 = -1$, то $\mu_1^2 = -\exp\left(i \frac{2\pi m}{N}\right)$, откуда $\mu_1 = i \exp\left(i \frac{\pi m}{N}\right)$, $\mu_2 = i \exp\left(-i \frac{\pi m}{N}\right)$. Поскольку $\mu_1 + \mu_2 = -2h\lambda = i \left(\exp\left(i \frac{\pi m}{N}\right) + \right.$

+ $\exp\left(-i\frac{\pi m}{N}\right) = 2i \cos \frac{\pi m}{N}$, то $\lambda^{(m)} = -\frac{i}{h} \cos \frac{\pi m}{N}$, $m = 0, 1, \dots, N-1$. Соответствующие решения исходной задачи таковы:

$$\begin{aligned} y_k^{(m)} &= C_1 (\mu_1^k - \mu_2^k) = C_1 i^k \left(\exp\left(i\frac{\pi km}{N}\right) - \exp\left(-i\frac{\pi km}{N}\right) \right) = \\ &= C_1 i^k 2i \sin \frac{\pi km}{N} = C i^k \sin \frac{\pi km}{N}, \quad m = 0, 1, 2, \dots, N-1. \end{aligned}$$

При $m = 0$ имеем $y_k^{(0)} \equiv 0$, поэтому решение $(\lambda^{(0)}, y_k^{(0)})$ следует отбросить.

22.12. Характеристическое уравнение разностной задачи имеет вид

$$\mu^2 - (2 - h^2 \lambda) \mu + 1 = 0.$$

Если корни характеристического уравнения вещественные, то разностная задача имеет только тривиальное решение. Действительно, если $\mu_1 \neq \mu_2$ — вещественные корни, то общее решение имеет вид $y_k = C_1 \mu_1^k + C_2 \mu_2^k$, а для определения C_1 и C_2 из краевых условий имеем систему $C_1 + C_2 = 0$, $C_1 \mu_1^N + C_2 \mu_2^N = 0$, из которой следует $C_1 \mu_1^N - C_1 \mu_2^N = 0$. Так как $\mu_1 \neq \mu_2$, то $C_1 = C_2 = 0$, т. е. общее решение является нулевым. Аналогично рассматривается случай равных вещественных корней.

Поэтому следует рассмотреть случай комплексно сопряженных корней $\mu_{1,2} = \cos \varphi \pm i \sin \varphi$. В этом случае общее решение разностной задачи представляется в виде $y_k = C_1 \cos k\varphi + C_2 \sin k\varphi$. Из краевых условий получаем $C_1 = 0$ и $\sin N\varphi = 0$. Отсюда $\varphi = \frac{\pi m}{N}$, $m = 0, \pm 1, \pm 2, \dots$. Так как $\mu_1 + \mu_2 = 2 - h^2 \lambda$, то $\cos \varphi = 1 - h^2 \lambda/2$. Следовательно,

$$\lambda^{(m)} = \frac{2}{h^2} \left(1 - \cos \frac{\pi m}{N} \right) = \frac{4}{h^2} \sin^2 \frac{\pi m}{2N}, \quad m = 0, 1, 2, \dots, N-1.$$

Все собственные значения различны. Собственные функции имеют вид

$$y_k^{(m)} = C_2 \sin \frac{\pi km}{N}, \quad m = 0, 1, 2, \dots, N-1.$$

При $m = 0$ имеем $y_k^{(0)} \equiv 0$, поэтому решение $(\lambda^{(0)}, y_k^{(0)})$ следует отбросить.

22.13. Введем обозначение $p = 1 - \lambda h^2/2$ и перепишем исходную задачу в виде:

$$\begin{aligned} y_{k+1} - 2py_k + y_{k-1} &= 0, \quad 1 \leq k \leq N-1, \\ y_1 - py_0 &= 0, \quad y_{N-1} - py_N = 0. \end{aligned}$$

Корни соответствующего характеристического уравнения $\mu_{1,2} = p \pm \sqrt{p^2 - 1}$. Отметим полезные соотношения:

$$\mu_1 \mu_2 = 1, \quad p = \frac{\mu_1 + \mu_2}{2}.$$

В случае различных (не обязательно вещественных) корней общее решение разностного уравнения имеет вид $y_k = C_1 \mu_1^k + C_2 \mu_2^k$. Из краевого условия при $k = 0$ получим $C_1 \mu_1 + C_2 \mu_2 - p(C_1 + C_2) = 0$. Поскольку $p = (\mu_1 + \mu_2)/2$, то $C_1 = C_2 = \tilde{C} \neq 0$. Это дает возможность переписать оставшееся краевое условие в удобной форме

$$\mu_1^{N-1} + \mu_2^{N-1} - \frac{\mu_1 + \mu_2}{2} (\mu_1^N + \mu_2^N) = 0.$$

Так как $\mu_1 \mu_2 = 1$, то справедливо равенство $\mu_1^{N-1} + \mu_2^{N-1} - \mu_1^{N+1} - \mu_2^{N+1} = 0$, или $\mu_1^{N-1} (1 - \mu_1^2) + \mu_2^{N-1} (1 - \mu_2^2) = 0$. Отсюда $\mu_1^N (\mu_2 - \mu_1) + \mu_2^N (\mu_1 - \mu_2) = 0$. В силу предположения о неравенстве корней, имеем $\mu_1^N / \mu_2^N = 1$, что дает $\mu_1^{2N} = 1$, или

$$\mu_{1,2} = \cos \frac{\pi m}{N} \pm i \sin \frac{\pi m}{N}, \quad m = 1, 2, \dots, N-1.$$

Отсюда получим $p = \cos \frac{\pi m}{N}$ и формулы для собственных значений

$$\lambda^{(m)} = \frac{4}{h^2} \sin^2 \frac{\pi m}{2N}, \quad m = 1, 2, \dots, N-1.$$

Собственные функции

$$y_k^{(m)} = \tilde{C} \left[\exp\left(i \frac{\pi m k}{N}\right) + \exp\left(-i \frac{\pi m k}{N}\right) \right] = C \cos \frac{\pi m k}{N}.$$

Осталось рассмотреть случай кратных корней: $\mu_1 = \mu_2 = p$. Возможны два случая: $p = \pm 1$, так как $\mu_1 \mu_2 = 1$. При этом соответствующие собственные значения равны $\lambda = 2(1-p)/h^2$. Их удобно включить в полученную ранее общую формулу, расширив границы индекса m от нуля до N , т. е. $\lambda^{(0)} = 0$ ($p = 1$), $\lambda^{(N)} = \frac{4}{h^2}$ ($p = -1$). Аналогично поступим и с соответствующими собственными функциями.

Окончательно имеем:

$$\lambda^{(m)} = \frac{4}{h^2} \sin^2 \frac{\pi m}{2N}, \quad m = 0, 1, \dots, N,$$

$$y_k^{(m)} = C \cos \frac{\pi m k}{N}, \quad k = 0, 1, \dots, N.$$

$$22.14. \lambda^{(m)} = \frac{4}{h^2} \sin^2 \frac{\pi(2m-1)}{4N}, \quad m = 1, 2, \dots, N,$$

$$y_k^{(m)} = C \sin \frac{\pi(2m-1)k}{2N}, \quad k = 0, 1, \dots, N.$$

$$22.15. \lambda^{(m)} = \frac{4}{h^2} \sin^2 \frac{\pi(2m-1)}{4N}, \quad m = 1, 2, \dots, N,$$

$$y_k^{(m)} = C \cos \frac{\pi(2m-1)k}{2N}, \quad k = 0, 1, \dots, N.$$

$$22.16. \lambda^{(m)} = \frac{4}{h^2} \sin^2 \frac{\pi m}{N}, \quad m = 0, 1, \dots, N-1,$$

$$y_k^{(m)} = C_1^{(m)} \cos \frac{2\pi m k}{N} + C_2^{(m)} \sin \frac{2\pi m k}{N}, \quad k - \text{целое.}$$

$$22.17. \lambda^{(m)} = \frac{4}{h^2} \sin^2 \frac{\pi(m-1)}{2(N-1)}, \quad m = 1, \dots, N-1,$$

$$y_k^{(m)} = C \cos \frac{\pi(m-1)(k-1/2)}{N-1}, \quad k = 0, 1, \dots, N.$$

$$22.18. \lambda^{(m)} = \frac{4}{h^2} \sin^2 \frac{\pi m}{2(N-1)}, \quad m = 1, \dots, N-1,$$

$$y_k^{(m)} = C \sin \frac{\pi m(k-1/2)}{N-1}, \quad k = 0, 1, \dots, N.$$

$$22.19. \lambda^{(m)} = \frac{4}{h^2} \sin^2 \frac{\pi(2m-1)}{2(2N-1)}, \quad m = 1, \dots, N-1,$$

$$y_k^{(m)} = C \sin \frac{\pi(2m-1)(N-k)}{2N-1}, \quad k = 0, 1, \dots, N.$$

$$22.20. \lambda^{(m)} = \frac{4}{h^2} \sin^2 \frac{\pi(2m-1)}{2(2N-1)}, \quad m = 1, \dots, N-1,$$

$$y_k^{(m)} = C \sin \frac{\pi(2m-1)k}{2N-1}, \quad k = 0, 1, \dots, N.$$

$$22.21. \lambda^{(m)} = -\frac{i}{h} \cos \frac{\pi m}{N-1}, \quad m = 1, \dots, N-1,$$

$$y_k^{(m)} = C_1^k \left[\sin \frac{\pi m k}{N-1} - i \sin \frac{\pi m(k-1)}{N-1} \right], \quad k = 0, 1, \dots, N.$$

К главе VII. Решение дифференциальных уравнений

23.1. 1) Нет, 2) да.

23.2. 1) $O(h)$; 2) $O(h^2)$.

23.3. Порядок аппроксимации равен $O(h^2)$. Ответ не изменится, если использовать следующие аппроксимации для коэффициента и правой части уравнения: $a_i = (a(x_{i+1}) + a(x_i))/2$, $f_i = (f(x_{i+1}) + f(x_i))/2$.

23.4. $u(x) = 1, \quad \varphi_i^h = \begin{cases} 1 & \text{при } i \neq 0, \\ 1 + h^{-1/4} & \text{при } i = 0. \end{cases}$

23.5. Порядок сходимости равен: $o(1)$ при $k = 0$, $O(h)$ при $k = 1$, $O(h^2)$ при $k \geq 2$. Если $u(x) = \text{const}$, то порядок сходимости — бесконечный.

23.6. Будем строить оператор L_h в виде

$$(L_h u_h)_i = \frac{a u_{i+1} + b u_i + c u_{i-1}}{h^2},$$

где h — постоянный шаг шаблона.

Для достаточно гладкой функции $u(x)$ выпишем разложения по формуле Тейлора в окрестности точки $x = x_i$: $u(x_i \pm h) = u(x_i) \pm h u'(x_i) + \frac{h^2}{2!} u''(x_i) \pm \frac{h^3}{3!} u'''(x_i) + \frac{h^4}{4!} u^{(4)}(x_i) \pm \frac{h^5}{5!} u^{(5)}(x_i) + \frac{h^6}{6!} u^{(6)}(\xi_i^\pm)$.

Подставим полученные выражения в формулу для $L_h(u)_h$ и сгруппируем множители при одинаковых производных $u(x)$ (или, что то же самое, при степенях h):

$$\begin{aligned} L_h(u)_h \Big|_{x=x_i} &= \\ &= \frac{1}{h^2} \left[(a + b + c)u(x_i) + h(a - c)u'(x_i) + \right. \\ &\quad + \frac{h^2}{2!}(a + c)u''(x_i) + \frac{h^3}{3!}(a - c)u^{(3)}(x_i) + \frac{h^4}{4!}(a + c)u^{(4)}(x_i) + \\ &\quad \left. + \frac{h^5}{5!}(a - c)u^{(5)}(x_i) + \frac{h^6}{6!}(a u^{(6)}(\xi_i^+) + c u^{(6)}(\xi_i^-)) \right]. \end{aligned}$$

По определению локальной аппроксимации $L_h(u)_h \Big|_{x=x_i} = u''(x_i) + O(h^p)$, $p > 0$, откуда выводим систему уравнений для коэффициентов a , b и c :

$$a + b + c = 0, \quad a - c = 0, \quad \frac{1}{2}(a + c) = 1.$$

Решив эту систему, получим

$$L_h(u)_h \Big|_{x=x_i} \equiv \frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1}))}{h^2} = u''(x_i) + \frac{h^2}{12} u^{(4)}(x_i) + O(h^4),$$

т. е. L_h локально аппроксимирует оператор второй производной L в точке $x = x_i$ со вторым порядком.

Разностная схема для поставленной задачи

$$\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} = f_i, \quad 0 < i < N, \quad u_0 = u_N = 0, \quad Nh = 1.$$

23.7. Обозначим $h = \max_{1 \leq i \leq N} h_i$, тогда разностная схема

$$-\frac{2}{h_{i+1} + h_i} \left(\frac{u_{i+1} - u_i}{h_{i+1}} - \frac{u_i - u_{i-1}}{h_i} \right) = f_i,$$

$$0 < i < N, u_0 = a, u_N = b$$

на неравномерной сетке

$$0 = x_0 < x_1 < \dots < x_{N-1} < x_N = 1, h_i = x_i - x_{i-1}, 1 \leq i \leq N,$$

имеет на решении порядок аппроксимации $O(h)$ при $f_i = f(x_i)$ и порядок $O(h^2)$ при

$$f_i = f(x_i) + \frac{1}{3} \frac{h_{i+1} - h_i}{h_{i+1} + h_i} (f(x_{i+1}) - f(x_{i-1})).$$

23.8. Для приближенного вычисления интеграла по отрезку $[x_{i-1}, x_{i+1}]$ использовать формулу Симпсона, а для получения недостающего начального условия $u_1 \approx u(h)$ применить формулу Тейлора (необходимые производные при $x = 0$ можно получить, дифференцируя уравнение требуемое число раз).

23.9. Запишем интегральное тождество Марчука:

$$-\frac{u(x_{i+1}) - u(x_i)}{\int_{x_i}^{x_{i+1}} \frac{dx}{k(x)}} + \frac{u(x_i) - u(x_{i-1})}{\int_{x_{i-1}}^{x_i} \frac{dx}{k(x)}} - \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x) dx =$$

$$= \frac{1}{\int_{x_i}^{x_{i+1}} \frac{dx}{k(x)}} \int_{x_i}^{x_{i+1}} \frac{dx}{k(x)} \int_{x_{i+\frac{1}{2}}}^x f(\xi) d\xi - \frac{1}{\int_{x_{i-1}}^{x_i} \frac{dx}{k(x)}} \int_{x_{i-1}}^{x_i} \frac{dx}{k(x)} \int_{x_{i-\frac{1}{2}}}^x f(\xi) d\xi.$$

Предположим для удобства, что точка $x = 1/2$ — узел сетки при любом h , т. е. $h = 1/N$, $N = 2K$. При этом $i = N/2$ — соответствующее значение индекса i . Вычислим величины

$$t_i = \int_{x_i}^{x_{i+1}} \frac{dx}{k(x)} = \begin{cases} h & \text{при } 0 \leq i < N/2, \\ h/2 & \text{при } N/2 \leq i < N. \end{cases}$$

Заменим интеграл в левой части тождества по формуле прямоугольников

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x) dx \approx h f(x_i) \equiv h f_i.$$

Теперь рассмотрим выражения в правой части тождества. Одно из них, например,

$$\int_{x_i}^{x_{i+1}} \frac{dx}{k(x)} \int_{x_{i+\frac{1}{2}}}^x f(\xi) d\xi,$$

применяя квадратурную формулу прямоугольников, запишем так:

$$\int_{x_i}^{x_{i+1}} \frac{dx}{k(x)} \int_{x_{i+\frac{1}{2}}}^x f(\xi) d\xi = \frac{h}{k(x_{i+\frac{1}{2}})} \int_{x_{i+\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(\xi) d\xi + O(h^3) = O(h^3).$$

Множитель при рассматриваемом интеграле в тождестве равен $O(h^{-1})$, поэтому все выражение для гладких функций имеет порядок $O(h^2)$ и его можно отбросить. Аналогично можно поступить и с другим выражением из правой части.

Окончательный результат можно записать в виде

$$-\frac{a_i u_{i+1} - b_i u_i + c_i u_{i-1}}{h^2} = f_i, \quad 0 < i < N, u_0 = u_N = 0,$$

где коэффициенты определяются по формулам: $b_i = a_i + c_i$,

$$\begin{cases} a_i = c_i = 1 & \text{при } 1 \leq i < N/2, \\ a_i = c_i = 2 & \text{при } N/2 < i < N, \\ a_i = 2, c_i = 1 & \text{при } i = N/2. \end{cases}$$

23.10. При $c > -\pi^2$.

23.11. Следует вычислить величины

$$a_{ij} = (L\varphi_j^h, \varphi_i^h) = \int_0^1 [k(x)(\varphi_j^h)'(\varphi_i^h)' + p(x)\varphi_j^h\varphi_i^h] dx, \quad 1 \leq i, j \leq N-1,$$

$$b_i = (f, \varphi_i^h) = \int_0^1 f \varphi_i^h dx = \int_{x_{i-1}}^{x_{i+1}} f \varphi_i^h dx, \quad 1 \leq i \leq N-1$$

с кусочно-линейными базисными функциями $\varphi_j^h(x)$ на равномерной сетке $\overline{D}_h = \{x_j = jh, j = 0, \dots, N; Nh = 1\}$.

Так как

$$(\varphi_j^h)' = \begin{cases} \frac{1}{h} & \text{при } x_{j-1} < x < x_j, \\ -\frac{1}{h} & \text{при } x_j < x < x_{j+1}, \\ 0 & \text{при } x \notin [x_{j-1}, x_{j+1}], \end{cases}$$

то непосредственные вычисления дают

$$a_{ij} = \begin{cases} -\frac{1}{h} \left(1 + \frac{x_{i-1} + x_i}{2} \right) + \frac{h}{6} & \text{при } j = i - 1, \\ \frac{2}{h} (1 + x_i) + \frac{2h}{3} & \text{при } j = i, \\ -\frac{1}{h} \left(1 + \frac{x_{i+1} + x_i}{2} \right) + \frac{h}{6} & \text{при } j = i + 1, \\ 0 & \text{при остальных } j. \end{cases}$$

Компоненты вектора правой части

$$b_i = \int_{x_{i-1}}^{x_{i+1}} f \varphi_i^h dx \approx f(x_i) \int_{x_{i-1}}^{x_{i+1}} \varphi_i^h dx = h f_i.$$

Разделив обе части уравнения на h , окончательно получим

$$-\frac{d_i \alpha_{i+1} - e_i \alpha_i + c_i \alpha_{i-1}}{h^2} + \frac{\alpha_{i+1} + 4\alpha_i + \alpha_{i-1}}{6} = f_i, \quad 0 < i < N, \quad \alpha_0 = \alpha_N = 0,$$

где коэффициенты определяются формулами

$$c_i = 1 + \frac{x_{i-1} + x_i}{2}, \quad d_i = 1 + \frac{x_{i+1} + x_i}{2}, \quad e_i = d_i + c_i.$$

23.13. Запишем функционал, соответствующий исходной задаче,

$$J(v) = \int_0^1 \left(k(x) (v')^2 + p(x) v^2 - 2f(x)v \right) dx,$$

и будем искать приближенное решение в виде $u_h = \sum_{j=1}^N \alpha_j \varphi_j(x)$, где базисные функции взяты из решения **23.11**. Подставим u_h в J и рассмотрим систему

$$\frac{\partial J(u_h)}{\partial \alpha_i} = 0, \quad 1 \leq i \leq N.$$

Нас интересует последнее уравнение системы (при $i = N$)

$$d \alpha_{N-1} + b \alpha_N = c,$$

где $d = \int_0^1 (k(x) \varphi'_{N-1} \varphi'_N + p(x) \varphi_{N-1} \varphi_N) dx$, $b = \int_0^1 \left(k(x) (\varphi'_N)^2 + p(x) \varphi_N^2 \right) dx$, $c = \int_0^1 f(x) \varphi_N dx$. Запишем формулу для $\varphi_N(x)$:

$$\varphi_N(x) = \begin{cases} 0 & \text{при } 0 \leq x \leq x_{N-1} = 1 - h, \\ \frac{x - x_{N-1}}{h} & \text{при } x_{N-1} \leq x \leq x_N = 1. \end{cases}$$

Эта базисная функция отлична от нуля только на отрезке $[1-h, 1]$, поэтому нам потребуется только часть функции $\varphi_{N-1}(x)$:

$$\varphi_{N-1}(x) = \frac{1-x}{h} \quad \text{при} \quad 1-h \leq x \leq 1.$$

В случае постоянных коэффициентов $k(x) \equiv k$, $p(x) \equiv p$ величины d , b , c определяются так:

$$d = -\frac{k}{h} + \frac{ph}{6}, \quad b = \frac{k}{h} + \frac{ph}{3}, \quad c = \int_{1-h}^1 f(x) \frac{x-1+h}{h} dx.$$

23.14. Рассмотреть коэффициент при 2ε в неравенстве $J(u) \leq J(u + \varepsilon w)$, справедливым при ε любого знака и любом $w \in U$.

24.1. Пусть $y(x)$ — произвольная гладкая функция. Тогда условия аппроксимации для левой и правой частей уравнения $y' = f(x, y)$ означают справедливость соотношений

$$\lim_{h \rightarrow 0} \frac{1}{h} \sum_{i=0}^n a_{-i} y_{k-i} = y'(x_k), \quad \lim_{h \rightarrow 0} \sum_{i=0}^n b_{-i} f(x_{k-i}, y_{k-i}) = f(x_k, y_k)$$

в произвольном узле x_k , $k \geq n$. Согласно формуле Тейлора

$$y(x-ih) = y(x) - ih y'(x) + O(h^2), \quad f(x-ih, y(x-ih)) = f(x, y(x)) + O(h).$$

Подставляя эти выражения в условия аппроксимации, получим

$$\lim_{h \rightarrow 0} \left[\left(\frac{1}{h} \sum_{i=0}^n a_{-i} \right) y(x_k) - \left(\sum_{i=0}^n i a_{-i} \right) y'(x_k) + O(h) \right] = y'(x_k),$$

$$\lim_{h \rightarrow 0} \left[\left(\sum_{i=0}^n b_{-i} \right) f(x_k, y(x_k)) + O(h) \right] = f(x_k, y(x_k)),$$

откуда (в силу произвольности функции $y(x)$) следует необходимость и достаточность указанных в условии задачи равенств.

24.2. 1) Да; 2) нет; 3) да.

24.3. Первый; можно, если положить $y_1 = h$, то порядок аппроксимации равен двум. В отличие от дифференциального случая, для разностной задачи необходимы два начальных условия. Поэтому аппроксимация решения в точке $x = h$ является частью формальной аппроксимации дифференциального оператора L .

24.4. Схема имеет второй порядок аппроксимации на решении. Исследуем ее сходимость. Несложно показать, что точные решения

дифференциальной и разностной задач имеют вид

$$y(x) = e^{-5x} + 1,$$

$$y_k = 1 + C_1 \mu_1^k + C_2 \mu_2^k, \quad \mu_{1,2} = -5h \pm \sqrt{1 + 25h^2}, \quad |\mu_1| < 1, \quad |\mu_2| > 1.$$

Из начальных условий $y_0 = 2, y_1 = 2 - 5h$ имеем систему уравнений для определения коэффициентов C_1, C_2

$$1 + C_1 + C_2 = 2, \quad 1 + C_1 \mu_1 + C_2 \mu_2 = 2 - 5h,$$

откуда следует, что $C_1 \neq 0$ и $C_2 \neq 0$. Это означает, что решение разностной задачи содержит растущую компоненту, и разностная схема на больших промежутках времени неверно отражает решение дифференциальной задачи, хотя формально схема α -устойчива и разностное решение сходится на любом конечном интервале к решению дифференциальной задачи.

24.5. Все коэффициенты равны $1/2$.

24.6. Использовать метод неопределенных коэффициентов построения разностных схем, заменив f на y' и сдвинув для удобства вычислений индексы заменой $j = k - 1$. Получить: $a_1 = 1/6, a_0 = 2/3, a_{-1} = 1/6; p = 4$.

24.7. Схема устойчива при $\theta = 0$ и $1 \geq \theta \geq 1/2$.

24.8. Учитывая необходимые условия аппроксимации (см. **24.1**), запишем систему для определения коэффициентов

$$2a + 4 = 1, \quad 2b + c = 1, \quad 8 + a = 3b,$$

или $a = -3/2, b = 13/6, c = -10/3$. При этом характеристическое уравнение имеет вид

$$(\mu^2 - 1) \left(\mu^2 - \frac{3}{2} \mu + 1 \right) = 0,$$

т. е. условие α -устойчивости выполнено.

Без учета нормировки $\lim_{h \rightarrow 0} \|f_h\|_{F_h} = \|f\|_F$ можно прийти к неверному ответу: $a = 28, b = 12, c = 36$, для которого условие α -устойчивости не выполнено.

24.9. Схема имеет первый порядок сходимости.

24.10. Решение разностной задачи имеет вид

$$y_k = (1 + h)y_{k-1} = (1 + h)^k y_0 = (1 + h)^k,$$

а точное решение дифференциальной задачи при $x = x_k$ равно $y(x_k) = \exp(x_k)$. Пусть $x_k = kh = 1$, тогда

$$\begin{aligned} y(x_k) - y_k &= e - (1+h)^{1/h} = e - \exp\left[\frac{1}{h} \ln(1+h)\right] = \\ &= e \left(1 - \exp\left[-\frac{h}{2} + O(h^2)\right]\right) = \frac{e}{2} h + O(h^2). \end{aligned}$$

Окончательно имеем: $c_1 = e/2$.

24.11. $c_1 = 0$.

24.12. Вывести формулу

$$y_k = y_0 \left[\frac{\mu_2}{\mu_2 - \mu_1} \mu_1^k - \frac{\mu_1}{\mu_2 - \mu_1} \mu_2^k \right] + y_1 \left[-\frac{1}{\mu_2 - \mu_1} \mu_1^k + \frac{1}{\mu_2 - \mu_1} \mu_2^k \right],$$

где $\mu_{1,2}$ — корни уравнения $\mu^2 + 2h\mu - 1 = 0$:

$$\mu_1 = -h + \sqrt{1+h^2} = 1 - h + \frac{h^2}{2} + O(h^4), \quad \mu_2 = -\left(1 + h + \frac{h^2}{2}\right) + O(h^4).$$

Получить $c_1 = 0$.

24.13. Эта схема неустойчива, сходимости нет.

24.14. Например,

$$\frac{y_{k+1} - y_{k-1}}{2h} + y_k = \cos(2hk), \quad y_0 = 0, \quad y_1 = h, \quad k \geq 1.$$

24.15. Например,

$$\frac{y_{k+1} - y_k}{h} + 5 \frac{y_{k+1} + y_k}{2} = \frac{\sin(2h(k+1)) + \sin(2hk)}{2}, \quad y_0 = 2, \quad k \geq 0.$$

24.16. Например,

$$\frac{y_{k+1} - y_{k-1}}{2h} - y_k = \exp(2hk), \quad y_0 = 1, \quad y_1 = 1 + 2h, \quad k \geq 1.$$

24.17. Например,

$$\frac{y_{k+1} - y_k}{h} - (y_{k+1} + y_k) = \frac{\exp(h(k+1)) + \exp(hk)}{2}, \quad y_0 = 1, \quad k \geq 0.$$

24.18. Например, можно взять заведомо α -неустойчивую схему

$$4 \frac{y_{k+1} - y_{k-1}}{2h} - 3 \frac{y_{k+1} - y_k}{h} = c f_{k-1} + d f_k + e f_{k+1},$$

и методом неопределенных коэффициентов получить заданный порядок аппроксимации на решении.

- 24.19.** 1) $-h^4 y^{(5)}(\xi)/180$, схема α -устойчива,
 2) $h^3 y^{(4)}(\xi)/24$, схема α -устойчива,
 3) $7h^4 y^{(5)}(\xi)/90$, схема α -устойчива,
 4) $h^3 y^{(4)}(\xi)/36$, схема α -неустойчива.

25.1. $p = 2$.

25.2. $\alpha = \gamma = 1/12$, $\beta = 5/6$.

25.3. Из формулы Тейлора $u(h) = u(0) + h u'(0) + \frac{h^2}{2} u''(0) + O(h^3)$ следует

$$u'(0) = \frac{u(h) - u(0)}{h} - \frac{h}{2} u''(0) + O(h^2).$$

Из исходного уравнения имеем $-u''(0) = f(0) - p(0)u(0)$, поэтому

$$a u(0) + b \left(\frac{u(h) - u(0)}{h} + \frac{h}{2} (f(0) - p(0)u(0)) \right) = c + O(h^2).$$

Искомая аппроксимация имеет вид

$$\left(a - \frac{h}{2} b p_0 \right) u_0 + b \frac{u_1 - u_0}{h} = c - \frac{h}{2} b f_0.$$

25.4. $-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + p(x_i)u_i = f(x_i)$, $1 \leq i \leq N-1$, $\frac{u_1 - u_0}{h} = 1$, $u_N = 0$. Отметим, что u_0 аппроксимирует $u(-h/2)$, u_N аппроксимирует $u(1)$.

25.5. Так как (см. **22.12**) собственные значения разностной задачи

$$-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} = \lambda u_i, \quad 0 < i < N, \quad u_0 = u_N = 0, \quad h = 1/N,$$

имеют вид

$$\lambda^{(m)}(A) = \frac{4}{h^2} \sin^2 \frac{\pi m h}{2}, \quad m = 1, \dots, N-1,$$

то справедливы оценки

$$\lambda_{\min} = \lambda^{(1)}(A) = \frac{4}{h^2} \sin^2 \frac{\pi h}{2} \geq 4, \quad \lambda_{\max} = \lambda^{(N-1)}(A) \leq \frac{4}{h^2}.$$

Выше было использовано неравенство $\sin |\beta| \geq 2|\beta|/\pi$ при $|\beta| \leq \pi/2$. Отсюда следует порядок обусловленности системы.

Исходная задача записывается в виде $A \mathbf{u} = \mathbf{f}$, или, в силу невырожденности матрицы A , $\mathbf{u} = A^{-1} \mathbf{f}$. Отсюда получаем неравенство для евклидовой нормы векторов $\|\mathbf{u}\|_2 \leq \|A^{-1}\|_2 \|\mathbf{f}\|_2$. Подчиненная

матричная норма (см. **13.5**) имеет вид $\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$. В рассматриваемом случае $\|A^{-1}\|_2 = 1/\lambda_{\min}$, и выполнение неравенства $\lambda_{\min} \geq C$ с постоянной C , не зависящей от h , по определению, означает устойчивость схемы в евклидовой норме.

25.6. Введем обозначение оператора

$$\Lambda u_i = \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}$$

и покажем, что если $\Lambda u_i \leq 0$ при $i = 1, \dots, N-1$ и $u_0 = u_N = 0$, то $u_i \geq 0$ при всех i .

Пусть $d = \min_i u_i < 0$ и q — такое наименьшее целое, что $u_q = d$. Тогда $u_{q-1} > d$, $u_{q+1} \geq d$ и

$$\Lambda u_q = \frac{(u_{q+1} - d) + (u_{q-1} - d)}{h^2} > 0.$$

Полученное противоречие доказывает, что $u_i \geq 0$ при всех i .

Следующий шаг — доказательство неравенства

$$\max_{0 \leq i \leq N} |u_i| \leq \frac{1}{8} U, \quad \text{где } U = \max_{0 < i < N} |\Lambda u_i|.$$

Введем для этого функцию

$$w_i = \frac{ih(1-ih)}{2} U, \quad i = 0, \dots, N,$$

удовлетворяющую условиям $w_i \geq 0$ и $\Lambda w_i = -U$. Теперь для функций $w_i \pm u_i$ справедливо

$$\Lambda(w_i \pm u_i) = -U \pm \Lambda u_i \leq 0, \quad w_0 \pm u_0 = w_N \pm u_N = 0.$$

Поэтому, используя доказанное выше свойство, имеем $w_i \pm u_i \geq 0$, откуда и следует требуемое неравенство

$$|u_i| \leq w_i \leq \max_{0 < i < N} w_i \leq \frac{1}{8} U.$$

Последний этап — определение величины U . Используя формулу Тейлора, запишем уравнение для погрешности $u(x_i) - u_i$

$$\Lambda(u(x_i) - u_i) = \frac{h^2}{12} u^{(4)}(\xi_i) = -\frac{h^2}{12} f''(\xi_i).$$

Отсюда $U = \frac{h^2}{12} \max_{[0,1]} |f''(x)|$, что и приводит к искомой оценке.

25.7. $\frac{u_N - u_{N-1}}{h} + \frac{h}{2} (\cos(1) + 1) - 3 u_N = 1.$

25.8. $\frac{u_1 - u_0}{h} - \frac{h}{2} + 4u_0 = 1.$

25.9. $\frac{u_N - u_{N-1}}{h} + \frac{h}{2} (3u_N + \exp(1)) = 0.$

25.10. $\frac{u_1 - u_0}{h} - \frac{h}{2} (2u_0 - 1) - u_0 = 0.$

25.11. Введем обозначение $A\mathbf{u} = \mathbf{f}$. Все схемы устойчивы, так как (см. решение 25.5)

1) $\lambda^{(m)}(A) = \frac{4}{h^2} \sin^2 \frac{\pi(m-1)}{2(N-1)} + 1, m = 1, \dots, N-1, \lambda_{\min} = 1;$

2) $\lambda^{(m)}(A) = \frac{4}{h^2} \sin^2 \frac{\pi(2m-1)}{2(2N-1)}, m = 1, \dots, N-1, \lambda_{\min} \geq 1;$

3) $\lambda^{(m)}(A) = \frac{4}{h^2} \sin^2 \frac{\pi(2m-1)}{2(2N-1)}, m = 1, \dots, N-1, \lambda_{\min} \geq 1;$

4) для разностного аналога второй производной

$$\Lambda u_i = \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}, \quad 1 \leq i \leq N-1$$

при условиях $u_0 = u_N = 0, Nh = 1$ имеем следующие собственные значения:

$$\lambda^{(m)}(-\Lambda) = \frac{4}{h^2} \sin^2 \frac{\pi m}{2N}, \quad m = 1, \dots, N-1, \quad \lambda_{\min}(-\Lambda) \geq 4,$$

поэтому для исходной задачи $\lambda_{\min}(A) \geq 4 - 3 = 1.$

25.12. При их проверке достаточно учесть, что $\Delta\varphi_{i-1} = \nabla\varphi_i.$

25.14. Воспользоваться формулой суммирования по частям.

25.15. $(\psi, \Lambda\varphi) = -(a\nabla\varphi, \nabla\psi) - (d\varphi, \psi) + (a\psi\nabla\varphi)_N - \psi_0(a\nabla\varphi)_1,$

$$(\varphi, \Lambda\psi) - (\psi, \Lambda\varphi) = a_N(\psi\nabla\varphi - \varphi\nabla\psi)_N - a_1(\psi_0\nabla\varphi_1 - \varphi_0\nabla\psi_1).$$

25.16. Умножим i -е уравнение на u_i и просуммируем их от 1 до $(N-1)$. Учитывая, что $u_0 = u_N = 0$, получим

$$\begin{aligned} & -\frac{1}{h^2} \sum_{i=1}^{N-1} (u_{i+1} - u_i)u_i + \frac{1}{h^2} \sum_{i=1}^{N-1} (u_i - u_{i-1})u_i = \\ & = -\frac{1}{h^2} \sum_{i=1}^N (u_i - u_{i-1})u_{i-1} + \frac{1}{h^2} \sum_{i=1}^N (u_i - u_{i-1})u_i = \frac{1}{h^2} \sum_{i=1}^N (u_i - u_{i-1})^2. \end{aligned}$$

Отсюда следует сумматорное (аналог интегрального) тождество

$$\frac{1}{h^2} \sum_{i=1}^N (u_i - u_{i-1})^2 + \sum_{i=1}^{N-1} p_i u_i^2 = \sum_{i=1}^{N-1} f_i u_i,$$

которое, учитывая обозначения $\nabla u_i = u_i - u_{i-1}$, $(u, v) = \sum_{i=1}^N u_i v_i$,

$(u, v) = \sum_{i=1}^{N-1} u_i v_i$, можно записать в виде

$$\frac{1}{h^2} (\nabla u, \nabla u] + (p u, u) = (f, u).$$

Докажем сеточный аналог неравенства для функции и ее производной. Представим значение u_k в виде суммы $u_k = \sum_{i=1}^k (u_i - u_{i-1})$.

Тогда

$$u_k^2 \leq \sum_{i=1}^k 1^2 \sum_{i=1}^k (u_i - u_{i-1})^2 \leq N \sum_{i=1}^N (u_i - u_{i-1})^2 \quad \forall k = 1, 2, \dots, N-1.$$

Суммируя по переменной k , получим требуемое неравенство

$$\sum_{k=1}^{N-1} u_k^2 \leq N(N-1) \sum_{i=1}^N (u_i - u_{i-1})^2 \leq \frac{1}{h^2} \sum_{i=1}^N (u_i - u_{i-1})^2.$$

Используя его в левой части сумматорного тождества, будем иметь

$$\sum_{k=1}^{N-1} u_k^2 \leq \frac{1}{2} \left(\sum_{i=1}^{N-1} f_i^2 + \sum_{i=1}^{N-1} u_i^2 \right).$$

Таким образом, $\sum_{i=1}^{N-1} u_i^2 \leq \sum_{i=1}^{N-1} f_i^2$, и априорная оценка решения разностной задачи в норме пространства $L_{2,h}$: $\|u_h\|_h^2 = h(u_h, u_h)$, согласованной с непрерывной нормой L_2 , имеет вид $\|u_h\|_h \leq \|f_h\|_h$.

25.17. Запишем на сетке $x_i = ih$, $0 \leq i \leq N$ для функции y_i тождество

$$y_i^2 \equiv (1 - x_i)y_i^2 + x_i y_i^2.$$

В силу условий $y_0 = y_N = 0$ справедливы равенства

$$y_i^2 = \left(\sum_{k=1}^i (y_{\bar{x}})_k h \right)^2, \quad y_i^2 = \left(\sum_{k=i+1}^N (y_{\bar{x}})_k h \right)^2.$$

Подставив эти равенства в тождество, оценим его правую часть с использованием неравенства Коши — Буняковского

$$\begin{aligned} y_i^2 &\leq (1 - x_i) \sum_{k=1}^i h \sum_{k=1}^i (y_{\bar{x}})_k^2 h + x_i \sum_{k=i+1}^N h \sum_{k=i+1}^N (y_{\bar{x}})_k^2 h = \\ &= x_i(1 - x_i) \sum_{k=1}^N (y_{\bar{x}})_k^2 h \equiv x_i(1 - x_i) \|y_{\bar{x}}\|^2. \end{aligned}$$

Максимум выражения $x(1-x)$ на отрезке $[0, 1]$ равен $1/4$ и достигается при $x = 1/2$, поэтому $y_i^2 \leq \|y_{\bar{x}}\|^2/4$ и $\|y\|_C \leq \|y_{\bar{x}}\|/2$.

25.18. Сделать замену $x' = lx$ и использовать решение **25.17**.

25.20. Так как $\|y_{\bar{x}}\|^2 = (-\Lambda y, y)$, где

$$(\Lambda y)_i = \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2}, \quad 1 \leq i \leq N-1, \quad y_0 = y_N = 0,$$

то постоянные в сеточных неравенствах можно получить, определив экстремумы собственных значений оператора Λ . Из решения задачи на собственные значения $\Lambda y = -\lambda y$ следует

$$\lambda_{\min} = \frac{4}{h^2} \sin^2 \frac{\pi h}{2l}, \quad \lambda_{\max} = \frac{4}{h^2} \cos^2 \frac{\pi h}{2l}.$$

Постоянные в искомом неравенстве получаются из оценок снизу для λ_{\min} и сверху для λ_{\max} .

26.1. $O(\tau + h)$. При $\tau/h = a$.

26.2. $O\left(\tau + h^2 + \frac{h^2}{\tau}\right) = O\left(\tau + \frac{h^2}{\tau}\right)$.

26.3. При $a < 0$ и $r = 1/|a|$ существует разностная схема с порядком аппроксимации $O(h^2)$.

26.4. Имеется однопараметрическое семейство схем первого порядка, коэффициенты которого удовлетворяют системе уравнений

$$a^0 r h = 1, \quad a^0 + a_0 + a_1 + a_{-1} = 0, \quad a^0 r + a_1 - a_{-1} = 0.$$

Схемы второго порядка не существуют.

26.5. $u_m^1 = \varphi(mh) + \tau[\varphi_x(mh) + f(mh, 0)]$.

26.6. Исключая $u_m^{n+1/2}$ при $m \pm 1/2$, получим схему

$$\frac{u_m^{n+1} - u_m^n}{\tau} + a \frac{u_{m+1}^n - u_{m-1}^n}{2h} - a^2 \frac{\tau}{2} \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2} = 0$$

второго порядка аппроксимации на решении при $\tau = rh$, $r = \text{const}$.

26.7. При $\theta = 1/2 - h/(2\tau)$.

26.8. Взять разностную схему

$$\frac{u_m^{n+1} - u_m^n}{\tau} + \frac{1}{2} \left[\frac{u_{m+1}^{n+1} - u_{m-1}^{n+1}}{2h} + \frac{u_{m+1}^n - u_{m-1}^n}{2h} \right] = 0,$$

аппроксимирующую уравнение в точке $(x_m, t_{n+\frac{1}{2}})$ с порядком $O(\tau^2 + h^2)$. В разложении в ряд Тейлора в окрестности этой

точки исключить главный член погрешности по h , аппроксимируя с четвертым порядком производную по переменной x на заданном шаблоне.

26.9. Рассмотреть шаблоны из l узлов: 1) $l = 4$; 2) $l = 4$; 3) $l = 4$; 4) $l = 3$.

26.10. Подставим в разностную схему частное решение $u_m^n = \lambda^n e^{im\varphi}$:

$$\frac{\lambda^{n+1} e^{im\varphi} - \lambda^n e^{im\varphi}}{\tau} + a \frac{\lambda^n e^{im\varphi} - \lambda^n e^{i(m-1)\varphi}}{h} = 0.$$

Сокращая на $\lambda^n e^{im\varphi}$, получим

$$\frac{\lambda - 1}{\tau} + a \frac{1 - e^{-i\varphi}}{h} = 0,$$

откуда следует

$$\lambda(\varphi) = 1 - \frac{a\tau}{h} + \frac{a\tau}{h} e^{-i\varphi}.$$

Пусть $a > 0$. Тогда при $0 < \frac{a\tau}{h} \leq 1$ имеем $|\lambda(\varphi)| \leq 1 - \frac{a\tau}{h} + \frac{a\tau}{h} = 1$, т. е. схема устойчива. При $\frac{a\tau}{h} = 1 + \gamma > 1$, $\gamma = \text{const}$, получим $\lambda(\pi) = -1 - 2\gamma < -1$, т. е. в этом случае схема неустойчива. Таким образом, при $a > 0$ разностная схема условно устойчива.

При $a < 0$ схема неустойчива.

26.11. Аналогично **26.10** получим

$$\lambda(\varphi) = 1 - \frac{a\tau}{2h} (e^{i\varphi} - e^{-i\varphi}) = 1 - i \frac{a\tau}{h} \sin \varphi,$$

откуда следует, что $\max |\lambda(\varphi)| = |\lambda(\pi/2)| = \sqrt{1 + a^2 \tau^2 / h^2}$. Пусть $\tau = Ah^2$, тогда $|\lambda(\pi/2)| \leq \exp \left\{ \frac{a^2 A}{2} \tau \right\}$, т. е. при $\tau = O(h^2)$ схема устойчива.

26.12. $\lambda(\varphi) = \frac{1}{2} \left(1 - \frac{a\tau}{h}\right) e^{i\varphi} + \frac{1}{2} \left(1 + \frac{a\tau}{h}\right) e^{-i\varphi} = \cos \varphi - \frac{a\tau}{h} i \sin \varphi$.

Схема устойчива ($|\lambda(\varphi)| \leq 1$) при выполнении условия $|a|\tau/h \leq 1$.

26.13. $\lambda(\varphi) = 1 - i(a\tau/h) \sin \varphi + (a^2 \tau / h^2)(\cos \varphi - 1)$. Схема устойчива при выполнении условия $|a|\tau/h \leq 1$.

26.14. $\lambda(\varphi) = \left(1 + \frac{a\tau}{h} - \frac{a\tau}{h} e^{-i\varphi}\right)^{-1}$. Введем следующие обозначения:

$\delta = \sup_{0 \leq \varphi \leq 2\pi} |\lambda(\varphi)|$ и $\gamma = a\tau/h$. Тогда:

при $a > 0$ или при $\gamma \leq -1$ выполняется неравенство $\delta \leq 1$, т. е. схема устойчива;

при $-1 < \gamma < 0$ имеем $\delta = \frac{1}{|2\gamma + 1|} > 1$, т. е. схема неустойчива.

26.15. $\lambda(\varphi) = \left(1 + \frac{a\tau}{h} i \sin \varphi\right)^{-1}$. Так как $|\lambda(\varphi)| \leq 1$, то схема устойчива при любых τ и h .

26.16. $\lambda(\varphi) = \cos \varphi - \frac{a\tau}{h} i \sin \varphi$. Схема устойчива ($|\lambda(\varphi)| \leq 1$) при выполнении условия $|a|\tau/h \leq 1$ (ср. с **26.12**).

26.17. Схема безусловно устойчива.

26.18. При $0 \leq \theta \leq 1/2$.

26.19. При $1/2 \leq \theta \leq 1$.

27.1. При $\tau/h^2 = 1/6$.

27.2. При $\theta = 1/2 - h^2/(12\tau)$.

27.3. $O(\tau^2 + h^4)$.

27.5. Схема имеет порядок аппроксимации $O(\tau + h^2)$. Введем обозначение $\rho = \tau/h^2$ и перепишем схему в удобном для анализа виде

$$u_m^{n+1} = (1 - 2\rho) u_m^n + \rho (u_{m+1}^n + u_{m-1}^n) + \tau f_m^n.$$

При $\rho \leq 1/2$ имеем

$$\begin{aligned} \|u^{n+1}\|_C &\leq (1 - 2\rho)\|u^n\|_C + 2\rho\|u^n\|_C + \tau\|f^n\|_C = \|u^n\|_C + \tau\|f^n\|_C \leq \\ &\leq \|u^{n-1}\|_C + \tau(\|f^n\|_C + \|f^{n-1}\|_C) \leq \dots \end{aligned}$$

$$\dots \leq \|u^0\|_C + \sum_{k=0}^n \tau\|f^k\|_C \leq \|u^0\|_C + (n+1)\tau \max_n \|f^n\|_C.$$

Следовательно, схема удовлетворяет определению устойчивости с постоянной $K = T$ при условии $\tau/h^2 \leq 1/2$.

27.6. Порядок аппроксимации схемы $O(\tau + h^2)$. Удобная для анализа форма записи имеет вид

$$u_m^{n+1} + \rho(-u_{m-1}^{n+1} + 2u_m^{n+1} - u_{m+1}^{n+1}) = u_m^n + \tau f_m^{n+1}, \quad \rho = \tau/h^2.$$

Выбирая из всех значений u_m^{n+1} , по модулю равных $\|u^{n+1}\|_C$, такое, у которого индекс m принимает наименьшее значение, имеем

$$|u_m^{n+1}| > |u_{m-1}^{n+1}| \quad \text{и} \quad |u_m^{n+1}| \geq |u_{m+1}^{n+1}|.$$

Отсюда $|2u_m^{n+1}| > |u_{m-1}^{n+1}| + |u_{m+1}^{n+1}|$, и знак выражения $2u_m^{n+1} - u_{m-1}^{n+1} - u_{m+1}^{n+1}$ совпадает со знаком u_m^{n+1} , т. е. справедлива оценка снизу

$$\|u^{n+1}\|_C = |u_m^{n+1}| \leq |u_m^{n+1} + \rho(2u_m^{n+1} - u_{m-1}^{n+1} - u_{m+1}^{n+1})| = |u_m^n + \tau f_m^{n+1}|.$$

Таким образом, при любых шагах сетки τ и h справедливо неравенство $\|u^{n+1}\|_C \leq \|u^n\|_C + \tau \|f^{n+1}\|_C$. Дальнейший вывод оценки безусловной устойчивости разностной схемы аналогичен решению **27.5**.

- 27.7.** Сходимость имеет место только при выполнении условия устойчивости $\rho \leq 1/2$, при этом порядок сходимости $O(\tau + h^2)$ для $\rho \neq 1/6$ и $O(\tau^2 + h^4)$ для $\rho = 1/6$.
- 27.8.** При выполнении указанного условия среди частных решений вида $u_m^n = \mu_h^n(k) \sin(\pi m h k)$ найдется номер k такой, что $|\mu_h(k)|^{\frac{T}{\tau}} \rightarrow \infty$ при $\tau \rightarrow 0$.
- 27.9.** С помощью частных решений вида $u_m^n = \mu_h^n(k) \sin(\pi m h k)$ показать, что схема неустойчива для всех τ и h .
- 27.10.** Введем оператор

$$\Lambda u_m = \frac{u_{m+1} - 2u_m + u_{m-1}}{h^2}.$$

Под Λ будем также понимать матрицу, которая ставит в соответствие вектору $u^n = (u_1^n, \dots, u_{M-1}^n)^T$ вектор $\Lambda u^n = (\Lambda u_1^n, \dots, \Lambda u_{M-1}^n)^T$ с учетом условий $u_0^n = u_M^n = 0$. Тогда рассматриваемую схему можно записать в виде $u^{n+1} = S u^n = S^{n+1} u^0$, где $S = (I - \tau \theta \Lambda)^{-1} (I + \tau(1 - \theta)\Lambda)$, а I — единичная матрица.

Задача на собственные значения

$$\Lambda y_m = -\lambda y_m, \quad 1 \leq m \leq M - 1, \quad Mh = l, \quad y_0 = y_M = 0,$$

имеет решение (см. **22.12**)

$$y_m^{(k)} = \sqrt{\frac{2}{l}} \sin\left(\frac{\pi m k}{M}\right), \quad \lambda^{(k)} = \frac{4}{h^2} \sin^2 \frac{\pi k h}{2l}, \quad 1 \leq k \leq M - 1.$$

Матрицы $(I + \beta \Lambda)^{\pm 1}$, а вместе с ними и S , имеют ту же систему собственных векторов, что и Λ , поэтому можно выразить собственные значения $\mu^{(k)}$ матрицы S через $\lambda^{(k)}$:

$$\mu^{(k)}(S) = \frac{1 - \tau(1 - \theta)\lambda^{(k)}}{1 + \tau\theta\lambda^{(k)}}.$$

Матрица S является симметричной, так как имеет представление $S = Q D Q^{-1}$, где столбцами ортогональной матрицы Q являются векторы $y^{(k)}$, а D — диагональная матрица, состоящая из соответствующих $\mu^{(k)}$. Матричная норма, подчиненная векторной евклидовой норме, в случае симметричной матрицы равна ее максимальному собственному значению. Метрика $L_{2,h}$ отличается от

евклидовой только множителем h , поэтому $\|S\|_{L_{2,h}} = \max_k |\mu^{(k)}(S)|$.

Выясним теперь, при каких условиях на параметры τ , θ и h будет выполнено $\|S\|_{L_{2,h}} \leq 1$, т. е.

$$-1 \leq \frac{1 - \tau(1 - \theta)\lambda^{(k)}}{1 + \tau\theta\lambda^{(k)}} \leq 1$$

для $k = 1, 2, \dots, M - 1$. Так как $\tau, \lambda^{(k)} > 0$, $\theta \geq 0$, то знаменатель дроби всегда положителен и неравенства преобразуются к виду

$$-(1 + \tau\theta\lambda^{(k)}) \leq 1 - \tau(1 - \theta)\lambda^{(k)} \leq 1 + \tau\theta\lambda^{(k)}.$$

Правое неравенство выполняется всегда, значит, содержательным является левое неравенство. Перепишем его в виде $\tau(1 - 2\theta)\lambda^{(k)} \leq 2$. При $1/2 \leq \theta \leq 1$ это неравенство выполняется при любом τ , а при $0 \leq \theta < 1/2$ имеет место ограничение

$$\tau \leq \frac{2}{(1 - 2\theta) \max_k \lambda^{(k)}} \leq \frac{h^2}{2 - 4\theta}.$$

Из полученной выше формулы $u^n = S^n u^0$ следует, что

$$\|u^n\|_{L_{2,h}} \leq \|S\|_{L_{2,h}}^n \|u^0\|_{L_{2,h}}.$$

Поэтому для всех $\theta \in [0, 1]$ справедливо неравенство устойчивости в метрике $L_{2,h}$ с постоянной $K = 1$.

Окончательно получаем: для $1/2 \leq \theta \leq 1$ схема устойчива при любых τ и h , а для $0 \leq \theta < 1/2$ — устойчива при выполнении условия $\tau \leq h^2/(2 - 4\theta)$.

Здесь в основу решения положен следующий принцип: все собственные значения оператора перехода должны по модулю не превышать единицу. Это ограничение можно ослабить до величины $1 + \gamma\tau$ с постоянной γ , не зависящей от τ и h (аналогично спектральному признаку для гиперболических уравнений). В этом случае величина K в определении устойчивости принимает значение $K = e^{\gamma T}$.

27.11. Преобразуем схему к виду

$$\frac{u_m^{n+1} - u_m^{n-1}}{2\tau} + \frac{\tau^2}{h^2} \frac{u_m^{n+1} - 2u_m^n + u_m^{n-1}}{\tau^2} = \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2}.$$

Отсюда следует свойство аппроксимации.

Для анализа устойчивости воспользуемся частными решениями $u_m^n = \mu_h^n(k) \sin(\pi m k h)$, для которых достаточно показать справед-

ливость неравенств $|\mu_h(k)| \leq 1$. Из преобразованной схемы имеем

$$\frac{\mu^2 - 1}{2\tau} + \frac{\tau^2}{h^2} \frac{\mu^2 - 2\mu + 1}{\tau^2} = -\lambda^{(k)} \mu, \quad \lambda^{(k)} = \frac{4}{h^2} \sin^2 \frac{\pi k}{2M}, \quad Mh = 1,$$

или, с использованием обозначения $\rho = 2\tau/h^2$, уравнение для корней параболы

$$(\rho + 1)\mu^2 - 2\mu(\rho - \tau\lambda^{(k)}) + (\rho - 1) = 0.$$

Если дискриминант этого уравнения отрицательный, то

$$|\mu_1|^2 = |\mu_2|^2 = (\rho - 1)/(\rho + 1) < 1.$$

В случае положительного дискриминанта парабола в точках $\mu = \pm 1$ принимает положительные значения, а координата μ_* ее вершины равна $(\rho - \tau\lambda^{(k)})/(\rho + 1)$ и располагается внутри интервала $(-1, 1)$. Это гарантирует оценку $|\mu_{1,2}| < 1$ для вещественных корней. Таким образом, в обоих случаях получаем, что схема устойчива для всех τ и h .

28.1. $\Delta^h u - \Delta u = \left(\frac{h_1^2}{12} L_1^2 + \frac{h_2^2}{12} L_2^2 \right) u + O(h_1^4 + h_2^4) = O(h_1^2 + h_2^2).$

28.2. Шаблон «косой крест» — это обычный шаблон «крест» с шагом $\sqrt{2}h$ в системе координат, полученной поворотом исходной системы на $\pi/4$. Поэтому $\Delta^h = \Lambda_1 + \Lambda_2 + \frac{h^2}{2} \Lambda_1 \Lambda_2$, или $a_{0,0} = -2$ и $a_{i,j} = 1/2$ в остальных случаях. Погрешность аппроксимации равна $O(h^2)$.

28.3. $\Delta^h u_0 = \frac{2}{3h^2} (u_1 + u_2 + \dots + u_6 - 6u_0)$, где u_i при $1 \leq i \leq 6$ — значения в соседних с x_0 узлах сетки, отстоящих на расстояние h .

28.4. $\Delta^h u(A_0) = \frac{16}{h^2} \sin^2 \frac{\pi}{n} \left[-u(A_0) + \frac{1}{n} \sum_{k=1}^n u(A_k) \right].$

28.5. $a_{0,0} = 4c - 4, \quad a_{1,1} = a_{1,-1} = a_{-1,1} = a_{-1,-1} = c,$

$$a_{1,0} = a_{-1,0} = a_{0,1} = a_{0,-1} = 1 - 2c,$$

или, что то же самое, $\Delta^h = \Lambda_1 + \Lambda_2 + c h^2 \Lambda_1 \Lambda_2$, где c — произвольная постоянная.

28.6. Все операторы при $c < 1/2$.

28.7. $(\Delta^h)^2 u = (\Lambda_1^2 + 2\Lambda_1 \Lambda_2 + \Lambda_2^2) u = (\Lambda_1 + \Lambda_2)^2 u$, погрешность аппроксимации равна $O(h^2)$.

$$28.8. \quad h \sum_{j=1}^{N_2-1} \left(\frac{u_{N_1,j} - u_{N_1-1,j}}{h} - \frac{u_{1,j} - u_{0,j}}{h} \right) + h \sum_{i=1}^{N_1-1} \left(\frac{u_{i,N_2} - u_{i,N_2-1}}{h} - \frac{u_{i,1} - u_{i,0}}{h} \right) = 0.$$

$$28.9. \quad \left(\Lambda_1 + \Lambda_2 + \frac{h^2}{6} \Lambda_1 \Lambda_2 \right) u = f + \frac{h^2}{12} (\Lambda_1 + \Lambda_2) f + O(h^4).$$

$$28.10. \quad \left(\Lambda_1 + \Lambda_2 + \frac{h^2}{6} \Lambda_1 \Lambda_2 \right) u = f + \frac{h^2}{12} (\Lambda_1 + \Lambda_2) f - \frac{h^4}{240} (\Lambda_1^2 + \Lambda_2^2) f + \frac{h^4}{90} \Lambda_1 \Lambda_2 f + O(h^6).$$

$$28.11. \quad \frac{u_{1,j} - u_{0,j}}{h_1} - \frac{h_1}{2} (f_{0,j} - \Lambda_2 u_{0,j}) - \alpha u_{0,j} = 0.$$

28.13. Будем считать, что функция $u_{i,j}$ отлична от константы, так как в этом случае утверждение является тривиальным. Предположим далее противное, т. е. что минимальное значение достигается во внутреннем узле сетки (вообще, таких узлов может быть несколько). Пусть его номер (m, n) , тогда из неравенства $\Delta^h u_{m,n} \leq 0$ следует

$$u_{m,n} \geq \frac{1}{4} (u_{m+1,n} + u_{m-1,n} + u_{m,n+1} + u_{m,n-1}).$$

Знак $>$ не имеет места, так как $u_{m,n}$ — минимальное значение функции на сетке. Знак равенства приводит к тому, что в окрестности узла (m, n) значения функции $u_{i,j}$ постоянны: $u_{m+1,n} = u_{m-1,n} = u_{m,n+1} = u_{m,n-1} = u_{m,n}$. Продолжая рассуждения для этих узлов, затем для их соседей, получим в силу связности сетки, что при выполнении неравенства $\Delta^h u_{i,j} \leq 0$ функция обязана быть константой. Это противоречит исходной посылке, значит, наименьшее значение должно достигаться на границе, где указанное в условии неравенство места не имеет.

28.18. Уравнение, соответствующее узлу (i, j) , при всех $0 < i, j < N; Nh = 1$, записывается в виде

$$4u_{i,j} - u_{i+1,j} - u_{i-1,j} - u_{i,j+1} - u_{i,j-1} = - \int_{S_{ij}} f \varphi_{ij} dx dy,$$

где S_{ij} — носитель базисной функции φ_{ij} .

28.19. Уравнение, соответствующее узлу (i, j) , при всех $0 < i, j < N; Nh = 1$, записывается в виде

$$8u_{i,j}/3 - (u_{i,j+1} + u_{i,j-1} + u_{i-1,j} + u_{i+1,j} + u_{i+1,j+1} + u_{i+1,j-1} + u_{i-1,j+1} + u_{i-1,j-1})/3 = - \int_{S_{ij}} f \varphi_{ij} dx dy,$$

где S_{ij} — носитель базисной функции φ_{ij} , т. е. квадрат со стороной $2h$ и центром в узле (i, j) .

К главе VIII. Решение интегральных уравнений

29.1. Для отыскания приближений к решению $y(x)$ в точках $x = 0, 1/2, 1$ выпишем систему уравнений

$$\begin{aligned} y_1 &= 1, \\ \frac{1}{3}(e^{0,25} - 1)y_2 + \frac{1}{12}(e^{0,5} - 1)y_3 &= e^{0,5} - \frac{1}{2}, \\ \frac{2}{3}(e^{0,5} - 1)y_2 + \frac{1}{6}(e + 5)y_3 &= e - 1, \end{aligned}$$

или (с точностью до четырех знаков после запятой)

$$\begin{aligned} y_1 &= 1, \\ 1,0947y_2 + 0,0541y_3 &= 1,1487, \\ 0,4325y_2 + 1,2864y_3 &= 1,7183. \end{aligned}$$

Решение этой системы: $y_1 = 1$, $y_2 \approx 0,9999$, $y_3 \approx 0,9996$. Несложно проверить, что точное решение уравнения $y(x) \equiv 1$, поэтому абсолютная погрешность не превышает 0,0004.

29.2. Точное решение $y(x) = x$.

29.3. С тремя верными десятичными знаками получить приближение $z(x) = 1,003x + e^{-x}$, точное решение $y(x) = x + e^{-x}$.

29.4. Квадратурная формула трапеций для вычисления интеграла

$I_i = \int_{x_{i-1}}^{x_i} \varphi(s) ds$ по отрезку длины $h_i = x_i - x_{i-1}$ имеет вид

$$I_i \approx h_i \frac{\varphi(x_{i-1}) + \varphi(x_i)}{2}.$$

Поэтому для определения приближенного решения имеем систему уравнений

$$\begin{aligned} y_1 &= f(x_1), \\ y_i - \lambda \sum_{j=2}^i h_j \frac{K_{i,j-1}y_{j-1} + K_{i,j}y_j}{2} &= f(x_i), \quad i = 2, 3, \dots, n, \end{aligned}$$

где $K_{i,j} = K(x_i, s_j)$. Решение этой системы можно получить рекуррентно: $y_1 = f(x_1)$, далее для $i = 2, 3, \dots, n$

$$y_i = \frac{2f(x_i) + \lambda h_2 K_{i,1}y_1 + \lambda \sum_{j=2}^{i-1} (h_j + h_{j+1}) K_{i,j}y_j}{2 - \lambda h_i K_{i,i}},$$

если знаменатель не обращается в нуль.

Считается, что в этой формуле сумма равна нулю, если верхний предел суммирования меньше нижнего (для $i = 2$).

29.5. Точное решение $y(x) \equiv 1$.

29.6. Точное решение $y(x) = e^{2x}$.

29.7. Пусть $h = (b - a)/n$, тогда составная формула прямоугольников

для вычисления интеграла $\int_a^b \varphi(s) ds$ такова:

$$S_1^n(\varphi) = h \sum_{j=1}^n \varphi\left(a + h\left(j - \frac{1}{2}\right)\right),$$

а систему (29.2) можно записать в виде

$$\sum_{j=1}^n a_{ij} y_j = f(x_i), \quad 1 \leq i \leq n,$$

где $a_{ij} = \delta_i^j - \lambda K h$, δ_i^j — символ Кронекера.

29.8. В методе простой итерации $\mathbf{y}^{k+1} = B\mathbf{y}^k + \mathbf{f}$ для решения системы (29.2) при $K(x, s) \equiv K > 0$ несложно найти явный вид элементов матрицы B : $b_{ij} = \lambda K h$. Матрица размера $n \times n$ с постоянными элементами $\lambda K h$ имеет ядро размерности $(n - 1)$ и одно собственное значение, равное $n\lambda K h$. В этом легко убедиться, если рассмотреть действие матрицы на базисные векторы ядра $(-1, 1, 0, \dots, 0)^T, (-1, 0, 1, 0, \dots, 0)^T, \dots, (-1, 0, 0, \dots, 0, 1)^T$ и на базисный вектор образа $(1, 1, \dots, 1)^T$. Отсюда, по теореме о необходимом и достаточном условии сходимости метода простой итерации, получим, что сходимость с произвольного начального приближения имеется при выполнении условия $|n\lambda K h| = |\lambda|K(b - a)| < 1$.

29.9. Обозначим через $R_1^n(\varphi)$ остаточный член составной квадратурной

формулы прямоугольников $\int_a^b \varphi(s) ds = S_1^n(\varphi) + R_1^n(\varphi)$. Тогда система уравнений для точных значений решения в узлах x_i имеет вид

$$y(x_i) - \lambda K \frac{b-a}{n} \sum_{j=1}^n y(x_j) = f(x_i) + \lambda K R_1^n(y)|_{x=x_i}, \quad 1 \leq i \leq n.$$

Введя для компонент погрешности и остаточного члена обозначения $e_i = y(x_i) - y_i$ и $r_i = R_1^n(y)|_{x=x_i}$, получим систему $A\mathbf{e} = \lambda K \mathbf{r}$. В силу неравенства $|\lambda|K(b - a) < \alpha < 1$ матрица A имеет диагональное преобладание, поэтому $\|\mathbf{e}\|_\infty \leq |\lambda|K \|A^{-1}\|_\infty \|\mathbf{r}\|_\infty$. Для остаточного члена составной квадратурной формулы прямоугольников справедлива оценка

$$|R_1^n(y)| \leq \|y''\|_C \frac{(b-a)^3}{24n^2},$$

откуда следует неравенство

$$\|\mathbf{r}\|_\infty \leq \|y''\|_C \frac{(b-a)^3}{24n^2}.$$

Оценим величину $\|A^{-1}\|_\infty$. Так как матрица обладает свойством диагонального преобладания, то на основании решения **14.24** имеем

$$\|A^{-1}\|_\infty \leq \frac{1}{1-\alpha} \frac{1}{1-|\lambda|Kh}, \quad h = (b-a)/n.$$

Отсюда следует окончательная оценка

$$\|e\|_\infty \leq \frac{1}{1-\alpha} \frac{|\lambda|K}{1-|\lambda|Kh} \|y''(x)\|_C \frac{(b-a)^3}{24n^2} = O(h^2).$$

Таким образом, если система линейных уравнений «не слишком плоха» и решение обладает достаточной гладкостью, то порядок погрешности в узлах x_i совпадает с порядком остаточного члена составной квадратурной формулы.

29.11. Вместо исходного рассмотрим интегральное уравнение

$$z(x) + \int_0^1 H(x,s)z(s) ds = e^x - x,$$

решение которого будем искать в виде

$$z(x) = e^x - x + D_1x^2 + D_2x^3 + D_3x^4.$$

Для определения постоянных D_1, D_2, D_3 получим систему

$$\begin{aligned} \frac{5}{4} D_1 + \frac{1}{5} D_2 + \frac{1}{6} D_3 &= -\frac{2}{3}, \\ \frac{1}{5} D_1 + \frac{13}{6} D_2 + \frac{1}{7} D_3 &= \frac{9}{4} - e, \\ \frac{1}{6} D_1 + \frac{1}{7} D_2 + \frac{49}{8} D_3 &= 2e - \frac{29}{5}. \end{aligned}$$

Ее решение дает: $D_1 \approx -0,5010, D_2 \approx -0,1671, D_3 \approx -0,0422$. Точное решение уравнения $y(x) \equiv 1$. Несложно проверить, что абсолютная погрешность не превышает 0,008.

29.12. Точное решение $y(x) = x + 1/2$.

30.1. Для отыскания приближенного решения $z(x) = c_1 + c_2x$ получим систему

$$\frac{8}{3} c_1 - \frac{2}{9} c_2 = -\frac{2}{3}, \quad -\frac{2}{9} c_1 + \frac{2}{27} c_2 = \frac{2}{9}.$$

Ее решение имеет вид $c_1 = 0, c_2 = 3$, а найденное приближенное решение интегрального уравнения совпадает с точным: $z(x) = y(x) = 3x$.

30.2. Для отыскания приближенного решения $z(x) = 1 + c_1x + c_2x^2$ получим систему

$$0 \cdot c_1 + \frac{2}{9} c_2 = \frac{4}{3}, \quad \frac{2}{9} c_1 + 0 \cdot c_2 = 0.$$

Ее решение имеет вид $c_1 = 0$, $c_2 = 6$, а найденное приближенное решение интегрального уравнения совпадает с точным: $z(x) = y(x) = 1 + 6x^2$.

30.3. Для отыскания приближенного решения $z(x) = 1 + c_1x + c_2x^2$ получим систему

$$\frac{2}{9} c_1 + 0 \cdot c_2 = 0, \quad 0 \cdot c_1 + \frac{2}{15} c_2 = \frac{4}{5},$$

Ее решение имеет вид $c_1 = 0$, $c_2 = 6$, а найденное приближенное решение интегрального уравнения совпадает с точным: $z(x) = y(x) = 1 + 6x^2$.

30.4. Для приближенных собственных функций вида $z(x) = c_1 + c_2x(1-x) + c_3x(1-x)(1-2x)$ с неизвестными пока коэффициентами имеем уравнения:

$$\int_0^1 \varphi_i(x) R z(x) dx = 0, \quad 1 \leq i \leq 3.$$

Вычисляя интегралы, получаем систему

$$\begin{aligned} c_1 \left(1 - \frac{\lambda}{12}\right) + \frac{c_2}{6} \left(1 - \frac{\lambda}{10}\right) &= \frac{c_1}{6} \left(1 - \frac{\lambda}{10}\right) + \frac{c_2}{30} \left(1 - \frac{17\lambda}{168}\right) = \\ &= \frac{c_3}{210} \left(1 - \frac{\lambda}{40}\right) = 0. \end{aligned}$$

Приравнявая определитель полученной системы к нулю, приближенно (с четырьмя знаками после запятой) получим:

$$\lambda_1 = 9,8751, \quad \lambda_2 = 40, \quad \lambda_3 = 170,1249.$$

Отсюда, учитывая условие нормировки $\int_0^1 [z(x)]^2 dx = 1$, определим

$$z_1(x) = -0,0684 + 5,817x(1-x), \quad z_2(x) = 14,49x(1-x)(1-2x).$$

Точное решение задачи имеет вид: $\lambda_k = (k\pi)^2$, $y_k(x) = \sqrt{2} \sin k\pi x$, $k = 1, 2, \dots$

30.5. $\lambda_1 = \frac{4}{\pi}$, $y_1(x) = c_1 \cos^2 x$, $\lambda_2 = \frac{8}{\pi}$, $y_1(x) = c_2 \cos 3x$.

30.6. Для отыскания приближенного решения в виде

$$z(x) = c_1 + c_2x + c_3 \frac{3x^2 - 1}{2}$$

подставим его в исходное уравнение. После интегрирования в правой части получим

$$c_1 + c_2x + c_3 \frac{3x^2 - 1}{2} = 1 + 2xc_1 + \frac{2}{3}xc_2.$$

Последовательно умножаем равенство на функции 1 , x , $\frac{3x^2 - 1}{2}$ и интегрируем по отрезку $[-1, 1]$. Это дает систему линейных уравнений

$$2c_1 = 2 + \frac{4}{3}c_1, \quad \frac{2}{3}c_2 = \frac{4}{9}c_2, \quad \frac{5}{2}c_3 = \frac{8}{15}c_1,$$

откуда находим $c_1 = 3$, $c_2 = 0$, $c_3 = 4$, т. е. $z(x) = y(x) = 1 + 6x^2$.

30.8. Для отыскания приближенного решения в виде

$$z(x) = c_1 + c_2x + c_3 \frac{3x^2 - 1}{2}$$

получим невязку интегрального уравнения

$$Rz(x) = c_1 \left(1 + \frac{4}{3}x\right) + c_2x + c_3 \left(\frac{3x^2 - 1}{2} + \frac{7}{3}x\right) - 1 - \frac{4}{3}x.$$

Если потребовать, чтобы она обращалась в нуль в точках $-1, 0, 1$, получим систему линейных уравнений для отыскания неизвестных коэффициентов:

$$\frac{1}{3}c_1 + c_2 + \frac{4}{3}c_3 = \frac{1}{3}, \quad c_1 = 1, \quad \frac{7}{3}c_1 + c_2 + \frac{10}{3}c_3 = \frac{7}{3},$$

откуда находим $c_1 = 1$, $c_2 = 0$, $c_3 = 0$, т. е. $z(x) = y(x) \equiv 1$.

31.1. Будем искать решение уравнения (31.1) в виде формального ряда

$$y(x) = \sum_{i=1}^{\infty} y_i z_i(x), \quad (4)$$

где $\{z_i(x)\}_{i=1}^{\infty}$ — полная, ортонормированная в $L_2(a, b)$ система собственных функций интегрального оператора G_b . Из гладкости правой части следует сходимость ряда

$$f(x) = \sum_{i=1}^{\infty} f_i z_i(x), \quad f_i = \int_a^b f(x) z_i(x) dx.$$

После подстановки в исходное уравнение предполагаемого вида решения (4) и разложения ядра и правой части по ортонормированному базису $\{z_i(x)\}_{i=1}^{\infty}$ получим:

$$y_i = f_i / \lambda_i, \quad i = 1, 2, \dots$$

Таким образом, сходимость функционального ряда (4) определяется сходимостью числового ряда

$$S = \sum_{i=1}^{\infty} |y_i|^2 = \sum_{i=1}^{\infty} |f_i/\lambda_i|^2.$$

Будем считать, что неявно наложенное условие гладкости на правую часть $f(x)$ обеспечивает выполнение неравенства $S < \infty$. Тогда решение поставленной задачи существует и единственно.

31.2. При указанных в условии свойствах характеристических чисел ядро интегрального оператора имеет вид

$$K(x, s) = \sum_{i=1}^p \lambda_i z_i(x) z_i(s),$$

т. е. является вырожденным. Поэтому на основании § 29 решение уравнения (31.1) представимо конечной суммой

$$y(x) = \sum_{i=1}^p y_i z_i(x)$$

с некоторыми коэффициентами y_i . Пусть правая часть уравнения (31.1) имеет вид $f(x) = f_k z_k(x)$, $f_k \neq 0$ при некотором $k > p$. Тогда равенство $G_b y(x) = f(x)$, т. е.

$$\int_a^b \left(\sum_{j=1}^p \lambda_j z_j(x) z_j(s) \right) \left(\sum_{i=1}^p y_i z_i(s) \right) ds = f_k z_k(x),$$

противоречит свойству ортонормированности системы собственных функций $\{z_i(x)\}_{i=1}^{\infty}$ интегрального оператора G_b . Поэтому решение поставленной задачи $y(x)$ существует только для правых частей, представимых в виде

$$f(x) = \sum_{i=1}^p f_i z_i(x).$$

При этом, в силу **31.1**, неизвестные коэффициенты в $y(x)$ определяются равенствами $y_i = f_i/\lambda_i$, $i = 1, 2, \dots, p$.

Теперь, чтобы установить неединственность решения для таких правых частей, достаточно проверить, что произвольная функция

$$y(x) = \sum_{i=1}^p \frac{f_i}{\lambda_i} z_i(x) + \sum_{i=p+1}^{\infty} c_i z_i(x)$$

с коэффициентами, удовлетворяющими условию $\sum_{i=p+1}^{\infty} |c_i|^2 < \infty$, также будет являться решением исходного интегрального уравнения.

31.3. Запишем решения точного и регуляризованного уравнений в виде разложений по собственным функциям интегрального оператора (см. решение **31.1**):

$$y = \sum_{i=1}^{\infty} \frac{(f, z_i)}{\lambda_i} z_i, \quad y_{\alpha} = \sum_{i=1}^{\infty} \frac{(f_{\delta}, z_i)}{\alpha + \lambda_i} z_i.$$

Обозначим через u решение регуляризованного уравнения с точной правой частью:

$$\alpha u + G_b u = f, \quad u = \sum_{i=1}^{\infty} \frac{(f, z_i)}{\alpha + \lambda_i} z_i.$$

Оно нам потребуется для получения оценки по неравенству треугольника: $\|y_{\alpha} - y\| \leq \|y_{\alpha} - u\| + \|u - y\|$. В силу ортонормированности системы собственных функций $\{z_i(x)\}_{i=1}^{\infty}$, для оценки слагаемых в правой части можно воспользоваться равенством Парсеваля:

$$\|y_{\alpha} - u\|^2 = \sum_{i=1}^{\infty} \frac{(f_{\delta} - f, z_i)^2}{(\alpha + \lambda_i)^2} \leq \frac{\delta^2}{\alpha^2},$$

$$\|u - y\|^2 = \alpha^2 \sum_{i=1}^{\infty} \frac{(f, z_i)^2}{(\alpha + \lambda_i)^2 \lambda_i^2} \leq \alpha^2 \sum_{i=1}^{\infty} \frac{(f, z_i)^2}{\lambda_i^4} \leq \alpha^2 M^2.$$

В последней оценке было использовано неравенство из условия задачи. Из полученных выражений следует $\|y_{\alpha} - y\| \leq \delta/\alpha + \alpha M$. Минимум правой части достигается при $\alpha = \sqrt{\delta/M}$, что приводит к неравенству $\|y_{\alpha} - y\| \leq 2\sqrt{\delta M}$. Эта оценка означает сходимость: $y_{\alpha} \rightarrow y$ при $\delta \rightarrow 0$.

31.4. Уравнение имеет вид $\alpha y_{\alpha} + Q_b y_{\alpha} = \varphi$, где $\varphi(x) = \int_a^b K(s, x) f(s) ds$,

а ядро $H(x, s)$ в новом интегральном операторе $Q_b = G_b^* G_b$ определяется через исходное $K(x, s)$ по формуле

$$H(x, s) = \int_a^b K(t, x) K(t, s) dt \text{ и является симметричным.}$$

31.6. След четного порядка для симметричного ядра определяется по формуле: $S_{2i} = \int_a^b \int_a^b K_i^2(x, s) dx ds$, где $K_1 = K(x, s)$,

$$K_i(x, s) = \int_a^b K(x, t) K_{i-1}(t, s) dt, \quad i = 2, 3, \dots$$

31.8. $\lambda_{\min} = 3$.

Литература



1. Арушанян И. О., Чижонков Е. В. Материалы семинарских занятий по курсу «Методы вычислений» /под ред. О. Б. Арушаняна — М.: Изд-во ЦПИ при механико-математическом факультете МГУ, 1999.
2. Бабенко К. И. Основы численного анализа. — М.: Наука, 1986.
3. Бахвалов Н. С. Численные методы. — М.: Наука, 1975.
4. Бахвалов Н. С., Жидков Н. П., Кобельков Г. М. Численные методы. — 6-е изд. — М.: БИНОМ. Лаборатория знаний, 2008.
5. Бахвалов Н. С., Корнев А. А., Чижонков Е. В. Численные методы. Решения задач и упражнения. — М.: Дрофа, 2009.
6. Воеводин В. В., Кузнецов Ю. А. Матрицы и вычисления. — М.: Наука, 1984.
7. Годунов С. К., Рябенский В. С. Разностные схемы. — М.: Наука, 1977.
8. Голуб Дж., Ван Лоун Ч. Матричные вычисления. — М.: Мир, 1999.
9. Дробышев В. И., Дымников В. П., Ривин Г. С. Задачи по вычислительной математике. — М.: Наука, 1980.
10. Икрамов Х. Д. Несимметричная проблема собственных значений. Численные методы. — М.: Наука, 1991.
11. Лебедев В. И. Функциональный анализ и вычислительная математика. — М.: ФИЗМАТЛИТ, 2005.
12. Марчук Г. И. Методы вычислительной математики. — М.: Наука, 1980.
13. Самарский А. А. Теория разностных схем. — М.: Наука, 1983.
14. Самарский А. А., Вабищевич П. Н., Самарская Е. А. Задачи и упражнения по численным методам. — М.: Эдиториал УРСС, 2000.
15. Сборник задач по численным методам / Сост. Стрелков Н. А. — Ярославль: Изд-во Яросл. ун-та, 1988.
16. Тьртыйшиников Е. Е. Методы численного анализа. — М.: Издательский центр «Академия», 2007.
17. Уилкинсон Дж. Х. Алгебраическая проблема собственных значений. — М.: Наука, 1970.
18. Численные методы решения задач алгебры и дифференциальных уравнений / Сост. Стрелков Н. А. — Ярославль: Изд-во Яросл. ун-та, 1989.

Предметный указатель



- Абсолютная ошибка** 84
алгебраический порядок точности 40
алгоритм *Евклида* 97
анализ устойчивости схем
 в интегральной метрике 129
 --- в равномерной метрике 128
аппроксимация 102
 безусловная 103
 на решении 103
 условная 103
- Вектор невязки** 62
 ошибки 67
вычисление интегралов от функций с особенностями 55
- Дифференциальные уравнения**
 аппроксимация 102
 -- постановки задач 101
 -- решение 101
дифференцирование численное 21
- Задача Коши** 101, 114
 краевая линейная 118
 -- смешанная 101
задачи на собственные значения 78, 99
 некорректные 143
- Интегральное тождество Марчука** 107
интегрирование численное 34
исследование устойчивости
 методом априорных оценок 121
итерации высшего порядка 89
- Квадратура** 34
 коэффициент 34
 узел 34
квадратурная формула
 Маркова—Лобатто 49
 -- *Маркова—Радо* 49
квадратурные формулы *Гаусса* 44
 -- интерполяционного типа 34
 -- составные 36
 -- *Чебышёва* 40
константа *Лебега*
 интерполяционного процесса 12
кратность корня 84
- Мантисса** 6
 длина 6
матрица *Гильберта* 30
 -- *Коши* 65
 плохо обусловленная 62
 простой структуры 70
 со сдвигом 82
матрицы подобные 78
матричные вычисления 57
машинная точность 7
метод априорных оценок 121
 -- *Бубнова—Галёркина* 140
 -- верхней релаксации 74
 -- *Галёркина* 109
 -- *Гаусса—Зейделя* 74
 -- замены интеграла 136
 -- замены ядра 136
 -- интегро-интерполяционный 106
 -- коллокации 141
 -- конечных элементов 109
 -- моментов 141
 -- наименьших квадратов 139

- наискорейшего градиентного спуска 73
 - неопределенных коэффициентов 22, 39, 105
 - *Ньютона* 89
 - обобщенный простой итерации 73
 - обратной итерации 82
 - оптимальный линейный одношаговый 67
 - парабол 86
 - *Петрова—Галёркина* 140
 - простой итерации 66, 85
 - регуляризации *Тихонова* 144
 - релаксации 74
 - *Рунца* 108
 - с предобуславливателем 73
 - секущих 86
 - сеток 101
 - симметричный релаксации 74
 - степенной 82
 - хорд 86
 - *Чебышёва* 90
 - *Якоби* 74
 - методы неявные 73
 - релаксации 73
 - многочлен алгебраический 11
 - *Лагранжа* 11
 - наилучшего равномерного приближения 26
 - многочлены *Чебышёва* второго рода 21
 - первого рода 17, 45
 - приведенный 18
 - разностное уравнение 17
 - рекуррентное соотношение 17
 - тригонометрическая форма 17
 - множество нормализованное 6
- Неравенства для сеточных функций** 122
- неравенство *Йенсена* 59
- норма вектора 57
 - матрицы 57
 - равномерная 11
 - нормы эквивалентные 57
- Оператор Лапласа** 131
- ортогональные многочлены 45
- *Лагерра* 45
 - *Лежандра* 45
 - *Чебышёва* первого рода 45
 - *Эрмита* 45
- отношение *Рэлея* 82
- Параметр релаксации** 74
- переполнение 7
- погрешность абсолютная 6
- предельная 9
 - вычислительная 6
 - главный член 49
 - относительная 6
 - предельная 9
 - решения задачи 6
 - функции 9
- полиномиальная интерполяция 11
- правило *Рунге* 50
- приближение сплайнами 30
- функций и производных 11
- проекционные методы 139
- простейшая регуляризация 144
- Разностные схемы методы построения** 105
- разность вперед 22
- назад 22
 - центральная 22
- рекуррентное соотношение 17
- решение дифференциальных уравнений 101
- интегральных уравнений 135
 - нелинейных уравнений 84
 - разностное 102
 - сходимость 104

- точное 62
- фундаментальное 99

- Спектральный признак**
 - устойчивости 125
- радиус матрицы 61
- сплайн интерполяционный** 30
 - локальный (аппроксимационный) 32
 - полиномиальный 30
- схема l -точечная** 117
 - Горнера 8
 - Дюфорта–Франкела 130
- сходимость решения** 104
 - скорость 84

- Теорема Чебышёва** 26
 - точки чебышёвского альтернанса 26

- Узлы равноотстоящие** 12
 - уравнение *Вольтерры* 135
 - Пуассона 131
 - разностное 17
 - Фредгольма 135
 - уравнения гиперболические 123
 - параболические 126
 - разностные 94, 97
 - эллиптические 130
 - устойчивость 103
 - безусловная 104
 - по граничным условиям 104
 - по начальным данным 104
 - по правой части 104
 - условная 104

- Формула Абеля** 121
 - прямоугольников 164
 - Симпсона 164
 - Стирлинга 183
 - трапеций 164
 - Филона 54
- формулы вспомогательные** 120
 - Грина 121
 - Ньютона–Котеса 35
- функции быстро осциллирующие** 53
 - координатные 139

- Числа с плавающей запятой** 6
 - численное дифференцирование 21
 - интегрирование 34
 - функций с особенностями 53
 - число обусловленности 62

- Элементы теории возмущений** 62

- Ядро интегрального оператора** 135

Оглавление



Предисловие ко второму изданию	3
Предисловие к первому изданию	4
Глава I. Погрешность решения задачи	6
§ 1. Вычислительная погрешность	6
§ 2. Погрешность функции	9
Глава II. Приближение функций и производных	11
§ 3. Полиномиальная интерполяция	11
§ 4. Многочлены Чебышёва	17
§ 5. Численное дифференцирование	21
§ 6. Многочлен наилучшего равномерного приближения	26
§ 7. Приближение сплайнами	30
Глава III. Численное интегрирование	34
§ 8. Квадратурные формулы интерполяционного типа	34
§ 9. Метод неопределенных коэффициентов	39
§ 10. Квадратурные формулы Гаусса	44
§ 11. Главный член погрешности	49
§ 12. Численное интегрирование функций с особенностями	53
Глава IV. Матричные вычисления	57
§ 13. Векторные и матричные нормы	57
§ 14. Элементы теории возмущений	62
§ 15. Метод простой итерации	66
§ 16. Методы релаксации	73
§ 17. Задачи на собственные значения	78
Глава V. Решение нелинейных уравнений	84
§ 18. Метод простой итерации и смежные вопросы	85
§ 19. Метод Ньютона. Итерации высшего порядка	89
Глава VI. Разностные уравнения	94
§ 20. Однородные разностные уравнения	94
§ 21. Неоднородные разностные уравнения	97
§ 22. Фундаментальное решение и задачи на собственные значения	99

Глава VII. Решение дифференциальных уравнений	101
§ 23. Методы построения разностных схем	105
§ 24. Задача Коши	114
§ 25. Линейная краевая задача	118
§ 26. Гиперболические уравнения	123
§ 27. Параболические уравнения	126
§ 28. Эллиптические уравнения	130
Глава VIII. Решение интегральных уравнений	135
§ 29. Методы замены интеграла и ядра	135
§ 30. Проекционные методы	139
§ 31. Некорректные задачи	143
Ответы, указания, решения	146
К главе I. Погрешность решения задачи	146
К главе II. Приближение функций и производных	153
К главе III. Численное интегрирование	164
К главе IV. Матричные вычисления	174
К главе V. Решение нелинейных уравнений	192
К главе VI. Разностные уравнения	200
К главе VII. Решение дифференциальных уравнений	209
К главе VIII. Решение интегральных уравнений	228
Литература	235
Предметный указатель	236

Минимальные системные требования определяются соответствующими требованиями программы Adobe Reader версии не ниже 11-й для платформ Windows, Mac OS, Android, iOS, Windows Phone и BlackBerry; экран 10"

Учебное электронное издание

Бахвалов Николай Сергеевич
Лапин Александр Васильевич
Чижонков Евгений Владимирович

**ЧИСЛЕННЫЕ МЕТОДЫ
В ЗАДАЧАХ И УПРАЖНЕНИЯХ**
Учебное пособие

Ведущий редактор *М. С. Стригунова*
Редактор *А. С. Попов*
Художник *С. Инфантэ*

Технический редактор *Е. В. Денюкова*
Оригинал-макет подготовлен *О. Г. Лапко* в пакете $\LaTeX 2\epsilon$

Подписано к использованию 19.03.15.
Формат 125×200 мм

Издательство «БИНОМ. Лаборатория знаний»
125167, Москва, проезд Аэропорта, д. 3
Телефон: (499) 157-5272
e-mail: info@pilotLZ.ru, <http://www.pilotLZ.ru>