

← **ВСЕ**

ОСНОВНЫЕ РАЗДЕЛЫ
ЧИСЛЕННОГО АНАЛИЗА

**ЗАДАЧИ И
УПРАЖНЕНИЯ**

Ч ПО

М

**ИСПЕННЫМ
МЕТОДАМ**

**А.А. Самарский
П.Н. Вабищевич
Е.А. Самарская**



УРСС

РОССИЙСКАЯ АКАДЕМИЯ НАУК
ИНСТИТУТ МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ
МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМ. М. В. ЛОМОНОСОВА

А.А. Самарский
П.Н. Вабищевич
Е.А. Самарская

**ЗАДАЧИ И
УПРАЖНЕНИЯ**

Ч ^{ПО}
ИСЛЕННЫМ
М **МЕТОДАМ**

Эдиториал УРСС • Москва • 2000





Настоящее издание осуществлено при финансовой поддержке Российского фонда фундаментальных исследований (проект № 99-01-14021)

**Самарский Александр Андреевич,
Вабищевич Петр Николаевич,
Самарская Елена Александровна**

Задачи и упражнения по численным методам: Учебное пособие. — М.: Эдиториал УРСС, 2000. — 208 с.

ISBN 5-8360-0158-8

Учебное пособие поддерживает курс по численным методам, который читается в вузах с повышенной математической подготовкой. Задачи и упражнения охватывают все основные разделы численного анализа: интерполирование функций, численное интегрирование, прямые и итерационные методы линейной алгебры, спектральные задачи, системы нелинейных уравнений, задачи минимизации функций, интегральные уравнения, краевые задачи и задачи с начальными данными для обыкновенных уравнений и уравнений с частными производными. Каждый раздел содержит небольшой справочный материал, упражнения (задачи с решениями) и набор задач для самостоятельной работы.

Книга рассчитана на студентов университетов и вузов, обучающихся по специальности «Прикладная математика».

Директор — Доминго Марин Рикой

Заместители директора — Наталья Финогенова, Ирина Макеева

Администратор — Леонид Иосилевич

Компьютерный дизайн — Виктор Романов

Верстка — Николай Вабищевич, Наталия Бекетова

Редакционно-корректорские работы — Елена Кудряшова, Сергей Шаракин

Дизайн обложки — Ирина Макеева

Техническая поддержка — Вадим Устьянский, Наталья Аринчева

Менеджер по продажам — Алексей Петяев

Издательство «Эдиториал УРСС» 113208, г Москва, ул Чертановская, д 2/11, к п
Лицензия ЛР № 064418 от 24 01 96 г Гигиенический сертификат на выпуск книжной
продукции № 77 ФЦ 8 953 П 270.3 99 от 30 03 99 г. Подписано к печати 06 07 2000 г
Формат 60×84/16 Тираж 1000 экз Печ л 13 Зак № 12 P

Отпечатано в ТОО «Типография ПЭМ» 121471, г Москва, Можайское шоссе, 25

Эдиториал УРСС
научная и учебная литература



Тел./факс: 7(095)135-44-23

Тел./факс: 7(095)135-42-46

E-mail: urss@urss.ru

Каталог изданий в Internet: <http://urss.ru>

ISBN 5-8360-0158-8

© А. А. Самарский,
П. Н. Вабищевич,
Е. А. Самарская, 2000
© Эдиториал УРСС, 2000

Оглавление

Предисловие	6
Глава 1. Интерполирование и приближение функций	8
1.1. Задачи интерполяции и приближения функций	8
1.2. Алгоритмы интерполяции и приближения функций	10
1.2.1. Полиномиальная интерполяция	10
1.2.2. Интерполяционные сплайны	11
1.2.3. Приближение функций в нормированном пространстве	12
1.3. Упражнения	13
1.4. Задачи	18
Глава 2. Численное интегрирование	23
2.1. Задачи приближенного вычисления интегралов	23
2.2. Алгоритмы приближенного вычисления интегралов	24
2.2.1. Классические квадратурные формулы составного типа	25
2.2.2. Квадратурные формулы интерполяционного типа	26
2.2.3. Квадратурные формулы Гаусса	27
2.3. Упражнения	28
2.4. Задачи	32
Глава 3. Прямые методы линейной алгебры	35
3.1. Задачи решения систем линейных уравнений	35
3.2. Алгоритмы решения систем линейных уравнений	36
3.2.1. Обусловленность матрицы и оценки точности решения систем линейных уравнений	37
3.2.2. Метод Гаусса для решения систем линейных уравнений	38
3.2.3. Метод квадратного корня	39
3.3. Упражнения	40
3.4. Задачи	44
Глава 4. Итерационные методы линейной алгебры	48
4.1. Итерационное решение систем линейных уравнений	48
4.2. Итерационные алгоритмы линейной алгебры	50
4.2.1. Классические итерационные методы	51
4.2.2. Двухслойные итерационные методы	53
4.2.3. Итерационные методы вариационного типа	55

4.3. Упражнения	56
4.4. Задачи	61
Глава 5. Спектральные задачи линейной алгебры	65
5.1. Собственные значения и собственные вектора матриц	65
5.2. Численные методы решения задач на собственные значения	66
5.2.1. Свойства собственных значений и собственных векторов	67
5.2.2. Итерационные методы решения частичной проблемы собственных значений	69
5.2.3. Решение полной проблемы собственных значений	70
5.3. Упражнения	72
5.4. Задачи	77
Глава 6. Нелинейные уравнения и системы	79
6.1. Решение нелинейных уравнений и систем	79
6.2. Итерационные методы решения нелинейных уравнений	80
6.2.1. Алгоритмы для решения нелинейного уравнения	80
6.2.2. Методы решения систем нелинейных уравнений	82
6.3. Упражнения	84
6.4. Задачи	88
Глава 7. Задачи минимизации функций	90
7.1. Поиск минимума функции многих переменных	90
7.2. Методы решения задач оптимизации	91
7.2.1. Поиск минимума функции одной переменной	91
7.2.2. Минимизация функций многих переменных	93
7.2.3. Задачи условной минимизации	95
7.3. Упражнения	96
7.4. Задачи	99
Глава 8. Интегральные уравнения	101
8.1. Задачи для интегральных уравнений	101
8.2. Методы решения интегральных уравнений	103
8.2.1. Интегральные уравнения Фредгольма второго рода	103
8.2.2. Интегральные уравнения с переменными пределами интегрирования	106
8.2.3. Интегральное уравнение Фредгольма первого рода	107
8.3. Упражнения	109
8.4. Задачи	114

Глава 9. Задача Коши для дифференциальных уравнений	118
9.1. Задачи с начальными условиями для систем обыкновенных дифференциальных уравнений	118
9.2. Численные методы решения задачи Коши	119
9.2.1. Методы Рунге—Кутты	119
9.2.2. Многошаговые методы	122
9.2.3. Жесткие системы обыкновенных дифференциальных уравнений	124
9.3. Упражнения	125
9.4. Задачи	131
Глава 10. Краевые задачи для дифференциальных уравнений	134
10.1. Краевые задачи	134
10.2. Численные методы решения краевых задач	137
10.2.1. Аппроксимация краевых задач	137
10.2.2. Сходимость разностных схем	141
10.2.3. Другие задачи	144
10.2.4. Решение сеточных уравнений	145
10.3. Упражнения	146
10.4. Задачи	152
Глава 11. Краевые задачи для эллиптических уравнений	158
11.1. Двумерные краевые задачи	158
11.2. Численное решение краевых задач	160
11.2.1. Аппроксимация краевых задач для эллиптических уравнений	160
11.2.2. Принцип максимума	161
11.2.3. Разностные уравнения в гильбертовом пространстве	163
11.2.4. Решение сеточных уравнений	165
11.3. Упражнения	169
11.4. Задачи	175
Глава 12. Нестационарные задачи математической физики	180
12.1. Нестационарные краевые задачи	180
12.2. Разностные методы решения нестационарных задач	183
12.2.1. Устойчивость двухслойных операторно-разностных схем	183
12.2.2. Устойчивость трехслойных разностных схем	187
12.2.3. Разностные схемы для параболического уравнения	189
12.2.4. Гиперболические уравнения	191
12.2.5. Многомерные задачи	192
12.3. Упражнения	194
12.4. Задачи	202
Литература	206

Предисловие

Курс по численным методам является основным при подготовке специалистов по прикладной и вычислительной математике. В нем излагаются основы численных методов решения задач алгебры, анализа, обыкновенных дифференциальных уравнений и уравнений с частными производными, с необходимой полнотой изучаются вопросы построения и теоретического обоснования вычислительных алгоритмов. Ставится и решается задача подготовки слушателей к практическому использованию численных методов при решении прикладных задач.

Поддержка курса по численным методам проводится как в теоретическом, так и в практическом плане. Закрепление базового материала по теории происходит на семинарских занятиях по численным методам. Навыки грамотного практического использования численных методов закладываются в вычислительном практикуме. С использованием современных библиотек численного анализа на компьютерах проводится содержательный анализ возможностей вычислительных алгоритмов при решении типовых задач. Высокая техническая оснащенность, рост возможностей вычислительной техники позволяет существенно обогатить содержание вычислительного практикума по численным методам.

Предлагаемое учебное пособие ориентировано на закрепление слушателями теоретического материала по курсу численных методов. Семинарские и самостоятельные занятия направлены на формирование навыков построения вычислительных алгоритмов для решения базовых задач численного анализа, теоретического исследования свойств алгоритма (точность, устойчивость, вычислительная работа на реализацию и т. д.).

Предлагаемая книга построена по следующему плану. Выделены основные, относительно самостоятельные разделы численного анализа. В отдельных главах рассмотрены задачи интерполирования и приближения функций, численного интегрирования, прямые и итерационные методы линейной алгебры, спектральные задачи линейной алгебры, системы нелинейных уравнений, задачи минимизации функций, интегральные уравнения, краевые задачи и задачи с начальными данными для обыкновенных уравнений, стационарные и нестационарные задачи математической физики.

Каждая глава (раздел численного анализа) начинается с формулировки задачи и приведения основных фактов по построению и исследованию вычислительных алгоритмов для выделенного класса задач. Этот материал не претендует на полноту, а лишь ориентирует читателя при изучении материала курса по численным методам. Дано небольшое число задач с решениями (упражнений) демонстрационного плана. Основное внимание уделяется задачам, предназначенным для самостоятельного решения. В ряде случаев задачи сформулированы в достаточно общем плане, который допускает исследование проблемы с различной глубиной исследования.

Проблемы подготовки задачника по численным методам широко и заинтересованно обсуждалась на кафедре вычислительных методов факультета вычислительной математики и кибернетики МГУ им. М. В. Ломоносова в течении длительного времени. Особенно полезными для нас были соображения А. В. Гулина и Е. С. Николаева. Авторы с благодарностью воспримут конструктивные замечания по нашей работе, особенно в части уточнения набора задач.

А. А. Самарский

П. Н. Вабищевич

Е. А. Самарская

Москва, июнь 2000 г.

Глава 1

Интерполирование и приближение функций

Рассматриваются задачи приближенного восстановления значений функции одной переменной по ее значениям в некоторых точках. Традиционный подход для одномерной интерполяции связан с построением алгебраических многочленов, принимающих заданные значения в точках интерполяции. Более перспективными являются подходы с использованием кусочно-гладких полиномов. Отдельно выделены задачи приближения функций в нормированных пространствах.

1.1. Задачи интерполяции и приближения функций

Задача интерполяции ставится следующим образом. Пусть на отрезке $[a, b]$ в узлах интерполирования $x_0 < x_1 < \dots < x_n$ известны значения функции $y_i = f(x_i)$, $i = 0, 1, \dots, n$. Необходимо найти значения функции $f(x)$ в точках $x \neq x_i$, $i = 0, 1, \dots, n$.

Пусть на отрезке $[a, b]$ задана система функций $\{\varphi_i(x)\}_{i=0}^n$ и определим

$$\varphi(x) = \sum_{i=0}^n c_i \varphi_i(x) \quad (1.1)$$

с действительными коэффициентами c_i , $i = 0, 1, \dots, n$. При интерполировании функции $f(x)$ для нахождения коэффициентов используются условия

$$\varphi(x_i) = f(x_i), \quad i = 0, 1, \dots, n. \quad (1.2)$$

В частном случае алгебраической интерполяции $\varphi_i(x) = x^i$, $i = 0, 1, \dots, n$.

Основные обозначения

	$f(x)$	— интерполируемая функция
$x_0 < x_1 < \dots < x_n$		— узлы интерполирования
	$\{\varphi_i(x)\}_{i=0}^n$	— система Чебышева
	$\varphi(x)$	— обобщенный интерполяционный многочлен
	$L_n(x)$	— интерполяционный многочлен n -го порядка
	$f(x_i, x_{i+1})$	— разделенная разность первого порядка
	$f(x_i, x_{i+1}, \dots, x_{i+k})$	— разделенная разность k -го порядка
	$S_m(x)$	— интерполяционный сплайн m -го порядка

При интерполяции сплайнами функция $f(x)$ приближается многочленами невысокой степени на частичных отрезках $[x_i, x_{i+1}]$, где $i = 0, 1, \dots, n-1$.

Рассматривается также задача построения обобщенного многочлена $\varphi(x)$, приближающего заданную функцию $f(x)$. В линейном нормированном пространстве коэффициенты обобщенного многочлена $\varphi(x)$ определяются из условия минимальности нормы погрешности интерполирования:

$$\left\| f(x) - \sum_{i=0}^n c_i \varphi_i(x) \right\|. \quad (1.3)$$

Аналогично ставятся задачи интерполяции и приближения многомерных функций.

1.2. Алгоритмы интерполяции и приближения функций

Для одномерных функций задачи интерполяции решаются с использованием алгебраических многочленов Лагранжа и Ньютона, параболических и кубических сплайнов, рассмотрена задача наилучшего приближения в гильбертовом пространстве.

1.2.1. Полиномиальная интерполяция

При аппроксимации полиномами используются функции

$$\varphi_i(x) = x^i, \quad i = 0, 1, \dots, n$$

и интерполяционный многочлен (см. (1.1)) имеет вид

$$\varphi(x) = L_n(x) = \sum_{i=0}^n c_i x^i.$$

Интерполяционный многочлен Лагранжа записывается в виде

$$L_n(x) = \sum_{i=0}^n \frac{\omega(x)}{(x - x_i) \omega'(x_i)} f(x_i), \quad (1.4)$$

где $\omega(x)$ — многочлен степени $n + 1$:

$$\omega(x) = \prod_{i=0}^n (x - x_i),$$

а

$$\omega'(x) \equiv \frac{d\omega}{dx}(x).$$

Можно использовать другую запись интерполяционного многочлена в виде интерполяционного многочлена Ньютона, которая строится с помощью разделенных разностей. Разделенной разностью первого порядка называется отношение

$$f(x_i, x_{i+1}) = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}.$$

Разделенная разность k -го порядка определяется по рекуррентной формуле

$$f(x_i, x_{i+1}, \dots, x_{i+k}) = \frac{f(x_{i+1}, x_{i+2}, \dots, x_{i+k}) - f(x_i, x_{i+1}, \dots, x_{i+k-1})}{x_{i+k} - x_i}.$$

С использованием таких обозначений получим

$$\begin{aligned} [L_n(x)] &= f(x_0) + (x - x_0) f(x_0, x_1) + \\ &+ (x - x_0)(x - x_1) f(x_0, x_1, x_2) + \dots + \\ &+ (x - x_0)(x - x_1) \dots (x - x_{n-1}) f(x_0, x_1, \dots, x_n). \end{aligned} \quad (1.5)$$

1.2.2. Интерполяционные сплайны

Пусть функция $f(x)$ задана в узлах $a = x_0 < x_1 < \dots < x_n = b$. Интерполяционный сплайн $S_m(x)$ порядка m определяется из условий:

1. на каждом отрезке $[x_i, x_{i+1}]$, $i = 0, 1, \dots, n - 1$ $S_m(x)$ является полиномом степени m ;
2. на всем отрезке $[a, b]$ $S_m(x)$ имеет непрерывные производные до порядка $m - 1$;
3. в узлах интерполяции

$$S_m(x_i) = f(x_i), \quad i = 0, 1, \dots, n.$$

При $m \geq 2$ единственность $S_m(x)$ обеспечивается $m - 1$ дополнительными условиями. Обычно эти условия формулируются на концах отрезка интерполяции $[a, b]$.

Интерполяционный кубический сплайн $S_3(x)$ на отрезке $[x_i, x_{i+1}]$ задается полиномом третьей степени:

$$S_3^{(i)} = a_i + b_i(x - x_i) + \frac{c_i}{2}(x - x_i)^2 + \frac{d_i}{6}(x - x_i)^3,$$

$$x_i \leq x \leq x_{i+1}, \quad i = 0, 1, \dots, n - 1, \quad (1.6)$$

причем

$$a_i = S_3^{(i)}(x_i), \quad b_i = \frac{dS_3^{(i)}}{dx}(x_i),$$

$$c_i = \frac{d^2 S_3^{(i)}}{dx^2}(x_i), \quad d_i = \frac{d^3 S_3^{(i)}}{dx^3}(x_i).$$

По определению для $S_3(x)$ выполнены условия:

$$S_3^{(i)}(x_i) = f(x_i), \quad i = 0, 1, \dots, n-1,$$

$$S_3^{(i)}(x_{i+1}) = f(x_{i+1}), \quad i = 0, 1, \dots, n-1,$$

$$\frac{dS_3^{(i)}}{dx}(x_{i+1}) = \frac{dS_3^{(i+1)}}{dx}(x_{i+1}), \quad i = 0, 1, \dots, n-2,$$

$$\frac{d^2 S_3^{(i)}}{dx^2}(x_{i+1}) = \frac{d^2 S_3^{(i+1)}}{dx^2}(x_{i+1}), \quad i = 0, 1, \dots, n-2.$$

Два дополнительных условия можно взять в виде (естественные кубические сплайны)

$$\frac{d^2 S_3^{(0)}}{dx^2}(x_0) = 0, \quad \frac{d^2 S_3^{(n-1)}}{dx^2}(x_n) = 0.$$

1.2.3. Приближение функций в нормированном пространстве

Пусть H — вещественное гильбертово пространство со скалярным произведением (f, g) и нормой $\|f\| = \sqrt{(f, f)}$. В случае $H = L_2(a, b)$ имеем

$$(f, g) = \int_a^b f(x)g(x) dx, \quad \|f\| = \left(\int_a^b |f(x)|^2 dx \right)^{1/2}.$$

В задаче о наилучшем приближении по системе функций

$$\varphi_i(x) \in H, \quad i = 0, 1, \dots, n$$

строится обобщенный многочлен (1.1) (элемент наилучшего приближения), который для заданной приближаемой функции $f(x) \in H$ минимизирует норму отклонения (1.3).

Коэффициенты элемента наилучшего приближения находятся из решения следующей системы линейных уравнений:

$$\sum_{j=0}^n c_j (\varphi_i, \varphi_j) = (f, \varphi_i), \quad i = 0, 1, \dots, n. \quad (1.7)$$

1.3. Упражнения

Рассмотрены примеры решения некоторых проблем теории интерполяции и приближения функций.

Упражнение 1.1. *Покажите однозначную разрешимость задачи интерполяции алгебраическими многочленами.*

Решение. Для определения коэффициентов $c_i, i = 0, 1, \dots, n$ получим систему линейных алгебраических уравнений

$$\sum_{j=0}^n c_j x_i^j = f(x_i), \quad i = 0, 1, \dots, n. \quad (1.8)$$

Определитель этой системы есть определитель Вандермонда:

$$W(x_0, x_1, \dots, x_n) = \begin{vmatrix} 1 & x_0 & \dots & x_0^n \\ 1 & x_1 & \dots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \dots & x_n^n \end{vmatrix} = \prod_{n \geq j > i \geq 0} (x_j - x_i).$$

Он отличен от нуля в рассматриваемом случае несовпадающих узлов интерполяции ($x_0 < x_1 < \dots < x_n$) и тем самым система уравнений (1.8) имеет единственное решение.

Упражнение 1.2. *Рассмотрите различные способы вычисления значений интерполяционного многочлена $L_n(x)$.*

Решение Для вычисления значения полинома

$$L_n(x) = c_0 + c_1 x + \dots + c_n x^n$$

в одной точке требуется $n(n+1)/2$ умножений и n сложений.

При использовании схемы Горнера полином переписывается в виде

$$L_n(x) = c_0 + x \left(c_1 + x \left(c_2 + x \left(\cdots (c_{n-1} + c_n x) \cdots \right) \right) \right).$$

В этом случае требуется только n умножений и n сложений.

Упражнение 1.3. Получите расчетные формулы для коэффициентов естественного кубического сплайна.

Решение Введем обозначения

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

Для кубического сплайна $S_3(x)$ с учетом представления (1.6) получим следующую систему уравнений:

$$a_i = f(x_i), \quad i = 0, 1, \dots, n-1, \quad (1.9)$$

$$a_i + b_i h_{i+1} + \frac{c_i}{2} h_{i+1}^2 + \frac{d_i}{6} h_{i+1}^3 = f(x_{i+1}), \quad i = 0, 1, \dots, n-1, \quad (1.10)$$

$$b_i + c_i h_{i+1} + \frac{d_i}{2} h_{i+1}^2 = b_{i+1}, \quad i = 0, 1, \dots, n-2, \quad (1.11)$$

$$c_i + d_i h_{i+1} = c_{i+1}, \quad i = 0, 1, \dots, n-2, \quad (1.12)$$

$$c_0 = 0, \quad c_{n-1} + d_{n-1} h_n = 0. \quad (1.13)$$

Формально доопределим $c_n = 0$, тогда из (1.12) и второго условия (1.13) получим

$$d_i = \frac{c_{i+1} - c_i}{h_{i+1}}, \quad i = 0, 1, \dots, n-1, \quad (1.14)$$

а вместо (1.13) будем иметь

$$c_0 = 0, \quad c_n = 0. \quad (1.15)$$

Подстановка (1.9), (1.14) в (1.10) дает следующее представление для коэффициентов b_i :

$$b_i = \frac{f(x_{i+1}) - f(x_i)}{h_{i+1}} - \frac{h_{i+1}}{6} (c_{i+1} + 2c_i), \quad i = 1, 2, \dots, n-1. \quad (1.16)$$

С учетом (1.14), (1.16) соотношения (1.11) приводят к уравнению

$$\begin{aligned} c_{i-1}h_i + 2c_i(h_i + h_{i+1}) + c_{i+1}h_{i+1} &= \\ &= 6 \left(\frac{f(x_{i+1}) - f(x_i)}{h_{i+1}} - \frac{f(x_i) - f(x_{i-1}))}{h_i} \right), \end{aligned} \quad (1.17)$$

$$i = 1, 2, \dots, n-1.$$

Тем самым приходим к линейной системе уравнений (1.15), (1.17) с трехдиагональной матрицей с диагональным преобладанием. Решение этой системы всегда существует и единственно.

Другие коэффициенты сплайна определяются в соответствии с (1.9), (1.14), (1.16).

Упражнение 1.4. Рассмотрим на отрезке $[a, b]$ класс функций, имеющих суммируемые с квадратом вторые производные, $W_2^2[a, b]$. Построим интерполирующую функцию

$$u(x) \in W_2^2[a, b], \quad u(x_i) = f(x_i), \quad i = 0, 1, \dots, n, \quad (1.18)$$

которая минимизирует функционал

$$J(u) = \int_a^b \left(\frac{d^2 u}{dx^2} \right)^2 dx. \quad (1.19)$$

Покажите, что такой функцией является естественный кубический сплайн $S_3(x)$ (экстремальное свойство кубического интерполирующего сплайна).

Решение. Для доказательства рассмотрим величину

$$J(u - S_3) = \int_a^b \left(\frac{d^2 u}{dx^2} - \frac{d^2 S_3}{dx^2} \right)^2 dx.$$

Имеет место тождество

$$J(u - S_3) = J(u) - J(S_3) - 2I, \quad (1.20)$$

где

$$I = \int_a^b \left(\frac{d^2 u}{dx^2} - \frac{d^2 S_3}{dx^2} \right) \frac{d^2 S_3}{dx^2} dx = \sum_{i=0}^{n-1} \int_{z_i}^{x_{i+1}} \left(\frac{d^2 u}{dx^2} - \frac{d^2 S_3}{dx^2} \right) \frac{d^2 S_3}{dx^2} dx.$$

Принимая во внимание постоянство $\frac{d^3 S_3}{dx^3}(x)$ при $x \in [x_i, x_{i+1}]$, интегрированием по частям получим

$$\begin{aligned} \int_{x_i}^{x_{i+1}} \left(\frac{d^2 u}{dx^2} - \frac{d^2 S_3}{dx^2} \right) \frac{d^2 S_3}{dx^2} dx &= \\ &= \left(\frac{du}{dx} - \frac{dS_3}{dx} \right) \frac{d^2 S_3}{dx^2} \Bigg|_{x_i}^{x_{i+1}} - \frac{d^3 S_3}{dx^3} (x_i + 0) (u(x) - S_3(x)) \Bigg|_{x_i}^{x_{i+1}}. \end{aligned}$$

Так как $u(x_i) = S_3(x_i)$, $i = 0, 1, \dots, n$, то

$$I = \left(\frac{du}{dx} - \frac{dS_3}{dx} \right) \frac{d^2 S_3}{dx^2} \Bigg|_a^b.$$

Для естественного кубического сплайна $\frac{d^2 S_3}{dx^2}(a) = \frac{d^2 S_3}{dx^2}(b) = 0$ и поэтому $I = 0$ в представлении (1.20), т. е.

$$J(u) = J(S_3) + J(u - S_3).$$

Очевидно, что $J(u - S_3) \geq 0$ и поэтому $J(S_3) \leq J(u)$, т. е. естественный кубический сплайн $S_3(x)$ доставляет минимум функционалу $J(u)$.

Упражнение 1.5. Установите свойство ортогональности погрешности $f - \varphi$ элементу наилучшего приближения

$$(f - \varphi, \varphi) = 0 \tag{1.21}$$

и получите оценку погрешности.

Решение. Элемент наилучшего приближения определяется выражением

$$\varphi(x) = \sum_{i=0}^n c_i \varphi_i(x),$$

где коэффициенты удовлетворяют условиям (см. (1.7))

$$\sum_{j=0}^n c_j (\varphi_i, \varphi_j) = (f, \varphi_i)$$

при $i = 0, 1, \dots, n$. Домножим уравнение на c_i и сложим по всем $i = 0, 1, \dots, n$, что дает

$$\|\varphi\|^2 = (f, \varphi). \quad (1.22)$$

Отсюда непосредственно вытекает доказываемое свойство (1.21).

Из тождества

$$\|f - \varphi\|^2 = \|f\|^2 - 2(f, \varphi) + \|\varphi\|^2$$

с учетом (1.22) следует оценка

$$\|f - \varphi\|^2 = \|f\|^2 - \|\varphi\|^2 \quad (1.23)$$

для погрешности наилучшего приближения.

Упражнение 1.6. Для ортонормированной системы функций $\{\varphi_i(x)\}_{i=0}^n$, т. е. для функций

$$(\varphi_i, \varphi_j) = \begin{cases} 0, & i \neq j, \\ 1, & i = j, \end{cases}$$

рассмотрите задачу среднеквадратичной аппроксимации.

Решение. В этом случае система уравнений (1.7) упрощается и для коэффициентов наилучшего приближения получим

$$c_i = (f, \varphi_i), \quad i = 0, 1, \dots, n. \quad (1.24)$$

В этом случае они называются коэффициентами Фурье разложения функции $f(x)$ по ортонормированной системе $\{\varphi_i(x)\}_{i=0}^n$. Для погрешности имеем представление

$$\|f - \varphi\|^2 = \|f\|^2 - \sum_{i=0}^n c_i^2,$$

которое следует из (1.23).

1.4. Задачи

Задача 1.1. На основе записи интерполяционного многочлена Лагранжа в форме (1.4) получите оценку погрешности интерполирования в виде

$$|f(x) - L_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |\omega(x)|$$

с постоянной

$$M_{n+1} = \sup_{x \in [a, b]} \left| \frac{d^{n+1} f}{dx^{n+1}}(x) \right|.$$

Задача 1.2. Покажите, что интерполяционный полином Лагранжа может быть построен по рекуррентным формулам:

$$L_0(x) = f(x_0),$$

$$L_k(x) = L_{k-1}(x) + (f(x_k) - L_{k-1}(x_k)) \frac{\omega_k(x)}{\omega_k(x_k)},$$

где

$$\omega_1(x) = x - x_0, \quad \omega_{k+1}(x) = \omega_k(x)(x - x_k).$$

Задача 1.3. В представлении интерполяционного полинома Лагранжа (см. (1.4))

$$L_n(x) = \sum_{i=0}^n l_{in}(x) f(x_i)$$

имеет место

$$\sum_{i=0}^n x_i^m l_{in}(x) = x^m, \quad m = 0, 1, \dots, n.$$

Задача 1.4. Погрешность (см. задачу 1) можно уменьшить за счет выбора узлов интерполяции — необходимо выбрать такие $x_i \in [a, b]$, $i = 0, 1, \dots, n$, для которых минимизируется

$$\max_{x \in [a, b]} \left| \prod_{i=0}^n (x - x_i) \right|.$$

Покажите, что оптимальными узлами являются корни приведенного многочлена Чебышева первого рода

$$T_{n+1}(x) = \frac{(b-a)^{n+1}}{2^{2n+1}} \cos \left((n+1) \arccos \frac{2x - (b+a)}{b-a} \right),$$

т. е. точки

$$x_i = \frac{a+b}{2} + \frac{b-a}{2} \cos \left(\frac{(2i+1)\pi}{2(n+1)} \right), \quad i = 0, 1, \dots, n.$$

Задача 1.5. Докажите следующие соотношения

$$f(x_0, x_1, \dots, x_n) = \sum_{i=0}^n \frac{f(x_i)}{\omega'(x_i)},$$

$$f(x_0, x_1, \dots, x_n) = \begin{vmatrix} 1 & x_0 & \dots & x_0^{n-1} & f(x_0) \\ 1 & x_1 & \dots & x_1^{n-1} & f(x_1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n & \dots & x_n^{n-1} & f(x_n) \end{vmatrix} \Bigg| \begin{vmatrix} 1 & x_0 & \dots & x_0^n \\ 1 & x_1 & \dots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \dots & x_n^n \end{vmatrix}.$$

Задача 1.6. Пусть $f(x_0, x_1, \dots, x_n) = 0$ для любых $a = x_0 < x_1 < \dots < x_n = b$. Тогда $f(x)$ на отрезке $[a, b]$ есть алгебраический полином степени не больше n .

Задача 1.7. Пусть $f(x) = g(x)h(x)$, тогда справедлива формула

$$f(x_0, x_1, \dots, x_n) = \sum_{i=0}^n g(x_0, x_1, \dots, x_i) h(x_i, x_{i+1}, \dots, x_n).$$

Задача 1.8. В узлах интерполяции $x_0 < x_1 < \dots < x_n$ заданы значения интерполирующей функции и ее производной: $f(x_i)$, $\frac{df}{dx}(x_i)$,

$i = 0, 1, \dots, n$. Покажите, что полином $Q(x)$ степени $2n + 1$, удовлетворяющий следующим условиям интерполяции (интерполяционный полином Эрмита)

$$Q(x_i) = f(x_i), \quad \frac{dQ}{dx}(x_i) = \frac{df}{dx}(x_i), \quad i = 0, 1, \dots, n,$$

может быть записан в виде

$$Q(x) = \sum_{i=0}^n \left(f(x_i) + (x - x_i) \left(\alpha_i f(x_i) + \beta_i \frac{df}{dx}(x_i) \right) \right) l_{in}^2(x),$$

где

$$\beta_i = 1, \quad \alpha_i = -2 \frac{dl_{in}}{dx}(x_i).$$

Задача 1.9. При интерполировании функций с равноотстоящими узлами $x_i = x_0 + ih$, $h > 0$, $i = 0, 1, \dots, n$ определим конечную разность первого порядка как величину

$$\Delta f(x_i) = f(x_{i+1}) - f(x_i).$$

Конечная разность $(k + 1)$ -го порядка определяется через разность k -го порядка рекуррентно:

$$\Delta^{k+1} f(x_i) = \Delta^k f(x_{i+1}) - \Delta^k f(x_i).$$

Покажите справедливость равенств

$$f(x_n) = f(x_0) + \frac{n}{1!} \Delta f(x_0) + \frac{n(n-1)}{2!} \Delta^2 f(x_0) + \dots + \Delta^n f(x_0),$$

$$f(x_0, x_1, \dots, x_n) = \frac{\Delta^n f(x_0)}{h^n n!}.$$

Задача 1.10. Получите следующую формулу для интерполирования функции в точке x , лежащей вблизи внутреннего узла k , с привлечением

значений интерполируемой функции в узлах $x_k \pm ih, i = 0, 1, \dots, m$:

$$\begin{aligned} [f(x)] &\approx f(x_k) + \frac{\theta}{1!} \Delta f(x_k) + \frac{\theta(\theta-1)}{2!} \Delta^2 f(x_{k-1}) + \\ &+ \frac{\theta(\theta-1)(\theta+1)}{3!} \Delta^3 f(x_{k-2}) + \dots + \\ &+ \frac{1}{(2m-1)!} (\theta-m+1) \cdots \theta \cdots (\theta+m-1) \Delta^{2m-1} f(x_{k-m+1}) + \\ &+ \frac{1}{(2m)!} (\theta+m-1) \cdots \theta \cdots (\theta-m) \Delta^{2m} f(x_{k-m}), \end{aligned}$$

где $\theta = (x - x_k)/h$.

Задача 1.11. Постройте кубический сплайн с дополнительными условиями

- $\frac{d^p S_3}{dx^p}(a+0) = \frac{d^p S_3}{dx^p}(b-0), \quad p = 0, 1, 2$
(периодический кубический сплайн);
- $\frac{dS_3}{dx}(a) = \nu_1, \quad \frac{dS_3}{dx}(b) = \nu_2.$

Задача 1.12. Приведите расчетные формулы для построения периодического параболического сплайна $S_2(x)$.

Задача 1.13. Пусть $f(x) \in C^{(4)}[a, b]$ и

$$\frac{d^2 f}{dx^2}(a) = \frac{d^2 f}{dx^2}(b) = 0.$$

Покажите, что в этих условиях для погрешности интерполирования естественным кубическим сплайном верна оценка

$$\max_{x \in [a, b]} \left| \frac{d^p f}{dx^p}(x) - \frac{d^p S_3}{dx^p}(x) \right| \leq M h^{4-p}, \quad p = 0, 1, 2,$$

где

$$h = \max_{1 \leq i \leq n} |x_i - x_{i-1}|.$$

Задача 1.14. Постройте сглаживающий кубический сплайн из условия минимизации функционала

$$J_\alpha(u) = \int_a^b \left(\frac{d^2 u}{dx^2} \right)^2 dx + \frac{1}{\alpha} \sum_{i=0}^n (u(x_i) - f(x_i))^2$$

с параметром сглаживания $\alpha > 0$. Сформулируйте соответствующую пятидиагональную систему уравнений для определения коэффициентов сплайна и докажите ее однозначную разрешимость.

Задача 1.15. Система функций $\{\varphi_i(x)\}_{i=0}^n$ называется системой Чебышева на отрезке $[a, b]$, если

$$\det \begin{pmatrix} \varphi_0(x_0) & \varphi_1(x_0) & \dots & \varphi_n(x_0) \\ \varphi_0(x_1) & \varphi_1(x_1) & \dots & \varphi_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_0(x_n) & \varphi_1(x_n) & \dots & \varphi_n(x_n) \end{pmatrix} \neq 0$$

при любом расположении узлов $a = x_0 < x_1 < \dots < x_n = b$. Пусть $\alpha_0 < \alpha_1 < \dots < \alpha_n$ — произвольные действительные числа. Покажите, что функции $\{x_{\alpha_i}\}_{i=0}^n$, $\{\exp \alpha_i x\}_{i=0}^n$ образуют чебышевскую систему на произвольном интервале $[a, b]$.

Задача 1.16. Рассмотрите задачу наилучшего приближения для приближения функции, заданной на множестве точек $x_0 < x_1 < \dots < x_N$ с $n < N$ (см. (1.1)) (метод наименьших квадратов).

Глава 2

Численное интегрирование

Задача приближенного интегрирования состоит в вычислении определенного интеграла по значениям подынтегральной функции в отдельных точках. Рассматриваются классические квадратурные формулы прямоугольников, трапеций и Симпсона. Проводится рассмотрение формул интегрирования при заданных узлах квадратурной формулы. В более общем случае проводится оптимизация квадратурных формул за счет выбора узлов.

2.1. Задачи приближенного вычисления интегралов

Рассматривается задача приближенного вычисления определенных интегралов

$$\int_a^b \varrho(x) f(x) dx \quad (2.1)$$

на некотором классе функций $f(x)$ с заданной весовой функцией $\varrho(x)$.

С этой целью подынтегральная функция задается в отдельных точках x_i отрезка $[a, b]$, $i = 0, 1, \dots, n$. Под квадратурной формулой понимается приближенное равенство

$$\int_a^b \varrho(x) f(x) dx \approx \sum_{i=0}^n c_i f(x_i), \quad (2.2)$$

в которой c_i , $i = 0, \dots, n$ — коэффициенты квадратурной формулы. Через

Основные обозначения

- $f(x)$ — подынтегральная функция
 $\varrho(x)$ — весовая функция
 $x_0 < x_1 < \dots < x_n$ — узлы квадратурной формулы
 $c_i, i = 0, 1, \dots, n$ — коэффициенты квадратурной формулы
 $L_n(x)$ — интерполяционный многочлен n -го порядка

$$\psi_n = \int_a^b \varrho(x) f(x) dx - \sum_{i=0}^n c_i f(x_i),$$

определим погрешность квадратурной формулы.

Минимизация погрешности (увеличение точности) квадратурной формулы на выбранном классе функций может достигаться прежде всего за счет выбора коэффициентов квадратурной формулы, а также за счет выбора узлов интегрирования.

2.2. Алгоритмы приближенного вычисления интегралов

Рассматриваются простейшие квадратурные формулы прямоугольников, трапеций и Симпсона для приближенного вычисления определенных интегралов. Строятся квадратурные формулы интерполяционного типа с заданными узлами квадратурной формулы, обсуждаются вопросы оптимизации квадратурные формулы за счет выбора узлов.

2.2.1. Классические квадратурные формулы составного типа

Будем рассматривать задачу вычисления интеграла

$$\int_a^b f(x) dx,$$

т. е. весовая функция $\varrho(x) = 1$. Представим интеграл в виде суммы по частичным отрезкам:

$$\int_a^b f(x) dx = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} f(x) dx.$$

Приведем некоторые простейшие квадратурные формулы для приближенного вычисления интеграла на частичном отрезке $[x_{i-1}, x_i]$.

Формула

$$\int_{x_{i-1}}^{x_i} f(x) dx \approx f(x_{i-1/2})(x_i - x_{i-1}) \quad (2.3)$$

называется формулой прямоугольников на частичном отрезке $[x_{i-1}, x_i]$, где $x_{i-1/2} = (x_{i-1} + x_i)/2$.

Для случая равномерной сетки

$$\omega = \{x \mid x = x_i = a + ih, i = 0, 1, \dots, n, nh = b - a\}$$

суммирование (2.3) по i от 1 до N дает составную формулу прямоугольников

$$\int_a^b f(x) dx \approx \sum_{i=1}^n f(x_{i-1/2})h.$$

При использовании значений интегрируемой функции в узлах простейшей является квадратурная формула трапеций. В этом случае

$$\int_{x_{i-1}}^{x_i} f(x) dx \approx \frac{f(x_{i-1}) + f(x_i)}{2} (x_i - x_{i-1}), \quad (2.4)$$

а для всего отрезка при использовании равномерной сетки получим

$$\int_a^b f(x) dx \approx \sum_{i=1}^n \frac{f(x_{i-1}) + f(x_i)}{2} h.$$

К формуле трапеций мы приходим при замене подинтегральной функции $f(x)$ кусочно-линейной функцией, которая проходит через точки $(x_{i-1}, f(x_{i-1}))$, $(x_i, f(x_i))$. При интерполировании подинтегральной функции на частичном отрезке $[x_{i-1}, x_i]$ с использованием трех точек $(x_{i-1}, f(x_{i-1}))$, $(x_{i-1/2}, f(x_{i-1/2}))$ и $(x_i, f(x_i))$ получим квадратурную формулу Симпсона частичного отрезка:

$$\int_{x_{i-1}}^{x_i} f(x) dx \approx \frac{f(x_{i-1}) + 4f(x_{i-1/2}) + f(x_i)}{6} (x_i - x_{i-1}). \quad (2.5)$$

На всем отрезке соответствующая квадратурная формула составного типа имеет вид

$$\int_a^b f(x) dx \approx \sum_{i=1}^n \frac{f(x_{i-1}) + 4f(x_{i-1/2}) + f(x_i)}{6} h.$$

2.2.2. Квадратурные формулы интерполяционного типа

Приведенные выше квадратурные формулы построены на основе разбиения отрезка интегрирования $[a, b]$ на частичные отрезки $[x_{i-1}, x_i]$, $i = 1, 2, \dots, n$ и использованию некоторых простейших интерполяций для подинтегральной функции на этих отрезках. Можно ориентироваться на применение квадратурных формул интерполяционного типа, когда подинтегральная функция заменяется интерполяционным многочленом сразу на всем отрезке интегрирования $[a, b]$.

Подынтегральная функция $f(x)$ в (2.1) заменяется интерполяционным многочленом Лагранжа

$$L_n(x) = \sum_{i=0}^n \frac{\omega(x)}{(x - x_i)\omega'(x_i)} f(x_i),$$

где

$$\omega(x) = \prod_{i=0}^n (x - x_i), \quad \omega'(x) \equiv \frac{d\omega}{dx}(x).$$

Для коэффициентов квадратурной формулы (2.2) получим представление

$$c_i = \int_a^b \varrho(x) \frac{\omega(x)}{(x - x_i) \omega'(x_i)} dx, \quad i = 0, 1, \dots, n. \quad (2.6)$$

2.2.3. Квадратурные формулы Гаусса

Для повышения точности квадратурной формулы можно оптимизировать выбор не только коэффициентов квадратурной формулы c_i , $i = 0, 1, \dots, n$, но и узлов интерполяции x_i , $i = 0, 1, \dots, n$. Квадратурные формулы интерполяционного типа (2.2), (2.6) являются точными для алгебраических полиномов степени n . За счет выбора узлов интерполирования строятся квадратурные формулы наивысшей алгебраической степени точности (квадратурные формулы Гаусса), которые точны для любого алгебраического многочлена степени $2n + 1$.

Потребуем, чтобы квадратурная формула (2.1) была точна для любого алгебраического многочлена степени m . Это означает, что формула точна для функций $f(x) = x^\alpha$, $\alpha = 0, 1, \dots, m$:

$$\int_a^b \varrho(x) x^\alpha dx = \sum_{i=0}^n c_i x_i^\alpha, \quad i = 0, 1, \dots, m. \quad (2.7)$$

Для определения $2n + 2$ неизвестных c_i , x_i , $i = 0, 1, \dots, n$ имеем нелинейную систему (2.7) $m + 1$ уравнений.

Для знакопостоянной весовой функции $\varrho(x)$ система уравнений (2.7) при $m = 2n + 1$ имеет единственное решение. При этом квадратурная формула является интерполяционной, т. е. коэффициенты вычисляются согласно (2.6), а узлы должны быть такими, чтобы многочлен

$$\omega(x) = \prod_{i=0}^n (x - x_i)$$

был ортогонален с весом $\varrho(x)$ любому многочлену $q(x)$ степени меньше $n + 1$:

$$\int_a^b \varrho(x)\omega(x)q(x) dx = 0.$$

2.3. Упражнения

Ниже приведены упражнения, которые связаны с построением и исследованием точности некоторых квадратурных формул.

Упражнение 2.1. *Получите следующую оценку погрешности составной квадратурной формулы трапеций*

$$\left| \int_a^b f(x) dx - \sum_{i=1}^n \frac{f(x_{i-1}) + f(x_i)}{2} h \right| \leq \tag{2.8}$$

$$\leq \max_{x \in [a, b]} \left| \frac{d^2 f(x)}{dx^2} \right| \frac{h^2 (b-a)}{12}.$$

Решение. В случае квадратурной формулы трапеций (2.2) подынтегральная функция на каждом частичном отрезке заменяется на интерполяционный многочлен первой степени

$$S_1^{(i)} = \frac{1}{h} ((x - x_{i-1})f(x_i) - (x - x_i)f(x_{i-1})).$$

Для погрешности аппроксимации имеем

$$f(x) - S_1^{(i)}(x) = \frac{(x - x_{i-1})(x - x_i)}{2} \frac{d^2 f(\xi)}{dx^2},$$

$$\xi = \xi(x) \in [x_{i-1}, x_i].$$

Тем самым для погрешности квадратурной формулы на частичном интервале получим

$$\begin{aligned}\psi^{(i)} &= \int_{x_{i-1}}^{x_i} f(x) dx - \frac{f(x_{i-1}) + f(x_i)}{2} h = \\ &= \int_{x_{i-1}}^{x_i} \frac{(x - x_{i-1})(x - x_i)}{2} \frac{d^2 f(\xi)}{dx^2} dx\end{aligned}$$

и поэтому

$$|\psi^{(i)}| \leq \max_{x \in [x_{i-1}, x_i]} \left| \frac{d^2 f(x)}{dx^2} \right| \frac{h^3}{12}.$$

Для составной формулы получил искомую оценку (2.8) второго порядка точности квадратурной формулы трапеций.

Упражнение 2.2. *Получите формулу Симпсона на основе комбинирования квадратурных формул прямоугольников.*

Решение. Составим линейную комбинацию из формулы прямоугольников

$$I_0^{(i)} = f(x_{i-1/2})h$$

и формулы трапеций

$$I_1^{(i)} = \frac{f(x_{i-1}) + f(x_i)}{2} h$$

на частичном отрезке для того, чтобы получить квадратурную формулу большей точности. Для формул прямоугольников и трапеций погрешность интегрирования имеет вид

$$\begin{aligned}\psi_0^{(i)} &= \int_{x_{i-1}}^{x_i} f(x) dx - I_0^{(i)} = \frac{h^3}{24} \frac{d^2 f(x_{i-1/2})}{dx^2} + O(h^5), \\ \psi_1^{(i)} &= \int_{x_{i-1}}^{x_i} f(x) dx - I_1^{(i)} = -\frac{h^3}{12} \frac{d^2 f(x_{i-1/2})}{dx^2} + O(h^5),\end{aligned}$$

соответственно. В силу такого представления для погрешности квадратурная формула

$$I_2^{(i)} = \frac{1}{3}(2I_0^{(i)} + I_1^{(i)}) = \frac{f(x_{i-1}) + 4f(x_{i-1/2}) + f(x_i)}{6} h$$

будет иметь точность $O(h^5)$ на частичном отрезке. Это и есть квадратурная формула Симпсона.

Упражнение 2.3. Показать, что для квадратурных формул интерполяционного типа имеет место равенство

$$\sum_{i=0}^n c_i = \int_a^b \varrho(x) dx. \quad (2.9)$$

Решение. Квадратурная формула интерполяционного типа точна для всех многочленов степени n . В частности это имеет место и для $f(x) = 1$, что и приводит к доказываемому равенству (2.9) для суммы коэффициентов квадратурной формулы.

Упражнение 2.4. Получите выражение для погрешности квадратурной формулы интерполяционного типа.

Решение. Пусть $R_n(x)$ — погрешность интерполирования, так что

$$f(x) = L_n(x) + R_n(x).$$

В силу такого представления для погрешности квадратурной формулы получим представление

$$\psi_n = \int_a^b \varrho(x) f(x) dx - \sum_{i=0}^n c_i f(x_i) = \int_a^b \varrho(x) R_n(x) dx.$$

Для погрешности интерполяции имеем

$$R_n(x) = \frac{1}{(n+1)!} \omega(x) \frac{d^{n+1} f}{dx^{n+1}}(\xi(x))$$

и поэтому

$$\psi_n = \frac{1}{(n+1)!} \int_a^b \varrho(x) \omega(x) \frac{d^{n+1} f}{dx^{n+1}}(\xi(x)) dx.$$

В силу этого для погрешности квадратурной формулы следует оценка

$$|\psi_n| \leq \frac{M_{n+1}}{(n+1)!} \int_a^b |\varrho(x) \omega(x)| dx,$$

где постоянная

$$M_{n+1} = \sup_{x \in [a, b]} \left| \frac{d^{n+1} f}{dx^{n+1}}(x) \right|.$$

Упражнение 2.5. Получите квадратурную формулу Гаусса для случая $\rho(x) = 1$ при $n = 1$.

Решение. В данном случае система уравнений (2.7) ($m = 2n + 1 = 3$) имеет вид

$$\begin{aligned} c_0 + c_1 &= b - a, \\ c_0 x_0 + c_1 x_1 &= \frac{1}{2}(b^2 - a^2), \\ c_0 x_0^2 + c_1 x_1^2 &= \frac{1}{3}(b^3 - a^3), \\ c_0 x_0^3 + c_1 x_1^3 &= \frac{1}{4}(b^4 - a^4). \end{aligned}$$

Ее решение есть

$$\begin{aligned} c_0 = c_1 &= \frac{b - a}{2}, \\ x_0 &= \frac{a + b}{2} - \frac{\sqrt{3}}{3} \frac{b - a}{2}, \\ x_1 &= \frac{a + b}{2} + \frac{\sqrt{3}}{3} \frac{b - a}{2}. \end{aligned}$$

Тем самым имеем квадратурную формулу

$$\int_a^b f(x) dx = \frac{b - a}{2} \left(f \left(\frac{a + b}{2} - \frac{\sqrt{3}}{3} \frac{b - a}{2} \right) + f \left(\frac{a + b}{2} + \frac{\sqrt{3}}{3} \frac{b - a}{2} \right) \right).$$

Упражнение 2.6. Покажите, что все коэффициенты квадратурной формулы Гаусса с $\rho(x) > 0$ положительны.

Решение. Рассмотрим многочлены степени $2n$

$$\theta_i(x) = \left(\frac{\omega(x)}{(x - x_i)\omega'(x_i)} \right)^2, \quad i = 0, 1, \dots, n.$$

Для значений этих многочленов в узлах имеем

$$\theta_i(x_j) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Принимая во внимание, что для этих многочленов квадратурная формула Гаусса точна, получим

$$\int_a^b \varrho(x)\theta_i(x) dx = \sum_{j=0}^n c_j \theta_i(x_j) = c_i.$$

С учетом $\theta_i(x) > 0$, $\theta_i(x) \neq 0$ и положительности весовой функции из последнего равенства получим $c_i > 0$, $i = 0, 1, \dots, n$.

2.4. Задачи

Задача 2.1. Показать, что квадратурная формула Симпсона точна не только для многочленов второй степени, но и для многочленов третьей степени.

Задача 2.2. Получите квадратурную формулу

$$\int_{x_{i-1}}^{x_i} f(x) dx \approx \frac{f(x_{i-1}) + 3f(x_{i-2/3}) + 3f(x_{i-1/3}) + f(x_i)}{8} (x_i - x_{i-1})$$

(формула трех восьмых), где $x_{i-k/3} = x_i - (x_i - x_{i-1})k/3$, и исследуйте ее точность.

Задача 2.3. Получите составную квадратурную формулу Симпсона на основе двукратного применения формулы трапеций: один раз с шагом h , другой — с шагом $h/2$.

Задача 2.4. Покажите, что квадратурная формула (формула Эйлера, фор-

мула Эйлера—Маклорена)

$$\int_{x_{i-1}}^{x_i} f(x) dx \approx \frac{f(x_{i-1}) + f(x_i)}{2} (x_i - x_{i-1}) + \frac{(x_i - x_{i-1})^2}{12} (f'(x_i) - f'(x_{i-1}))$$

точна для многочленов третьей степени. Постройте на ее основе квадратурную формулу составного типа.

Задача 2.5. На основе стандартных квадратурных формул составного типа получите квадратурную формулу

$$\int_{-1}^1 \sqrt{1-x^2} f(x) dx \approx \frac{\pi}{n+1} \sum_{i=0}^n \sin^2 \frac{(i+1)\pi}{n+1} f\left(\cos \frac{(i+1)\pi}{n+1}\right)$$

и исследуйте ее погрешность.

Задача 2.6. Постройте составные квадратурные формулы (формулы Филлона) для вычисления интегралов от быстроосциллирующих функций $f(x) = g(x) \exp(ix/\varepsilon)$, где $g(x)$ — мало меняющаяся амплитуда колебаний.

Задача 2.7. На частичном отрезке $[x_i, x_{i-1}]$ подынтегральная функция аппроксимируется кубическим сплайном

$$S_3^{(i)} = a_i + b_i(x - x_i) + \frac{c_i}{2}(x - x_i)^2 + \frac{d_i}{6}(x - x_i)^3.$$

Получите формулу сплайн-квадратуры и исследуйте ее точность.

Задача 2.8. Получите квадратурные формулы интерполяционного типа при $\varrho(x) = 1$, $n = 1, 2, 3$.

Задача 2.9. Докажите, что если квадратурная формула (2.2) точна для любого многочлена степени n , то она является квадратурной формулой интерполяционного типа.

Задача 2.10. Покажите, что квадратурная формула

$$\int_{-\infty}^{\infty} \exp(-x^2) f(x) dx \approx \frac{\sqrt{\pi}}{6} \left(f\left(-\frac{\sqrt{3}}{2}\right) + f(0) + f\left(\frac{\sqrt{3}}{2}\right) \right)$$

точна на многочленах шестой степени.

Задача 2.11. Пусть подынтегральная функция $f(x)$ задана в узлах с погрешностью, т. е.

$$\tilde{f}(x_i) = f(x_i) + \delta_i, \quad i = 0, 1, \dots, n.$$

Получите оценку погрешности для вычисления интеграла (2.1) $\varrho(x) > 0$ при использовании квадратурной формулы Гаусса вида

$$|\tilde{I}_n - I_n| \leq \max_{0 \leq i \leq n} |\delta_i| \int_a^b \varrho(x) dx.$$

Задача 2.12. Рассмотрите проблему построения квадратурных формул с равными коэффициентами (формулы Чебышева):

$$\int_a^b \varrho(x) f(x) dx \approx c(n) \sum_{i=0}^n f(x_i).$$

Для $n = 1, 2, 3, 4$ получите выражения для узлов квадратурной формулы.

Глава 3

Прямые методы линейной алгебры

Одной из основных задач вычислительной математики является проблема решения систем линейных алгебраических уравнений с вещественными коэффициентами. Для нахождения приближенного решения систем уравнений используются прямые и итерационные методы. Математический аппарат линейной алгебры базируется на понятиях нормы вектора и матрицы, числа обусловленности. Рассматриваются классические методы исключения неизвестных, отмечаются особенности решения задач с симметричной вещественной матрицей.

3.1. Задачи решения систем линейных уравнений

Рассматривается задача нахождения решения системы линейных алгебраических уравнений

$$Ax = f. \tag{3.1}$$

Здесь A — квадратная матрица $n \times n$ с вещественными коэффициентами a_{ij} , $i, j = 1, 2, \dots, n$, f — заданный вектор с вещественными компонентами, x — искомый вектор. Будем считать, что определитель матрицы A отличен от нуля и поэтому система уравнений (3.1) имеет единственное решение. Для его нахождения будем использовать прямые (точные) методы, в которых решение находится за конечное число арифметических действий.

Входные данные (коэффициенты матрицы A и компоненты вектора f) заданы с погрешностью, т.е. вместо (3.1) решается система

Основные обозначения

$x = \{x_i\} = \{x_1, x_2, \dots, x_n\}$	— n -мерный вектор
$A = \{a_{ij}\}$	— матрица с элементами a_{ij}
E	— единичная матрица
$D = \text{diag}\{d_1, d_2, \dots, d_n\}$	— диагональная матрица
$\ x\ $	— норма вектора x
$\ A\ $	— норма матрицы A
$\det(A)$	— определитель матрицы A
$\text{cond}(A)$	— число обусловленности матрицы A

уравнений

$$\tilde{A}\tilde{x} = \tilde{f}. \quad (3.2)$$

Необходимо оценить влияние погрешностей в задании коэффициентов и правой части на решение задачи. Близость решения задачи к решению задачи с точными входными данными связывается с числом обусловленности матрицы.

3.2. Алгоритмы решения систем линейных уравнений

Рассматриваются основные понятия линейной алгебры: норма вектора, согласованная норма матрицы. Дается оценка погрешности решения системы линейных уравнений при возмущении правой части и коэффициентов матрицы на основе привлечения понятия числа обусловленности. Среди прямых методов выделяется метод Гаусса с и без выбора главного элемента, который связан с разложением матрицы на произведение треугольных матриц. Для задач с симметричными вещественными матрицами выделяется метод квадратного корня (метод Холецкого).

3.2.1. Обусловленность матрицы и оценки точности решения систем линейных уравнений

Среди норм векторов наиболее употребительны нормы:

$$\|x\|_1 = \sum_{i=1}^n |x_i|,$$

$$\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2},$$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Матричная норма $\|A\|$ подчинена векторной норме $\|x\|$, если

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

Для квадратной невырожденной матрицы A существует единственная матрица A^{-1} , называемая обратной, для которой $AA^{-1} = A^{-1}A = E$. Число обусловленности матрицы A есть

$$\text{cond}(A) = \|A\| \|A^{-1}\|.$$

При рассмотрении близости решений уравнений (3.1), (3.2) для погрешности в задании матрицы, решения и правой части используем обозначения

$$\delta A = \tilde{A} - A, \quad \delta x = \tilde{x} - x, \quad \delta f = \tilde{f} - f.$$

Если матрица A имеет обратную и выполнено условие

$$\|\delta A\| \|A\| < 1,$$

тогда для относительной погрешности справедлива оценка

$$\frac{\|\delta x\|}{\|x\|} \leq \frac{\text{cond}(A)}{1 - \text{cond}(A)\|\delta A\| \|A\|^{-1}} \left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta f\|}{\|f\|} \right). \quad (3.3)$$

Оценка (3.3) выражает устойчивость решения при возмущении правой части и коэффициентов уравнения (3.1) (корректность задачи).

3.2.2. Метод Гаусса для решения систем линейных уравнений

Классический алгоритм исключения неизвестных (метод Гаусса) связывается с использованием представления исходной матрицы A в виде произведения треугольных матриц. Матрица A называется нижней (левой) треугольной матрицей, если ее элементы $a_{ij} = 0$ при $i < j$; для верхней (правой) треугольной матрицы $A - a_{ij} = 0$, если $i > j$.

Если все главные миноры матрицы A отличны от нуля, т. е.

$$a_{11} \neq 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0, \quad \det |A| \neq 0,$$

тогда матрица A представима в виде

$$A = LU, \tag{3.4}$$

где L — нижняя треугольная матрица с единичной диагональю и U — верхняя треугольная матрица с ненулевыми диагональными элементами.

Приведем рекуррентные формулы для определения треугольных матриц L и U :

$$u_{11} = a_{11},$$

$$u_{1j} = a_{1j}, \quad l_{j1} = \frac{a_{j1}}{u_{11}}, \quad j = 2, 3, \dots, n,$$

$$u_{ii} = a_{ii} - \sum_{k=1}^{i-1} l_{ik} u_{ki}, \quad i = 2, 3, \dots, n,$$

$$u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj}, \quad l_{ji} = \frac{1}{u_{ii}} \left(a_{ji} - \sum_{k=1}^{j-1} l_{jk} u_{ki} \right),$$

$$i = 2, 3, \dots, n, \quad j = i + 1, i + 2, \dots, n.$$

Эти формулы определяют компактную схему метода Гаусса для разложения матрицы на множители.

После того как разложение (3.4) проведено, решение системы уравнений сводится к последовательному решению двух систем уравнений с треугольными матрицами:

$$Ly = f, \tag{3.5}$$

$$Ux = y. \tag{3.6}$$

Разложение (3.4) и решение системы (3.5) связывается с прямым ходом в методе исключения неизвестных, а решение системы (3.6) — с обратным ходом.

В методе Гаусса с выбором главного элемента на очередном шаге исключается неизвестное, коэффициент по модулю при котором является наибольшим. В этом случае метод Гаусса применим для любых невырожденных матриц A , т. е. матриц, для которых $\det(A) \neq 0$.

Матрицей перестановок P называется квадратная матрица, у которой в каждой строке и в каждом столбце только один элемент отличен от нуля и равен единице. При $\det(A) \neq 0$ существует матрица перестановок P такая, что справедливо разложение

$$PA = LU.$$

Тем самым метод Гаусса с выбором главного элемента соответствует применению обычного метода Гаусса, который применяется к системе, полученной из исходной системы перестановкой некоторых уравнений.

3.2.3. Метод квадратного корня

При решении системы уравнений (3.1) с симметричной вещественной невырожденной матрицей A используется разложение

$$A = S^*DS,$$

где S — верхняя треугольная матрица с положительными элементами на главной диагонали, S^* — транспонированная к ней ($s_{ij}^* = s_{ji}$), а D — диагональная матрица с элементами d_i , $i = 1, 2, \dots, n$, равными ± 1 . Вычисления на основе этого разложения определяют метод квадратного корня (метод Холецкого).

Для элементов матриц S и D используются расчетные формулы:

$$d_1 = \text{sign } a_{11}, \quad s_{11} = |a_{11}|^{1/2}, \quad s_{1j} = \frac{a_{1j}}{s_{11}}, \quad j = 2, 3, \dots, n,$$

$$d_i = \text{sign} \left(a_{ii} - \sum_{k=1}^{i-1} |s_{ki}|^2 d_k \right),$$

$$s_{ii} = \left| a_{ii} - \sum_{k=1}^{i-1} |s_{ki}|^2 d_k \right|^{1/2},$$

$$s_{ij} = \frac{1}{s_{i1}d_1} \left(a_{ij} - \sum_{k=1}^{i-1} s_{k1}s_{kj}d_k \right),$$

$$i = 2, 3, \dots, n, \quad j = i + 1, i + 2, \dots, n.$$

В методе квадратного корня вычислительные затраты примерно в два раза меньше, чем в стандартном методе Гаусса (эффект учета симметрии матрицы задачи).

3.3. Упражнения

Здесь рассматриваются основные характеристики квадратных вещественных матриц, свойств прямых методов решения систем линейных уравнений, базирующихся на треугольном разложении матрицы системы.

Упражнение 3.1. Показать, что норма матрицы

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|,$$

подчинена норме вектора

$$\|x\|_1 = \sum_{i=1}^n |x_i|.$$

Решение Для произвольного вектора x имеем

$$\begin{aligned} \|Ax\|_1 &= \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij}x_j \right| \leq \sum_{i=1}^n \sum_{j=1}^n |a_{ij}| |x_j| \leq \\ &\leq \sum_{j=1}^n |x_j| \sum_{i=1}^n |a_{ij}| \leq \left(\max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \right) \|x\|_1. \end{aligned}$$

Поэтому достаточно показать, что существует вектор x , для которого в последнем неравенстве достигается равенство. Пусть

$$\max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| = \sum_{i=1}^n |a_{ik}|,$$

тогда можно взять вектор x с компонентами $x_i = 0$, $i \neq k$, $x_k = 1$. Тем самым норма матрицы $\|A\|_1$ подчинена норме вектора $\|x\|_1$.

Упражнение 3.2. *Покажите, что*

$$\text{cond}(A) \geq \frac{|\lambda_{\max}(A)|}{|\lambda_{\min}(A)|}. \quad (3.7)$$

Решение. Для собственного вектора v , отвечающего наибольшему по модулю собственному значению, имеем равенство

$$Av = \lambda_{\max} v.$$

В силу этого

$$\|Av\| = |\lambda_{\max}| \|v\|,$$

что, принимая во внимание $\|Av\| \leq \|A\| \|v\|$, дает неравенство

$$\|A\| \geq |\lambda_{\max}|. \quad (3.8)$$

Для обратной матрицы A^{-1} максимальным по модулю является собственное значение λ_{\min}^{-1} и поэтому

$$\|A^{-1}\| \geq |\lambda_{\min}|^{-1}. \quad (3.9)$$

Из (3.9) и (3.8) следует доказываемое неравенство (3.7).

Упражнение 3.3. *Приведите расчетные формулы метода Гаусса для решения системы уравнений с трехдиагональной матрицей.*

Решение. На основе общих расчетных формул компактной схемы метода Гаусса для матриц L и U получим

$$\begin{aligned} u_{11} &= a_{11}, & u_{12} &= a_{12}, & l_{21} &= \frac{a_{21}}{u_{11}}, \\ u_{ii} &= a_{ii} - l_{i,i-1} u_{i-1,i}, & u_{i,i+1} &= a_{i,i+1}, & l_{i+1,i} &= \frac{a_{i+1,i}}{u_{ii}}, \quad i = 2, 3, \dots, n. \end{aligned}$$

Эти формулы приводят к следующему рекуррентному соотношению для элементов матрицы L :

$$l_{21} = \frac{a_{21}}{a_{11}}, \quad l_{i+1,i} = \frac{a_{i+1,i}}{a_{ii} - l_{i,i-1} a_{i-1,i}}, \quad i = 2, 3, \dots, n$$

Реализация (3.5) в нашем случае дает

$$y_1 = f_1, \quad y_i = f_i - l_{i,i-1}y_{i-1}, \quad i = 2, 3, \dots, n.$$

А из (3.6) находится решение системы уравнений:

$$x_n = \frac{y_n}{a_{nn} - l_{n,n-1}a_{n-1,n}},$$

$$x_i = \frac{l_{i+1,i}}{a_{i+1,i}}(y_i - a_{i+1,i}x_{i+1}), \quad i = n-1, n-2, \dots, 1.$$

Приведенные формулы могут рассматриваться как один из вариантов алгоритма прогонки.

Упражнение 3.4. Постройте алгоритм обращения квадратной матрицы на основе использования метода Гаусса.

Решение. Нахождение матрицы A^{-1} эквивалентно решению матричного уравнения

$$AX = E, \tag{3.10}$$

где X — искомая квадратная матрица. Перепишем уравнение (3.10) в виде системы n^2 уравнений для нахождения элементов x_{ij} , $i, j = 1, 2, \dots, n$ матрицы X :

$$\sum_{k=1}^n a_{ik}x_{kj} = \delta_{ij}, \quad i, j = 1, 2, \dots, n, \tag{3.11}$$

где δ_{ij} — символ Кронекера:

$$\delta_{ij} = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}$$

Система уравнений (3.11) в силу отмеченной специфики правой части распадается на n независимых систем уравнений с одной и той же матрицей A и различными правыми частями. Определим вектора

$$x^{(j)} = \{x_{ij}\}, \quad e^{(j)} = \{\delta_{ij}\}, \quad j = 1, 2, \dots, n$$

и перейдем к n системам уравнений:

$$Ax^{(j)} = e^{(j)}, \quad j = 1, 2, \dots, n.$$

После треугольного разложения (3.4) матрицы A решаются уравнения с треугольными матрицами:

$$Ly^{(j)} = e^{(j)}, \quad Ux^{(j)} = y^{(j)}, \quad j = 1, 2, \dots, n.$$

Упражнение 3.5. Подсчитайте число арифметических действий при решении системы уравнений методом квадратного корня.

Решение. Ограничимся случаем положительно определенной симметричной вещественной матрицы ($(Ax, x) > 0$, $A = A^*$). В этом случае треугольное разложение имеет вид $A = S^*S$, причем

$$s_{11} = (a_{11})^{1/2}, \quad s_{1j} = \frac{a_{1j}}{s_{11}}, \quad j = 2, 3, \dots, n,$$

$$s_{ii} = \left(a_{ii} - \sum_{k=1}^{i-1} (s_{ki})^2 \right)^{1/2},$$

$$s_{ij} = \frac{1}{s_{ii}} \left(a_{ij} - \sum_{k=1}^{i-1} s_{ki} s_{kj} \right),$$

$$i = 2, 3, \dots, n, \quad j = i + 1, i + 2, \dots, n.$$

Вычисления диагональных элементов требуют

$$\sum_{i=2}^n (i-1) = \frac{n(n-1)}{2}$$

умножений. Для каждого фиксированного j для вычисления внедиагональных элементов требуется

$$\sum_{i=2}^{j-1} (i-1) = \frac{(j-1)(j-2)}{2}$$

умножений, а всего

$$\sum_{j=2}^n \frac{(j-1)(j-2)}{2} = \frac{n(n-1)(n-2)}{6}$$

умножений. Число делений совпадает с числом внедиагональных элементов матрицы S и поэтому для реализации треугольного разложения требуется

$$\frac{n(n-1)(n+4)}{6} \approx \frac{n^3}{6}.$$

Для нахождения решения системы уравнений (3.1) после треугольного разложения решаются две системы уравнений

$$S^*y = f, \quad Sx = y,$$

что требует еще $n(n+1)$ операций умножения и деления. Аналогично подсчитывается число сложений.

Упражнение 3.6. Пусть

$$A = S^*S > 0.$$

Выразите число обусловленности матрицы A через число обусловленности матрицы S .

Решение. Для симметричной положительно определенной вещественной матрицы A норма определяется выражением

$$\|A\|_2 = \sup_{x \neq 0} \frac{(x, Ax)}{(x, x)}$$

С учетом этого

$$\begin{aligned} \text{cond}(A) &= \|A\|_2 \|A^{-1}\|_2 = \sup_{x \neq 0} \frac{(x, Ax)}{(x, x)} \sup_{x \neq 0} \frac{(x, A^{-1}x)}{(x, x)} = \\ &= \sup_{x \neq 0} \frac{(Sx, Sx)}{(x, x)} \sup_{x \neq 0} \frac{(S^{-1}x, S^{-1}x)}{(x, x)} = \|S\|_2^2 \|S^{-1}\|_2^2 \end{aligned}$$

и поэтому $\text{cond}(A) = (\text{cond}(S))^2$.

3.4. Задачи

Задача 3.1. Доказать следующие неравенства для норм векторов:

$$\begin{aligned} \|x\|_\infty &\leq \|x\|_1 \leq n \|x\|_\infty, \\ \frac{1}{\sqrt{n}} \|x\|_1 &\leq \|x\|_2 \leq \|x\|_1, \\ \|x\|_\infty &\leq \|x\|_2 \leq \sqrt{n} \|x\|_\infty. \end{aligned}$$

Задача 3.2. Покажите, что норма матрицы

$$\|A\|_2 = \sqrt{\varrho(A^*A)}$$

(через $\varrho(A)$ обозначен спектральный радиус матрицы A — максимальное по модулю собственное значение матрицы) подчинена норме вектора

$$\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}.$$

Задача 3.3. Покажите, что норма матрицы

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$$

является подчиненной норме вектора

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Задача 3.4. Покажите справедливость следующих неравенств для норм матриц:

$$\frac{1}{n} \|A\|_\infty \leq \|A\|_1 \leq n \|A\|_\infty,$$

$$\frac{1}{\sqrt{n}} \|A\|_2 \leq \|A\|_1 \leq \sqrt{n} \|A\|_2,$$

$$\frac{1}{\sqrt{n}} \|A\|_2 \leq \|A\|_\infty \leq \sqrt{n} \|A\|_2.$$

Задача 3.5. M -норма квадратной матрицы A определяется выражением

$$M(A) = n \max_{i \leq i, j \leq n} |a_{ij}|.$$

Покажите справедливость неравенств

$$\frac{1}{n} M(A) \leq \|A\|_\infty \leq M(A),$$

$$\frac{1}{n} M(A) \leq \|A\|_1 \leq M(A),$$

$$\frac{1}{n} M(A) \leq \|A\|_2 \leq M(A).$$

Задача 3.6. Доказать неравенство

$$\|A\|_2^2 \leq \|A\|_1 \|A\|_\infty.$$

Задача 3.7. Покажите, что для собственных значений симметричной матрицы A справедливы оценки

$$\lambda_{\max} \geq \max_{1 \leq i \leq n} a_{ii}, \quad \lambda_{\min} \leq \min_{1 \leq i \leq n} a_{ii}.$$

Задача 3.8. Покажите, что для 1-, 2-, ∞ -норм матрицы число обусловленности не меняется при перестановке строк и столбцов.

Задача 3.9. Показать, что для $n \times n$ матрицы имеет место неравенство

$$\frac{1}{n} \leq \frac{\text{cond}_\infty(A)}{\text{cond}_2(A)} \leq n.$$

Задача 3.10. Вещественная матрица A называется ортогональной, если сопряженная матрица A^* совпадает с обратной A^{-1} . Докажите, что для $\text{cond}(A) = 1$, если A — ортогональная матрица.

Задача 3.11. Пусть A — матрица со строгим диагональным преобладанием по строкам (по столбцам), т. е.

$$|a_{ii}| > \sum_{i \neq j=1}^n |a_{ij}| \quad \left(|a_{ii}| > \sum_{i \neq j=1}^n |a_{ji}| \right), \quad i = 1, 2, \dots, n.$$

Докажите, что в этом случае матрица A невырожденная.

Задача 3.12. Пусть вещественная матрица A симметричная и положительно определена ($A = A^* > 0$). Докажите, что

$$D = \text{diag} \{a_{11}, a_{22}, \dots, a_{nn}\} > 0.$$

Задача 3.13. Докажите, что, если A — симметричная и положительно определенная матрица, то $\max_{ij} |a_{ij}|$ достигается при $i = j$.

Задача 3.14. Пусть A — матрица с элементами

$$a_{ii} > \sum_{i \neq j=1}^n |a_{ij}|, \quad a_{ij} < 0, \text{ если } i \neq j, \quad i = 1, 2, \dots, n.$$

Покажите, что матрица A^{-1} имеет только положительные элементы.

Задача 3.15. Подсчитайте число арифметических действий при решении системы линейных уравнений методом Гаусса.

Задача 3.16. Получите расчетные формулы для определителя симметричной вещественной матрицы на основе использования разложения Холецкого.

Глава 4

Итерационные методы линейной алгебры

Для приближенного решения больших систем линейных алгебраических уравнений используются итерационные методы. Такие системы возникают при приближенном решении многомерных краевых задач математической физики. Рассмотрение начинается с классических итерационных методов Якоби и Зейделя. Приведены базовые понятия теории итерационных методов решения систем линейных уравнений, рассматриваемых в евклидовых пространствах. Обсуждаются проблемы выбора итерационных параметров, выбора матрицы перехода (переобуславливателя).

4.1. Итерационное решение систем линейных уравнений

Рассматриваются проблемы итерационного решения системы линейных уравнений

$$Ax = f \tag{4.1}$$

для нахождения вектора x . В теории итерационных методов матрица A , обычно, рассматривается как линейный оператор, действующий в евклидовом пространстве $H = l_2$, в котором скалярное произведение есть

$$(x, y) = \sum_{i=1}^n x_i y_i, \text{ а норма } \|x\| = (x, x)^{1/2}.$$

Итерационный метод основан на том, что начиная с некоторого начального приближения $x^0 \in H$ последовательно определяются приближенные решения уравнения (4.1) $x^1, x^2, \dots, x^k, \dots$, где k — номер

Основные обозначения

$x = \{x_i\} = \{x_1, x_2, \dots, x_n\}$	— n -мерный вектор
$A = \{a_{ij}\}$	— матрица с элементами a_{ij}
E	— единичная матрица
$D = \text{diag} \{d_1, d_2, \dots, d_n\}$	— диагональная матрица
$\ x\ $	— норма вектора x
$\ A\ $	— норма матрицы A
x^k	— приближенное решение на k -ой итерации
$z^k = x^k - x$	— погрешность приближенного решения
$r^k = Ax^k - f$	— невязка на k -ой итерации
τ, τ_k	— итерационные параметры

итерации. Значения x^{k+1} определяются по ранее найденным x^k, x^{k-1}, \dots . Если при вычислении x^{k+1} используются только значения на предыдущей итерации x^k , то итерационный метод называется одношаговым (двухслойным). Соответственно, при использовании x^k и x^{k-1} итерационный метод называется двухшаговым (трехслойным).

Двухслойный итерационный метод записывается в следующей канонической форме

$$B \frac{x^{k+1} - x^k}{\tau_{k+1}} + Ax^k = f, \quad k = 0, 1, \dots \quad (4.2)$$

Для характеристики точности приближенного решения естественно ввести погрешность $z^k = x^k - x$. Будем рассматривать сходимость итерационного метода в энергетическом пространстве H_R , порожденном симметричной и положительно определенной матрицей R . В H_R скалярное произведение и норма есть

$$(y, w)_R = (Ry, w), \quad \|y\|_R = \sqrt{(y, y)_R}.$$

Итерационный метод сходится в H_R , если $\|z^k\|_R \rightarrow 0$ при $k \rightarrow \infty$. В качестве меры сходимости итераций принимают относительную по-

грешность ε , так что на K -ой итерации

$$\|x^K - x\|_R \leq \varepsilon \|x^0 - x\|_R. \quad (4.3)$$

В силу того, что само точное решение x неизвестно, оценка точности приближенного решения проводится по невязке

$$r^k = Ax^k - f = Ax^k - Ax,$$

которая может быть вычислена непосредственно. Например, итерационный процесс проводится до выполнения оценки

$$\|r^K\| \leq \varepsilon \|r^0\|. \quad (4.4)$$

Использование критерия сходимости (4.4) соответствует выбору $R = A^*A$ в (4.3). Минимальное число итераций, которое гарантирует точность ε (выполнение (4.3) или (4.4)), обозначим $K(\varepsilon)$.

При построении итерационного метода мы должны стремиться к минимизации вычислительной работы по нахождению приближенного решения задачи (4.1) с заданной точностью. Пусть Q_k — число арифметических действий для нахождения приближения x^k и пусть делается $K \geq K(\varepsilon)$ итераций. Тогда общие затраты оцениваются величиной

$$Q(\varepsilon) = \sum_{k=1}^K Q_k.$$

Применительно к двухслойному итерационному методу (4.2) минимизация $Q(\varepsilon)$ может достигаться за счет выбора операторов B_k и итерационных параметров τ_{k+1} . Обычно матрицы B_k (переобуславливатели) задаются из каких-либо соображений близости к матрице A , а оптимизация итерационного метода (4.2) осуществляется за счет выбора итерационных параметров.

4.2. Итерационные алгоритмы решения систем линейных уравнений

Рассматриваются традиционные итерационные методы решения систем линейных уравнений — метод Якоби и метод Зейделя. Приведены основные результаты о скорости сходимости итерационных методов при решении задач с вещественной симметричной положительно определенной

матрицей. Приводится оптимальный выбор постоянных и переменных итерационных параметров. Второй класс итерационных методов связан с определением итерационных параметров на каждом итерационном шаге из минимума функционалов для невязки — итерационные методы вариационного типа.

4.2.1. Классические итерационные методы

В итерационном методе Якоби новое приближение на $(k+1)$ -ой итерации определяется из условий

$$\sum_{j=1}^{i-1} a_{ij}x_j^k + a_{ii}x_i^{k+1} + \sum_{j=i+1}^n a_{ij}x_j^k = f, \quad i = 1, 2, \dots, n. \quad (4.5)$$

Тем самым следующее приближение для отдельной компоненты вектора определяется из соответствующего уравнения системы, когда все другие компоненты берутся с предыдущей итерации.

Метод Зейделя основан на том, что найденное приближение для компонент вектора сразу же задействуются в вычислениях:

$$\sum_{j=1}^{i-1} a_{ij}x_j^{k+1} + a_{ii}x_i^{k+1} + \sum_{j=i+1}^n a_{ij}x_j^k = f, \quad i = 1, 2, \dots, n. \quad (4.6)$$

Для записи итерационных методов (4.5), (4.6) используется следующее разложение матрицы A :

$$A = L + D + U. \quad (4.7)$$

Здесь $D = \text{diag}\{a_{11}, a_{22}, \dots, a_{nn}\}$ — диагональная часть матрицы A , а L и U — нижняя и верхняя треугольные матрицы с нулевыми элементами на главной диагонали, т. е.

$$L = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ a_{21} & 0 & 0 & \dots & 0 \\ a_{31} & a_{32} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & 0 \end{bmatrix},$$

$$U = \begin{bmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & 0 & a_{23} & \dots & a_{2n} \\ 0 & 0 & 0 & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

С учетом (4.7) итерационный метод Якоби (4.5) записывается в каноническом виде (4.2) при

$$B = D, \quad \tau_{k+1} = 1.$$

Для итерационного метода Зейделя (4.6) имеем

$$B = D + L, \quad \tau = 1.$$

Наиболее естественным обобщением рассматриваемых итерационных методов является использование переменных итерационных параметров. В этом случае мы получим

$$D \frac{x^{k+1} - x^k}{\tau_{k+1}} + Ax^k = f, \quad k = 0, 1, \dots, \quad (4.8)$$

$$(D + L) \frac{x^{k+1} - x^k}{\tau_{k+1}} + Ax^k = f, \quad k = 0, 1, \dots \quad (4.9)$$

Отметим также метод верхней релаксации

$$(D + \tau L) \frac{x^{k+1} - x^k}{\tau} + Ax^k = f, \quad k = 0, 1, \dots, \quad (4.10)$$

который можно рассматривать как параметрическое обобщение итерационного метода Зейделя.

Запишем стационарный итерационный метод ($B_k = B$, $\tau_{k+1} = \tau$ в виде

$$x^{k+1} = Sx^k + B^{-1}f, \quad k = 0, 1, \dots, \quad (4.11)$$

где $S = E - \tau B^{-1}A$ — матрица перехода. Необходимым и достаточным условием сходимости итерационного метода (4.11) является условие, чтобы спектральный радиус матрицы перехода S был меньше единицы, т.е. когда все собственные значения матрицы S по модулю меньше единицы.

4.2.2. Двухслойные итерационные методы

Приведем некоторые факты теории итерационных методов при решении задачи (4.1) с симметричной вещественной положительно определенной матрицей A , т. е. когда

$$A = A^* > 0. \quad (4.12)$$

Метод простой итерации (стационарный итерационный метод) соответствует использованию в (4.2) постоянного итерационного параметра $\tau_{k+1} = \tau$, т. е.

$$B \frac{x^{k+1} - x^k}{\tau} + Ax^k = f, \quad k = 0, 1, \dots \quad (4.13)$$

Итерационный метод (4.13) для решения задачи (4.1), (4.12) сходится в H_A , т. е. $\|z\|_A \rightarrow 0$ при $k \rightarrow \infty$, если выполнено неравенство

$$B > \frac{\tau}{2} A. \quad (4.14)$$

Будем считать, что

$$B = B^* > 0 \quad (4.15)$$

и задана априорная информация об операторах B и A в виде двухстороннего операторного неравенства

$$\gamma_1 B \leq A \leq \gamma_2 B, \quad \gamma_1 > 0, \quad (4.16)$$

т. е. операторы B и A энергетически эквивалентны с постоянными энергетической эквивалентности γ_α , $\alpha = 1, 2$. Тогда итерационный метод (4.13) сходится в H_R , $R = A, B$ при $0 < \tau < 2/\gamma_2$. Оптимальным значением итерационного параметра является

$$\tau = \tau_0 = \frac{2}{\gamma_1 + \gamma_2}, \quad (4.17)$$

при котором для числа итераций K , необходимых для достижения точности ϵ , справедлива оценка

$$K \geq K_0(\epsilon) = \frac{\ln \epsilon}{\ln \rho_0}, \quad (4.18)$$

где

$$\varrho_0 = \frac{1 - \xi}{1 + \xi}, \quad \xi = \frac{\gamma_1}{\gamma_2}.$$

Заметим, что в (4.18) $K_0(\varepsilon)$, вообще говоря, нецелое и K — минимальное целое, при котором выполнено $K \geq K_0(\varepsilon)$. Этот результат указывает путь оптимизации сходимости итерационного процесса (4.13) за счет выбора оператора B в соответствии с (4.16), т.е. оператор B должен быть близок оператору A по энергии.

Оптимальный набор итерационных параметров в нестационарном итерационном методе (4.2) для приближенного решения задачи (4.1) при (4.12), (4.15) связан с корнями полиномов Чебышева, поэтому такой итерационный метод называется чебышевским итерационным методом (методом Ричардсона). Определим множество \mathcal{M}_K следующим образом:

$$\mathcal{M}_K = \left\{ -\cos\left(\frac{2i-1}{2K}\pi\right), \quad i = 1, 2, \dots, K \right\}.$$

Для итерационных параметров τ_k используется формула

$$\tau_k = \frac{\tau_0}{1 + \varrho_0 \mu_k}, \quad \mu_k \in \mathcal{M}_K, \quad k = 1, 2, \dots, K. \quad (4.19)$$

Чебышевский итерационный метод (4.2), (4.19) сходится в H_R , $R = A, B$ и для числа итераций K , необходимых для достижения точности ε , справедлива оценка

$$K \geq \underline{K}_0(\varepsilon) = \frac{\ln(2\varepsilon^{-1})}{\ln \varrho_1^{-1}}, \quad (4.20)$$

где

$$\varrho_1 = \frac{1 - \xi^{1/2}}{1 + \xi^{1/2}}, \quad \xi = \frac{\gamma_1}{\gamma_2}.$$

Заметим, что в чебышевском методе (см. (4.19)) расчет итерационных параметров осуществляется по заданному общему числу итераций K . Естественно, что вырожденный случай $K = 1$ соответствует рассмотренному выше методу простой итерации. Практическая реализация чебышевского итерационного метода связана с проблемой вычислительной устойчивости, которая решается специальным упорядочиванием итерационных параметров (выбором μ_k из множества \mathcal{M}_K).

4.2.3. Итерационные методы вариационного типа

Выше рассматривались итерационные методы решения задачи в условиях, когда задана априорная информация об операторах B и A в виде констант (см. (4.16)) энергетической эквивалентности γ_1 и γ_2 . Через эти постоянные определяются оптимальные значения итерационных параметров (см. (4.17), (4.19)). В итерационных методах вариационного типа, в которых итерационные параметры вычисляются без такой априорной информации.

Обозначая невязку $r^k = Ax^k - f$ и поправку $w^k = B^{-1}r^k$, для итерационных параметров при естественном предположении о минимизации погрешности в H_R получим формулу

$$\tau_{k+1} = \frac{(Rw^k, z^k)}{(Rw^k, w^k)}. \quad (4.21)$$

Итерационный процесс (4.2) запишется следующим образом

$$x^{k+1} = x^k - \tau_{k+1}w^k, \quad k = 0, 1, \dots$$

Конкретизация итерационного метода достигается за счет выбора оператора $R = R^* > 0$. Этот выбор должен быть подчинен, в частности, условию возможности вычисления итерационных параметров. В формулу (4.21) входит невычисляемая величина z^k и поэтому простейший выбор $R = B$ здесь не проходит. Вторая отмеченная выше возможность $R = A$ приводит нас к итерационному методу скорейшего спуска, когда

$$\tau_{k+1} = \frac{(w^k, r^k)}{(Aw^k, w^k)}. \quad (4.22)$$

Среди других возможностей выбора R отметим случай $R = AB^{-1}A$ — метод минимальных поправок, когда

$$\tau_{k+1} = \frac{(Aw^k, w^k)}{(B^{-1}Aw^k, Aw^k)}. \quad (4.23)$$

Двухслойный итерационный метод вариационного типа сходится не медленнее метода простой итерации, т.е. для числа итераций n , необходимых для достижения точности ϵ , справедлива оценка (4.18).

В вычислительной практике наибольшее распространение получили трехслойные итерационные методы вариационного типа. По скорости сходимости они не хуже итерационного метода с чебышевским набором итерационных параметров.

В трехслойном (двухшаговом) итерационном методе новое приближение находится по двум предыдущим. Для реализации метода требуются два начальных приближения x^0, x^1 . Обычно x^0 задается произвольно, а x^1 находится по двухслойному итерационному методу. Трехслойный метод записывается в следующей канонической форме трехслойного итерационного метода:

$$\begin{aligned} By^{k+1} &= \alpha_{k+1}(B - \tau_{k+1}A)y^k + (1 - \alpha_{k+1})By^{k-1} + \alpha_{k+1}\tau_{k+1}\varphi, \\ k &= 1, 2, \dots, \\ By^1 &= (B - \tau_1A)y^0 + \tau_1\varphi, \end{aligned} \quad (4.24)$$

где α_{k+1} и τ_{k+1} — итерационные параметры.

В методе сопряженных градиентов итерационные параметры рассчитываются по формулам

$$\begin{aligned} \tau_{k+1} &= \frac{(w^k, r^k)}{(Aw^k, w^k)}, \quad k = 0, 1, \dots, \\ \alpha_{k+1} &= \left(1 - \frac{\tau_{k+1}}{\tau_k} \frac{(w^k, r^k)}{(w^{k-1}, r^{k-1})} \frac{1}{\alpha_k}\right)^{-1}, \\ k &= 1, 2, \dots, \quad \alpha_1 = 1. \end{aligned}$$

Этот метод наиболее широко используется в вычислительной практике при решении задач с симметричной положительно определенной матрицей.

4.3. Упражнения

Приведены упражнения, которые иллюстрируют теоретические результаты по итерационному решению систем линейных алгебраических уравнений.

Упражнение 4.1. Пусть в матрице A в уравнения (4.1) для элементов имеет место неравенство

$$q|a_{ii}| > \sum_{i \neq j=1}^n |a_{ij}|, \quad i = 1, 2, \dots, n, \quad (4.25)$$

где $0 < q < 1$. Тогда итерационный метод Зейделя сходится и для погрешности справедлива оценка

$$\|z^k\|_{\infty} \leq q^k \|z^0\|_{\infty}, \quad z^k = x^k - x, \quad \|x\|_{\infty} = \max_{1 \leq i \leq n} |x_i|.$$

Решение. Из (4.6) для погрешности на новой итерации имеем оценку

$$|z_i^{k+1}| \leq \|z^{k+1}\|_{\infty} \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| + \|z^k\|_{\infty} \sum_{j=i+1}^n \left| \frac{a_{ij}}{a_{ii}} \right|. \quad (4.26)$$

Принимая во внимание (4.25), получим

$$\sum_{j=i+1}^n \left| \frac{a_{ij}}{a_{ii}} \right| \leq q - \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| \leq q \left(1 - \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| \right).$$

С учетом этого (4.26) дает

$$|z_i^{k+1}| \leq \|z^{k+1}\|_{\infty} \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| + q \|z^k\|_{\infty} \left(1 - \sum_{j=1}^{i-1} \left| \frac{a_{ij}}{a_{ii}} \right| \right). \quad (4.27)$$

Пусть максимум $|z_i^{k+1}|$ достигается при $i = m$, тогда (4.27) приводит к неравенству

$$\|z^{k+1}\|_{\infty} \left(1 - \sum_{j=1}^{m-1} \left| \frac{a_{mj}}{a_{mm}} \right| \right) \leq q \|z^k\|_{\infty} \left(1 - \sum_{j=1}^{m-1} \left| \frac{a_{mj}}{a_{mm}} \right| \right).$$

Отсюда и следует доказываемая оценка для погрешности.

Упражнение 4.2. Покажите, что метод верхней релаксации (4.10) при решении задачи (4.1) с симметричной положительно определенной матрицей A сходится при $0 < \tau < 2$.

Решение. Достаточно проверить выполнение неравенства (4.14). С учетом симметрии ($U = L^*$) имеем

$$(Ax, x) = (Dx, x) + 2(Lx, x)$$

и поэтому

$$(Bx, x) - \frac{\tau}{2}(Ax, x) = \left(1 - \frac{\tau}{2}\right)(Dx, x).$$

В силу положительной определенности матрицы A имеем $D > 0$ и поэтому неравенство (4.14) при отмеченных ограничениях на итерационный параметр выполнено.

Упражнение 4.3. Пусть $A = A^* > 0$, $B = B^* > 0$ и $C = A^{1/2}B^{-1}A^{1/2}$. Тогда неравенства

$$\gamma_1 B \leq A \leq \gamma_2 B, \quad \gamma_1 E \leq C \leq \gamma_2 E \quad (4.28)$$

эквивалентны.

Решение. Положим $y = C^{1/2}x$, $v = A^{-1/2}y$ и при постоянной γ рассмотрим выражение

$$\begin{aligned} (Cx, x) - \gamma(x, x) &= (y, y) - \gamma(C^{-1}y, y) = \\ &= (y, y) - \gamma(A^{-1/2}AA^{-1/2}y, y) = (Av, v) - \gamma(Bv, v). \end{aligned}$$

Следовательно матрицы $(C - \gamma E)$ и $(A - \gamma B)$ имеют одинаковые знаки. Полагая $\gamma = \gamma_1$ и $\gamma = \gamma_2$, получим эквивалентность двухсторонних операторных неравенств (4.28).

Упражнение 4.4. Пусть в итерационном методе (4.13) A — симметричная и положительно определенная матрица и выполнено неравенство

$$B - \frac{\tau}{2}A \geq \frac{1 - \varrho^2}{2\tau}B^*A^{-1}B \quad (4.29)$$

с постоянной $\varrho \in (0, 1)$. Тогда итерационный метод сходится и для погрешности справедлива оценка

$$\|z^k\|_A \leq \varrho^k \|z^0\|_A. \quad (4.30)$$

Решение. Неравенство (4.30) эквивалентно выполнению матричного неравенства

$$\varrho^2 A \geq S^*AS, \quad S = E - \tau B^{-1}A,$$

т. е.

$$\tau A((B^*)^{-1} + B^{-1})A \geq (1 - \varrho^2)A + \tau^2 A(B^*)^{-1}AB^{-1}A.$$

Это матричное неравенство останется в силе после умножения его справа на матрицу $G = A^{-1}B$, а слева на $G^* = B^*A^{-1}$:

$$\tau(B + B^*) \geq (1 - \varrho^2)B^*A^{-1}B + \tau^2A.$$

Последнее неравенство совпадает с (4.29). При $\varrho \in (0, 1)$ неравенство (4.30) обеспечивает сходимость итерационного метода (4.13).

Упражнение 4.5. Пусть

$$A = A_1 + A_2 = A^* > 0, \quad A_2^* = A_1 = \frac{1}{2}D + L.$$

В попеременно-треугольном методе переобуславливатель B задается в виде

$$B = (G + \omega A_1)G^{-1}(G + \omega A_2), \quad (4.31)$$

где $G = G^* > 0$. При априорной информации

$$A \geq \delta G, \quad \delta > 0, \quad A_1 G^{-1} A_2 \leq \frac{\Delta}{4} A \quad (4.32)$$

укажите оптимальный выбор параметра ω .

Решение. Прежде всего покажем, что матрица B — положительно определена и симметрична. В самом деле

$$\begin{aligned} (Bx, y) &= ((G + \omega A_1)G^{-1}(G + \omega A_2)x, y) = \\ &= ((G + \omega A_2)x, G^{-1}(G + \omega A_1)y) = \\ &= (x, (G + \omega A_1)G^{-1}(G + \omega A_2)y), \\ (Bx, x) &= ((G + \omega A_2)x, G^{-1}(G + \omega A_1)x) = \|(G + \omega A_2)x\|_{G^{-1}}^2, \end{aligned}$$

т. е. $B = B^* > 0$.

Скорость сходимости итерационного метода (4.13) в условиях (4.12), (4.15) определяется постоянными γ_1, γ_2 в (4.16). В попеременно-треугольном методе (4.31) имеем

$$B = G + \omega(A_1 + A_2) + \omega^2 A_1 G^{-1} A_2 = (G - \omega A_1)G^{-1}(G - \omega A_2) + 2\omega A$$

и поэтому

$$B \geq 2\omega A, \quad \gamma_2 = \frac{1}{2\omega}.$$

Для оценки матрицы B сверху привлекается априорная информация (4.32):

$$B = G + \omega A + \omega^2 A_1 G^{-1} A_2 \leq \frac{1}{\delta} \left(1 + \omega\delta + \frac{\omega\delta\Delta}{4} \right) A.$$

Тем самым

$$\gamma_1 = \delta G \left(1 + \omega\delta + \frac{\omega\delta\Delta}{4} \right)^{-1}.$$

Скорость сходимости будет наибольшей (см. (4.18), (4.20)) при максимальном $\xi(\omega) = \gamma_1/\gamma_2$. Максимум $\xi(\omega)$ достигается при

$$\omega = \omega_0 = \frac{2}{\sqrt{\Delta\delta}}.$$

Упражнение 4.6. Получите выражения для итерационного метода (4.2) с $B = B^* > 0$ для решения задачи (4.1) с матрицей $A > 0$ из условия минимума поправки в H_B .

Решение. Для погрешности итерационного метода имеем однородное уравнение

$$B \frac{z^{k+1} - z^k}{\tau_{k+1}} + Az^k = 0, \quad k = 0, 1, \dots$$

Аналогично записывается и уравнение для поправки

$$B \frac{w^{k+1} - w^k}{\tau_{k+1}} + Aw^k = 0, \quad k = 0, 1, \dots$$

Отсюда для нормы поправки на новой итерации получим

$$\begin{aligned} \|Bw^{k+1}\|_{B^{-1}}^2 &= \|w^{k+1}\|_B^2 = (B^{-1}(Bw^k - \tau_{k+1}Aw^k), (Bw^k - \tau_{k+1}Aw^k)) = \\ &= (Bw^k, w^k) - 2\tau_{k+1}(Aw^k, w^k) + \tau^2(B^{-1}Aw^k, w^k). \end{aligned}$$

Дифференцируя это выражение по τ_{k+1} , находим для определения итерационных параметров в методе минимальных поправок формулу (4.23).

4.4. Задачи

Задача 4.1. Докажите сходимость метода Якоби при решении задачи (4.1) с матрицей A , для элементов которой выполнены одно из следующих условий

$$\sum_{j \neq i=1}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1, \quad i = 1, 2, \dots, n,$$

$$\sum_{j \neq i=1}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1, \quad j = 1, 2, \dots, n,$$

$$\sum_{j=1}^n \sum_{j \neq i=1}^n \left(\frac{a_{ij}}{a_{ii}} \right)^2 < 1.$$

Задача 4.2. Исследуйте сходимость итерационного метода Зейделя, когда

$$q|a_{jj}| > \sum_{j \neq i=1}^n |a_{ij}| \quad j = 1, 2, \dots, n,$$

где $0 < q < 1$ (диагональное преобладание по столбцам).

Задача 4.3. Установите следующие свойства положительно определенных матриц:

- если $A > 0$, то матрица A — невырожденная и $A^{-1} > 0$;
- если $A, B > 0$, то для любых неотрицательных чисел α, β , не равных нулю одновременно, имеем $\alpha A + \beta B > 0$;
- для симметричной вещественной (эрмитовой) положительно определенной матрицы A существует единственная эрмитова положительно определенная матрица S такая, что $S^2 = A$. Матрица S называется квадратным корнем из матрицы A и обозначается $A^{1/2}$.

Задача 4.4. Покажите эквивалентность матричных неравенств:

$$A \geq 0, \quad B^* A B \geq 0,$$

если $A = A^*$ и B — невырожденная матрица и

$$\alpha A \geq \beta B, \quad \alpha B^{-1} \geq \beta A^{-1},$$

если $A = A^* > 0$, $B = B^* > 0$, α и β — любые действительные числа.

Задача 4.5. Покажите, что итерационный метод (4.13) при $B = E$ для задачи (4.1) с $A > 0$ сходится при всех τ , удовлетворяющих неравенству $\tau < 2/\|A\|$.

Задача 4.6. Пусть $S = E - \tau C$ и выполнены условия

$$C = C^*, \quad \gamma_1 E \leq C \leq \gamma_2 E, \quad \gamma_1 > 0.$$

Тогда $\|S\| < 1$ при $0 < \tau < 2/\gamma_2$ и нижняя грань нормы оператора достигается при

$$\tau = \tau_0 = \frac{2}{\gamma_1 + \gamma_2},$$

причем

$$\inf_{\tau} \|S\| = \|E - \tau_0 C\| = \frac{1 - \xi}{1 + \xi}, \quad \xi = \frac{\gamma_1}{\gamma_2}.$$

Задача 4.7. Пусть A и B — симметричные положительно определенные матрицы. Тогда неравенства

$$\frac{1 - \varrho}{\tau} B \leq A \leq \frac{1 + \varrho}{\tau} B$$

с $\varrho > 0$ необходимы и достаточны для того, чтобы для задачи

$$B \frac{z^{k+1} - z^k}{\tau} + Az^k = 0, \quad k = 0, 1, \dots$$

выполнялась оценка

$$\|z^{k+1}\|_A \leq \varrho \|z^k\|_A, \quad k = 0, 1, \dots$$

Задача 4.8. Получите оценку числа итераций попеременно-треугольного итерационного метода (4.31), (4.32) при выборе чебышевского набора итерационных параметров вида

$$K \geq K_0(\varepsilon) = \frac{1}{2\sqrt{2}\sqrt[3]{\eta}} \ln\left(\frac{2}{\varepsilon}\right), \quad \eta = \frac{\delta}{\Delta}.$$

Задача 4.9. Определите область значений итерационного параметра в методе минимальных невязок ($B = E$) при решении задачи (4.1) с матрицей $A = \alpha E + K$, где K — кососимметричная ($K = -K^*$) матрица, а $\alpha > 0$.

Задача 4.10. Докажите, что в итерационном методе сопряженных градиентов выполнены следующие свойства ортогональности для погрешностей на различных итерациях

$$(Gs^j, s^i) = 0, \quad j = 0, 1, \dots, i-1, \quad i = 1, 2, \dots,$$

где

$$G = A^{1/2} B^{-1} A^{1/2}, \quad s^k = A^{1/2} z^k.$$

Задача 4.11. При реализации трехслойного итерационного процесса (4.24) используется следующее представление для нового приближения

$$y^{k+1} = \alpha_{k+1} y^k + (1 - \alpha_{k+1}) y^{k-1} - \alpha_{k+1} \tau_{k+1} w^k,$$

где $w^k = B^{-1} r^k$. Рассмотрите возможность использования расчетных формул

$$y^{k+1} = y^k + \lambda_k p^k, \quad k = 0, 1, \dots,$$

$$p^k = w^k + \mu_k p^{k-1}, \quad k = 1, 2, \dots, \quad p^0 = w^0.$$

Задача 4.12. Пусть

$$S(\omega) = (E + \omega A)^{-1} (E - \omega A), \quad A = A^* > 0.$$

Докажите, что

$$\inf_{\omega} \|S(\omega)\| = \|S(\omega_0)\| = \frac{1 - \sqrt{\eta}}{1 + \sqrt{\eta}},$$

где

$$\omega_0 = \frac{1}{\sqrt{\delta \Delta}}, \quad \eta = \frac{\delta}{\Delta}.$$

Задача 4.13. Пусть

$$A = A_1 + A_2, \quad A_\alpha = A_\alpha^*,$$

$$\delta_\alpha E \leq A_\alpha \leq \Delta_\alpha E, \quad \delta_\alpha > 0, \quad \alpha = 1, 2.$$

Рассмотрите условия сходимости итерационного метода (итерационный метод переменных направлений)

$$\begin{aligned} (E + \tau A_1) \frac{x^{k+1/2} - x^k}{\tau} + Ax^k &= f, \\ (E + \tau A_2) \frac{x^{k+1} - x^{k+1/2}}{\tau} + Ax^{k+1/2} &= f \end{aligned} \tag{4.33}$$

и выберите оптимальный итерационный параметр τ .

Задача 4.14. Покажите, что при $A > 0$, $B = B^* > 0$ и выполнении неравенства

$$(B^{-1}Ax, Ax) \leq \gamma_2(Ax, x)$$

итерационный метод (4.13) сходится при $\tau \leq 2/\gamma_2$.

Задача 4.15. В условиях предыдущей задачи и при дополнительном условии

$$A \geq \gamma_1 B, \quad \gamma_1 > 0$$

выберите оптимальное значение итерационного параметра τ .

Задача 4.16. Рассмотрите итерационный метод (4.33) для решения задачи (4.1) при

$$A = A_1 + A_2,$$

$$A_\alpha \geq \delta_\alpha E, \quad \|A_\alpha x\|^2 \leq \Delta_\alpha (A^\alpha x, x), \quad \delta_\alpha > 0, \quad \alpha = 1, 2.$$

Глава 5

Спектральные задачи линейной алгебры

Важной и трудной задачей линейной алгебры является нахождение собственных значений и собственных векторов матриц. Рассматриваются проблема устойчивости собственных значений по отношению к малым возмущениям элементов матрицы. Для приближенного нахождения отдельных собственных значений широко используется степенной метод в различных модификациях. Для решения полной проблемы для симметричных матриц применяется итерационный метод Якоби и QR -алгоритм.

5.1. Собственные значения и собственные вектора матриц

Рассматриваются проблемы нахождения собственных значений и собственных векторов квадратной вещественной матрицы A . Собственным числом называется число λ такое, что для некоторого ненулевого вектора (собственного вектора) φ имеет место равенство

$$A\varphi = \lambda\varphi. \quad (5.1)$$

Собственные вектора определены с точностью до числового множителя. Множество всех собственных значений матрицы A называется спектром матрицы A .

С учетом того, что ищется нетривиальное решение уравнения (5.1), то

$$\det(A - \lambda E) = 0. \quad (5.2)$$

Основные обозначения

$x = \{x_i\} = \{x_1, x_2, \dots, x_n\}$ — n -мерный вектор

$A = \{a_{ij}\}$ — матрица с вещественными
элементами a_{ij}

E — единичная матрица

$D = \text{diag}\{d_1, d_2, \dots, d_n\}$ — диагональная матрица

$\lambda_i, i = 1, 2, \dots, n$ — собственные значения

$\varphi^i, i = 1, 2, \dots, n$ — собственные вектора

$(x, y) = \sum_{i=1}^n x_i y_i$ — скалярное произведение

Тем самым собственные значения λ матрицы A являются корнями характеристического многочлена n -ой степени (5.2). Задача отыскания собственных значений и собственных векторов матрицы сводится к построению характеристического многочлена, отысканию его корней и последующему нахождению нетривиальных решений уравнения (5.1) для найденных собственных значений.

В вычислительной практике рассматривается как полная проблема собственных значений, когда необходимо найти все собственные значения матрицы A , так и частичная проблема собственных значений, когда ищутся только некоторые собственные значения, например, максимальные (минимальные) по модулю.

5.2. Численные методы решения задач на собственные значения

Начнем с приведения некоторых полезных фактов о свойствах собственных значений и собственных векторов квадратной матрицы. Далее рассматриваются методы решения частичной и полной проблемы собственных значений.

5.2.1. Свойства собственных значений и собственных векторов

Квадратная вещественная матрица порядка n имеет n собственных значений, при этом каждое собственное значение считается столько раз, какова его кратность как корня характеристического многочлена. Для симметричной матрицы A собственные значения вещественны, а собственные вектора, соответствующие различным собственным значениям, ортогональны, т. е. $(\varphi^i, \varphi^j) = 0$, если $i \neq j$.

Отметим также некоторые свойства собственных значений и собственных векторов для сопряженной матрицы A^* :

$$A^* \psi = \mu \psi. \quad (5.3)$$

Для спектральных задач (5.1), (5.3) имеем

$$\begin{aligned} \lambda_i &= \mu_i, \quad i = 1, 2, \dots, n, \\ (\varphi^i, \psi^j) &= 0, \quad i \neq j. \end{aligned}$$

Две квадратные матрицы A и B одинаковых размеров называются подобными, если существует такая невырожденная матрица P , что $A = P^{-1}BP$. Подобные матрицы имеют одни и те же собственные значения, так как из (5.1) непосредственно следует

$$B\phi = \lambda\phi, \quad \phi = P\varphi.$$

Поэтому вычислительные алгоритмы решения спектральных задач базируются на подобном преобразовании матрицы к матрице B , для которой спектральная задача решается проще. В качестве B естественно выбирать диагональную матрицу, причем в данном случае это будет

$$\Lambda = \text{diag} \{ \lambda_1, \lambda_2, \dots, \lambda_n \}.$$

Упорядочим собственные значения симметричной матрицы A по возрастанию, т. е. $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Свойства собственных значений и собственных функций связаны с отношением Релея $\frac{(Ax, x)}{(x, x)}$. Отметим, например, что для любого ненулевого вектора x справедливы неравенства

$$\lambda_1 \leq \frac{(Ax, x)}{(x, x)} \leq \lambda_n.$$

Важны также экстремальные свойства Релея

$$\lambda_1 = \min_{x \neq 0} \frac{(Ax, x)}{(x, x)}, \quad \lambda_n = \max_{x \neq 0} \frac{(Ax, x)}{(x, x)}.$$

Для локализации собственных значений произвольной матрицы A привлекаются круги Гершгорина: любое собственное значение матрицы A лежит по крайней мере в одном из кругов

$$|\lambda - a_{ii}| \leq \sum_{i \neq j=1}^n |a_{ij}|, \quad i = 1, 2, \dots, n.$$

Приведем теперь некоторые сведения о возмущении собственных значений при возмущении элементов матрицы. Помимо (5.1) рассмотрим задачу

$$\tilde{A}\tilde{\varphi} = \tilde{\lambda}\tilde{\varphi}. \quad (5.4)$$

Ограничимся случаем, когда все собственные значения простые. С точностью до членов второго порядка возмущение собственных значений за счет возмущения матрицы дается оценкой

$$|\tilde{\lambda} - \lambda| \leq c_i \|\tilde{A} - A\|, \quad (5.5)$$

где $\|x\| = \sqrt{(x, x)}$. Мерой устойчивости собственного значения λ , служит величина

$$c_i = \frac{\|\varphi^i\| \|\psi^i\|}{|(\varphi_i, \psi_i)|}, \quad (5.6)$$

которая называется коэффициентом перекоса матрицы A , соответствующим данному собственному значению. Здесь ψ_i — собственное значение матрицы A^* . Для нормированных собственных векторов задач (5.1) и (5.5) соответствующая оценка устойчивости имеет вид

$$\|\tilde{\varphi}^i - \varphi^i\| \leq \sum_{i \neq j=1}^n \frac{c_j}{|\lambda_i - \lambda_j|} \|\tilde{A} - A\|.$$

В частности, для симметричной матрицы все коэффициенты перекоса равны единице и оценки устойчивости вычисления собственных значений оптимальны.

5.2.2. Итерационные методы решения частичной проблемы собственных значений

Для нахождения минимального по модулю (максимального) собственного значения используются прямые и обратные итерации. Пусть матрица A является симметричной, все ее собственные значения простые и упорядочены следующим образом

$$|\lambda_1| < |\lambda_2| < \dots < |\lambda_n|.$$

Определим последовательность векторов

$$x^{k+1} = Ax^k, \quad k = 0, 1, \dots \quad (5.7)$$

при некотором заданном x^0 (прямые итерации). Рассматривая последовательности скалярных произведений (x^k, x^k) , (x^{k+1}, x^k) , при ограничении $(x^0, \varphi^n) \neq 0$ получим

$$\frac{(x^{k+1}, x^k)}{(x^k, x^k)} = \lambda_n + O\left(\left|\frac{\lambda_{n-1}}{\lambda_n}\right|^{2k}\right). \quad (5.8)$$

Тем самым при использовании итерационного процесса (5.7) находится максимальное по модулю собственное значение матрицы A .

Принимая во внимание, что собственные значения матрицы A^{-1} есть $1/\lambda_i$, $i = 1, 2, \dots, n$, при использовании последовательности (обратные итерации)

$$y^{k+1} = A^{-1}y^k, \quad (y^0, \varphi^1) \neq 0, \quad k = 0, 1, \dots \quad (5.9)$$

имеем

$$\lim_{k \rightarrow \infty} \frac{(y^{k+1}, y^k)}{(y^k, y^k)} = \frac{1}{\lambda_1}.$$

Тем самым при обратных итерациях находится минимальное по модулю собственное значение матрицы.

Заметим, что в прямых и обратных итерациях нет необходимости в специальном вычислении соответствующих собственных векторов, так как

$$\lim_{k \rightarrow \infty} x^k = \varphi^n, \quad \lim_{k \rightarrow \infty} y^k = \varphi^1. \quad (5.10)$$

Вычислительная реализация обратных итераций (5.9) может быть основана на однократном LU разложении матрицы A . После этого каждая обратная итерация по вычислительным затратам эквивалентна прямой итерации.

Отметим процедуру ускорения сходимости обратных итераций при известном хорошем приближении собственного значения $\tilde{\lambda}_1$ к собственному значению λ_1 . В этом случае рассматриваются обратные итерации со сдвигом, когда

$$z^{k+1} = (A - \tilde{\lambda}_1 E)^{-1} z^k, \quad (z^0, \varphi^1) \neq 0, \quad k = 0, 1, \dots$$

Скорость сходимости обратных итераций со сдвигом и без сдвига определяется отношениями

$$\left| \frac{\lambda_1 - \tilde{\lambda}_1}{\lambda_2 - \lambda_1} \right|, \quad \left| \frac{\lambda_1}{\lambda_2} \right|$$

соответственно. В более общей ситуации обратные итерации со сдвигом используются для нахождения ближайшего к заданному числу собственного значения и соответствующего собственного вектора.

5.2.3. Решение полной проблемы собственных значений

Прямые и обратные итерации хорошо приспособлены для определения отдельных собственных значений и собственных векторов. Для решения спектральной задачи в целом используется QR -алгоритм. Он основан на представлении матрицы A в виде произведения $A = QR$, где Q — ортогональная $Q^*Q = E$, а R — верхняя треугольная матрицы.

Строится последовательность ортогональных матриц Q_k и верхних треугольных матриц R_k по рекуррентным формулам

$$\begin{aligned} A &= Q_1 R_1, & A_1 &= R_1 Q_1, \\ A_1 &= Q_2 R_2, & A_2 &= R_2 Q_2, \\ &\dots & & \\ A_{k-1} &= Q_k R_k, & A_k &= R_k Q_k, \\ &\dots & & \end{aligned}$$

Процесс построения по матрице A матриц Q_k , R_k , A_k называется QR -алгоритмом.

Для любой матрицы A и любой пары индексов i, j , $i \neq j$ всегда можно найти такую матрицу вращения $T(ij)$, что элемент b_{ij} матрицы $B = T^*(ij)AT(ij)$ равен нулю. Определим последовательность матриц

$$A_0 = A, A_1, A_2, \dots,$$

каждая из которых получается из предыдущей с помощью преобразования подобия, определяемой матрицей вращения. На каждом шаге этого процесса обнуляется отдельный внедиагональный элемент. При таком преобразовании сумма квадратов внедиагональных элементов убывает. Последовательность матриц A_k сходится к диагональной матрице и диагональные элементы матрицы A_k в рассматриваемом методе вращений (метод Якоби) являются соответствующими приближениями для собственных значений матрицы A .

Оптимизация метода вращений достигается за счет выбора элемента для уничтожения на каждом шаге преобразований. Это может быть максимальный по модулю внедиагональный элемент всей матрицы A_k или на выбранном столбце.

5.3. Упражнения

Рассматриваются некоторые спектральные свойства квадратных вещественных матриц и алгоритмов нахождения собственных значений.

Упражнение 5.1. Для матрицы $A = A^* > 0$, для которой $\lambda_1 < \lambda_2 < \dots < \lambda_n$, докажите экстремальные свойства отношения Релея.

Решение. В нашем предположении о свойствах матрицы собственные значения $\varphi^1, \varphi^2, \dots, \varphi^n$ образуют ортонормированный базис. Произвольный вектор X разложим по этому базису, т. е.

$$x = \sum_{i=1}^n \alpha_i \varphi^i.$$

Для отношения Релея получим

$$\frac{(Ax, x)}{(x, x)} = \frac{\left(A \sum_{i=1}^n \alpha_i \varphi^i, \sum_{i=1}^n \alpha_i \varphi^i \right)}{\left(\sum_{i=1}^n \alpha_i \varphi^i, \sum_{i=1}^n \alpha_i \varphi^i \right)} = \frac{\sum_{i=1}^n \lambda_i \alpha_i^2}{\sum_{i=1}^n \alpha_i^2}.$$

С учетом наших предположений о спектре матрицы

$$\lambda_1 \leq \frac{(Ax, x)}{(x, x)} \leq \lambda_n. \quad (5.11)$$

Равенство в левой (правой) части неравенства (5.11) достигается на собственном векторе φ^1 (φ^n).

Упражнение 5.2. Получите оценку (5.5) для возмущения собственного значения простой матрицы.

Решение. Будем рассматривать задачу возмущения i -го собственного значения, тем самым

$$\begin{aligned} A\varphi^i &= \lambda_i\varphi^i, \\ \bar{A}\bar{\varphi}^i &= \bar{\lambda}_i\bar{\varphi}^i. \end{aligned}$$

С точностью до членов второго порядка малости имеем

$$A(\bar{\varphi}^i - \varphi^i) + (\bar{A} - A)\varphi^i \approx (\bar{\lambda}_i - \lambda_i)\varphi^i + \lambda_i(\bar{\varphi}^i - \varphi^i). \quad (5.12)$$

Разложим возмущение собственного вектора по собственным векторам невозмущенной задачи:

$$\bar{\varphi}^i - \varphi^i = \sum_{j=1}^n \beta_{ij}\varphi^j$$

и домножим (5.12) скалярно на i -ый собственный вектор сопряженной задачи (5.3). Принимая во внимание свойство ортогональности собственных векторов задач (5.1) и (5.3), из (5.12) получим

$$(\psi^i, (\bar{A} - A)\varphi^i) \approx (\bar{\lambda}_i - \lambda_i)(\psi^i, \varphi^i).$$

С учетом оценки

$$(\psi^i, (\bar{A} - A)\varphi^i) \leq \|\bar{A} - A\| \|\psi^i\| \|\varphi^i\|$$

это дает искомую оценку (5.5) при определении коэффициента перекоса матрицы A согласно (5.6).

Упражнение 5.3. Покажите справедливость асимптотического представления (5.8) для максимального по модулю собственного значения симметричной матрицы A , у которой

$$|\lambda_1| < |\lambda_2| < \dots < |\lambda_n|,$$

при использовании прямых итераций (5.7).

Решение. Система собственных векторов в силу наших предположений о матрице A образует ортонормированный базис. С учетом (5.7) имеем

$$y^0 = \sum_{i=1}^n \alpha_i \varphi^i, \quad y^k = \sum_{i=1}^n \alpha_i \lambda_i^k \varphi^i$$

и поэтому

$$(y^k, y^k) = \sum_{i=1}^n \alpha_i^2 \lambda_i^{2k}, \quad (y^{k+1}, y^k) = \sum_{i=1}^n \alpha_i^2 \lambda_i^{2k+1}.$$

Для отношения этих скалярных произведений имеем

$$\frac{(y^{k+1}, y^k)}{(y^k, y^k)} = \lambda_n \sum_{i=1}^{n-1} \alpha_i^2 \frac{\lambda_i^{2k+1}}{\lambda_n^{2k+1}} \left(\sum_{i=1}^{n-1} \alpha_i^2 \frac{\lambda_i^{2k}}{\lambda_n^{2k}} \right)^{-1}.$$

Это и приводит нас к искомому представлению (5.8).

Упражнение 5.4. Пусть в предположениях предыдущего упражнения известно максимальное по модулю собственное значение λ_n . Пусть

$$y^{k+1} = x^{k+1} - \lambda_n x^k, \quad k = 0, 1, \dots,$$

где x^k определяются в соответствии с (5.7), тогда

$$\frac{(y^{k+1}, y^k)}{(y^k, y^k)} = \lambda_{n-1} + O\left(\left|\frac{\lambda_{n-2}}{\lambda_{n-1}}\right|^{2k}\right), \quad (5.13)$$

т. е. находится следующее собственное значение.

Решение. Разложение по собственным векторам дает

$$y^k = \sum_{i=1}^n \alpha_i \lambda_i^{k-1} (\lambda_i - \lambda_n) \varphi^i$$

и поэтому

$$(y^k, y^k) = \sum_{i=1}^n \alpha_i^2 \lambda_i^{2(k-1)} (\lambda_i - \lambda_n)^2,$$

$$(y^{k+1}, y^k) = \sum_{i=1}^n \alpha_i^2 \lambda_i^{2k-1} (\lambda_i - \lambda_n)^2.$$

После подстановки этих представлений получим выражение (5.13), которое дает

$$\lim_{k \rightarrow \infty} \frac{(y^{k+1}, y^k)}{(y^k, y^k)} = \lambda_{n-1}.$$

Упражнение 5.5. Покажите, что для вещественной матрицы A и ортогональной матрицы Q имеет место

$$\|QA\|_E = \|A\|_E, \quad (5.14)$$

для евклидовой (сферической) нормы:

$$\|A\|_E = \left(\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 \right)^{1/2}.$$

Решение. Непосредственными выкладками убеждаемся в справедливости равенства

$$\|A\|_E^2 = \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 = \text{tr}(A^*A),$$

где

$$\text{tr}(A) = \sum_{i=1}^n a_{ii}.$$

Принимая во внимание, что для ортогональной матрицы $Q^* = Q^{-1}$, получим

$$\|QA\|_E^2 = \text{tr}((QA)^*QA) = \text{tr}(A^*Q^*QA) = \text{tr}(A^*A).$$

Тем самым имеет место равенство (5.14).

Упражнение 5.6. Для симметричной матрицы A постройте матрицу вращения $T(kl)$, которая обращает в нуль элемент b_{kl} матрицы $B = T^*(kl)AT(kl)$.

Решение. Для элементов матрицы вращения $T(kl)$ используем обозначения

$$T_{kk}(k, l) = T_{ll}(k, l) = \cos \theta, \quad T_{kl}(k, l) = -T_{lk}(kl) = -\sin \theta.$$

Рассмотрим вначале матрицу $C = AT_{kl}$. Она отличается от матрицы A только столбцами с номерами k и l :

$$c_{ik} = a_{ik} \cos \theta + a_{il} \sin \theta,$$

$$c_{il} = -a_{ik} \sin \theta + a_{il} \cos \theta,$$

$$b_{ij} = a_{ij}, \quad i = 1, 2, \dots, n, \quad k, l \neq j = 1, 2, \dots, n.$$

В силу такого определения элементов матрицы имеем

$$c_{ik}^2 + c_{il}^2 = a_{ik}^2 + a_{il}^2, \quad i = 1, 2, \dots, n. \quad (5.15)$$

Аналогично, для элементов матрицы $B = T_{kl}^* C$ получим

$$b_{ki} = c_{ki} \cos \theta + c_{li} \sin \theta,$$

$$b_{li} = -c_{ki} \sin \theta + c_{li} \cos \theta,$$

$$c_{ji} = b_{ji}, \quad i = 1, 2, \dots, n, \quad k, l \neq j = 1, 2, \dots, n,$$

причем

$$b_{ki}^2 + b_{li}^2 = c_{ki}^2 + c_{li}^2, \quad i = 1, 2, \dots, n. \quad (5.16)$$

Полагая $i = k, l$ в (5.15), (5.16), с учетом симметричности матриц получим

$$b_{kk}^2 + 2b_{kl}^2 + b_{ll}^2 = a_{kk}^2 + 2a_{kl}^2 + a_{ll}^2.$$

Для диагональных элементов имеем

$$b_{kk} = a_{kk} \cos^2 \theta + a_{ll} \sin^2 \theta + 2a_{kl} \cos \theta \sin \theta,$$

$$b_{ll} = a_{kk} \sin^2 \theta + a_{ll} \cos^2 \theta - 2a_{kl} \cos \theta \sin \theta.$$

С учетом этого требование $b_{kl} = 0$ дает нам

$$\cos 2\theta = \left(1 - \frac{4a_{kl}^2}{(a_{kk} - a_{ll})^2 + a_{kl}^2} \right)^{1/2}.$$

Отсюда мы и получим выражения для элементов матрицы вращения.

5.4. Задачи

Задача 5.1. Пусть $\lambda_1, \lambda_2, \dots, \lambda_n$ — собственные значения матрицы A . Докажите, что

$$\operatorname{tr}(A) = \sum_{i=1}^n \lambda_i, \quad \det(A) = \prod_{i=1}^n \lambda_i.$$

Задача 5.2. Покажите, что для любых квадратных матриц A и B матрицы AB и BA имеют одинаковые характеристические многочлены и, следовательно, одни и те же собственные значения.

Задача 5.3. Докажите, что при $A = A^*$, $B = B^*$ собственные значения матрицы $AB + BA$ вещественны, а матрицы $AB - BA$ чисто мнимые.

Задача 5.4. Пусть вещественная матрица S кососимметричная ($S = -S^*$). Доказать, что матрица преобразования Кели

$$T = (E - S)(E + S)^{-1}$$

ортогональна.

Задача 5.5. Докажите, что каждое собственное значение положительно определенной матрицы положительно.

Задача 5.6. Пусть A — положительно определенная матрица. Показать, что имеет место неравенство Адамара

$$\det(A) \leq \prod_{i=1}^n a_{ii},$$

причем равенство достигается тогда и только тогда, когда A — диагональная матрица.

Задача 5.7. Доказать следующую оценку для евклидовой нормы матрицы:

$$\sum_{i=1}^n |\lambda_i|^2 \leq \|A\|_F^2.$$

Задача 5.8. Пусть при некотором i для всех $k \neq j$ выполняются неравенства

$$|a_{kk} - a_{ii}| > \sum_{k \neq j=1}^n |a_{kj}| + \sum_{i \neq j=1}^n |a_{ij}|.$$

Покажите, что круг Гершгорина

$$|\lambda - a_{ii}| \leq \sum_{i \neq j=1}^n |a_{ij}|$$

содержит только одно собственное значение.

Задача 5.9. Для матрицы $A = A^*$, у которой

$$|\lambda_1| < |\lambda_2| < \dots < |\lambda_n|,$$

для нахождения λ_n используется итерационный процесс (5.7). Установите сходимость y^k при $k \rightarrow \infty$ к собственному вектору φ^n .

Задача 5.10. Исследуйте итерационный процесс обратных итераций для нахождения собственных значений $\lambda_1 = \lambda_2$ и соответствующих собственных векторов, когда

$$|\lambda_1| > |\lambda_3| > \dots > |\lambda_n|. \quad (5.17)$$

Задача 5.11. Рассмотрите возможности решения частичной проблемы значений для симметричной матрицы A при $\lambda_1 = -\lambda_2$ в предположениях (5.17).

Задача 5.12. Докажите, что в QR -алгоритме все матрицы A_k будут верхними почти треугольными, если верхней почти треугольной является матрица A .

Глава 6

Нелинейные уравнения и системы

Многие прикладные задачи приводят к необходимости нахождения приближенного решения нелинейных уравнений и систем нелинейных уравнений. С этой целью используются итерационные методы. Приведены алгоритмы решения нелинейных уравнений с одним неизвестным и систем нелинейных уравнений. Применяются итерационные метод последовательных приближений (простой итерации) и метод Ньютона в различных модификациях.

6.1. Решение нелинейных уравнений и систем

Ищется решение нелинейного уравнения

$$f(x) = 0, \quad (6.1)$$

где $f(x)$ — заданная функция. Корни уравнения (6.1) могут быть комплексными и кратными. Выделяют как самостоятельную проблему разделения корней, когда проводится выделение области в комплексной плоскости, содержащей один корень. После этого на основе тех или иных итерационных процессов при выбранном начальном приближении находится решение нелинейного уравнения (6.1).

В более общем случае мы имеем не одно уравнение (6.1), а систему нелинейных уравнений

$$f_i(x_1, x_2, \dots, x_n) = 0, \quad i = 1, 2, \dots, n. \quad (6.2)$$

Обозначим через $x = \{x_1, x_2, \dots, x_n\}$ вектор неизвестных и определим вектор-функцию $F(x) = \{f_1, f_2, \dots, f_n\}$. Тогда система (6.2) записывается

Основные обозначения

$f(x)$	— функция одной переменной
$f_i(x), i = 1, 2, \dots, n$	— функции n переменных ($x = \{x_i\} = \{x_1, x_2, \dots, x_n\}$)
$F(x) = \{f_1, f_2, \dots, f_n\}$	— вектор-функция с компонентами f_1, f_2, \dots, f_n
$F'(x)$	— матрица Якоби
x^k	— приближенное решение на k -ой итерации

в виде уравнения

$$F(x) = 0. \quad (6.3)$$

Частным случаем (6.3) является уравнение (6.1) ($n = 1$). Второй пример (6.3) — система линейных алгебраических уравнений, когда $F(x) = Ax - f$.

6.2. Итерационные методы решения нелинейных уравнений

Для приближенного решения нелинейных уравнений и систем нелинейных уравнений используются итерационные методы. Среди основных подходов можно выделить метод последовательных приближений (простой итерации) и метод Ньютона.

6.2.1. Алгоритмы для решения нелинейного уравнения

При итерационном решении уравнений (6.1), (6.3) задается некоторое начальное приближение, достаточно близкое к искомому решению x^* . В одношаговых итерационных методах новое приближение x^{k+1} определяется по предыдущему приближению x^k . Говорят, что итерационный метод сходится с линейной скоростью, если $x^{k+1} - x^* = O(x^k - x^*)$ и итерационный метод имеет квадратичную сходимость, если $x^{k+1} - x^* = O((x^k - x^*)^2)$.

Заменяем уравнение (6.1) эквивалентным уравнением

$$x = \varphi(x), \quad (6.4)$$

полагая, например,

$$\varphi(x) = x + g(x)f(x),$$

где функция $g(x)$ не меняет знака на отрезке, на котором ищется решение уравнения (6.1). Для приближенного решения уравнения (6.4) используется метод простой итерации, когда

$$x^{k+1} = \varphi(x^k), \quad k = 0, 1, \dots, \quad (6.5)$$

при некотором заданном начальном приближении x^0 .

Пусть в некоторой окрестности $R = \{x \mid |x - x^*| \leq r\}$ корня $x = x^*$ уравнения (6.4) функция $\varphi(x)$ удовлетворяет условию Липшица

$$|\varphi(x) - \varphi(y)| \leq q|x - y| \quad x, y \in R \quad (6.6)$$

с постоянной $q < 1$. Тогда метод простой итерации (6.5) сходится и для погрешности верна оценка

$$|x^k - x^*| \leq q^k |x^0 - x^*|. \quad (6.7)$$

Можно сформулировать условия, гарантирующие, что имеется единственный корень в окрестности начального приближения x^0 . Пусть теперь $R = \{x \mid |x - x^0| \leq r\}$ и

$$|x^0 - \varphi(x^0)| \leq (1 - q)r, \quad (6.8)$$

тогда при выполнении (6.6) с $q < 1$ уравнение (6.4) имеет единственное решение в R .

В итерационном методе Ньютона (методе касательных) для нового приближения имеем

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)}, \quad k = 0, 1, \dots, \quad f'(x) \equiv \frac{df}{dx}(x). \quad (6.9)$$

Пусть x^* — простой вещественный корень уравнения (6.1) и определим $R = \{x \mid |x - x^*| \leq r\}$ — окрестность этого корня. Предположим также, что

$$\inf_{x \in R} |f'(x)| = m > 0, \quad \sup_{x \in R} |f''(x)| = M,$$

причем

$$|x^0 - x^*| < \frac{2m}{M}.$$

Тогда при $x^0 \in R$ метод Ньютона (6.9) сходится и для погрешности справедлива оценка

$$|x^k - x^*| \leq q^{2k-1} |x^0 - x^*|.$$

Тем самым метод Ньютона имеет квадратичную сходимость.

Модификации метода Ньютона направлены на минимизацию вычислительной работы, на увеличение окрестности корня, в которой можно задавать начальное приближение. Примером выступает метод секущих, который получается из метода Ньютона заменой производной в знаменателе на соответствующую разделенную разность:

$$x^{k+1} = x^k - \frac{x^k - x^{k-1}}{f(x^k) - f(x^{k-1})} f(x^k), \quad k = 0, 1, \dots \quad (6.10)$$

Этот метод является простейшим двухшаговым итерационным методом, когда новое приближение x^{k+1} находится по двум предыдущим x^k и x^{k-1} .

6.2.2. Методы решения систем нелинейных уравнений

При приближенном решении систем нелинейных уравнений (6.3) можно ориентироваться на многомерные аналогии метода простой итерации и метода Ньютона. Многие одношаговые методы для приближенного решения (6.3) по аналогии с двухслойными итерационными методами для решения систем линейных алгебраических уравнений можно записать в виде

$$B_{k+1} \frac{x^{k+1} - x^k}{\tau_{k+1}} + F(x^k) = 0, \quad k = 0, 1, \dots, \quad (6.11)$$

где τ_{k+1} — итерационные параметры, а B_{k+1} — квадратная матрица $n \times n$, имеющая обратную.

Для стационарного итерационного метода (6.11) (B и τ не зависят от k) имеем

$$x^{k+1} = S(x^k), \quad (6.12)$$

где $S(x) = x - \tau B^{-1}F(x)$. Тем самым (6.12) соответствует применению метода простой итерации для преобразованного уравнения

$$x = S(x). \quad (6.13)$$

Пусть в окрестности $R = \{x \mid \|x - x^0\| \leq r\}$ заданного начального приближения x^0 выполнены условия

$$\begin{aligned} \|S(x) - S(y)\| &\leq q\|x - y\|, \quad x, y \in R, \\ \|x^0 - S(x^0)\| &\leq (1 - q)r, \quad q < 1. \end{aligned}$$

Тогда уравнение (6.13) имеет в R единственное решение x^* , которое дается итерационным процессом (6.12), причем для погрешности справедлива оценка

$$\|x^{k+1} - x^*\| \leq q^k \|x^0 - x^*\|.$$

В методе Ньютона новое приближение для решения системы уравнений (6.2) определяется из решения системы линейных уравнений

$$\sum_{j=1}^n (x_j^{k+1} - x_j^k) \frac{\partial f_i(x^k)}{\partial x_j} + f_i(x^k) = 0, \quad (6.14)$$

$$i = 1, 2, \dots, n, \quad k = 0, 1, \dots$$

Определим матрицу Якоби

$$F'(x) = \begin{bmatrix} \frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \dots & \frac{\partial f_1(x)}{\partial x_n} \\ \frac{\partial f_2(x)}{\partial x_1} & \frac{\partial f_2(x)}{\partial x_2} & \dots & \frac{\partial f_2(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(x)}{\partial x_1} & \frac{\partial f_n(x)}{\partial x_2} & \dots & \frac{\partial f_n(x)}{\partial x_n} \end{bmatrix}$$

и запишем (6.14) в виде

$$F'(x) (x^{k+1} - x^k) + F(x^k) = 0, \quad k = 0, 1, \dots \quad (6.15)$$

При использовании записи (6.11) метод Ньютона (6.15) соответствует выбору

$$B_{k+1} = F'(x^k), \quad \tau_{k+1} = 1.$$

Система линейных уравнений (6.15) для нахождения нового приближения x^{k+1} может решаться итерационно. В этом случае мы имеем двухступенчатый итерационный процесс со внешними и внутренними итерациями. Например, внешний итерационный процесс может осуществляться по методу Ньютона, а внутренние итерации — на основе итерационного метода Зейделя.

При решении систем нелинейных уравнений можно использовать прямые аналоги стандартных итерационных методов, которые применяются для решения систем линейных уравнений. Нелинейный метод Зейделя применительно к решению (6.2) дает

$$f_i(x_1^{k+1}, x_2^{k+1}, \dots, x_i^{k+1}, x_{i+1}^k, \dots, x_n^k) = 0, \quad i = 1, 2, \dots, n.$$

В этом случае каждая компонента нового приближения находится из решения нелинейного уравнения, что можно сделать на основе итерационных методов простой итерации и метода Ньютона в различных модификациях. Тем самым снова приходим к двухступенчатому итерационному методу, в котором внешние итерации проводятся в соответствии с методом Зейделя, а внутренние — методом Ньютона.

Основные вычислительные сложности применения метода Ньютона для приближенного решения систем нелинейных уравнений связаны с необходимостью решения линейной системы уравнений с матрицей Якоби на каждой итерации, причем от итерации к итерации эта матрица меняется. В модифицированном методе Ньютона

$$F'(x^0)(x^{k+1} - x^k) + F(x^k) = 0, \quad k = 0, 1, \dots$$

матрица Якоби обращается только один раз.

6.3. Упражнения

Приведены примеры построения и исследования итерационных методов для решения нелинейных уравнений. Для простоты изложения мы ограничились случаем одного уравнения.

Упражнение 6.1. Рассмотрите условия сходимости метода релаксации

$$\frac{x^{k+1} - x^k}{\tau} + f(x^k) = 0, \quad k = 0, 1, \dots$$

при решении уравнения (6.1), если $f'(x) > 0$.

Решение. Будем ориентироваться на использование оценки (6.7) для скорости сходимости метода последовательных приближений (6.5). В нашем случае

$$\varphi(x) = x - \tau f(x).$$

Оценим теперь величину q в (6.6). В нашем случае имеем

$$|\varphi(x) - \varphi(y)| = |x - y - \tau(f(x) - f(y))| \leq q|x - y|,$$

где постоянная

$$q = \max |1 - \tau f'(x + \theta(x - y))| \quad (6.16)$$

с некоторым $\theta \in [0, 1]$.

Для конкретизации условий сходимости и оптимизации выбора итерационного параметра будем считать, что

$$0 < m \leq f'(x) \leq M.$$

В этих предположениях из (6.16) следует, что

$$q(\tau) = \max \{|1 - \tau m|, |1 - \tau M|\}.$$

Итерационный метод будет сходиться ($q < 1$), если $\tau < 2/M$, а для оптимального значения итерационного параметра $\tau = \tau_0$, при котором минимально q , получим

$$\tau_0 = \frac{2}{m + M}, \quad q(\tau_0) = \frac{1 - \xi}{1 + \xi}, \quad \xi = \frac{m}{M}.$$

Упражнение 6.2. В интерполяционных методах нахождение корней уравнения (6.1) основывается на замене функции интерполяционным многочленом. Метод секущих (6.10) связывается с интерполированием многочленом первой степени. Постройте метод решения уравнения (6.1) на основе интерполяционного многочлена второго порядка (метод парабол).

Решение. Пусть известны три приближения x^{k-2} , x^{k-1} и x^k . Новое приближение находится как решение уравнения

$$L_2(x) = 0, \quad (6.17)$$

где $L_2(x)$ — интерполяционный многочлен второго порядка, построенный по узлам x^{k-2} , x^{k-1} и x^k .

Интерполяционный многочлен Ньютона имеет вид

$$L_2(x) = f(x^k) + (x - x^k)f(x^k, x^{k-1}) + \\ + (x - x^k)(x - x^{k-1})f(x^k, x^{k-1}, x^{k-2}),$$

где

$$f(x^k, x^{k-1}) = \frac{f(x^k) - f(x^{k-1})}{x^k - x^{k-1}},$$

$$f(x^k, x^{k-1}, x^{k-2}) = \frac{f(x^{k-1}, x^{k-2}) - f(x^k, x^{k-1})}{x^{k-2} - x^k}$$

— разделенные разности первого и второго порядка соответственно.

Уравнение (6.16) принимает вид

$$a(x - x^k)^2 + b(x - x^k) + c = 0, \quad (6.18)$$

где

$$a = f(x^k, x^{k-1}, x^{k-2}),$$

$$b = f(x^k, x^{k-1}) + (x^k - x^{k-1})f(x^k, x^{k-1}, x^{k-2}), \quad c = f(x^k).$$

Решая квадратное уравнение (6.18) найдем два, в общем случае комплексных, корня. В качестве нового приближения x^{k+1} в методе парабол берется корень уравнения (6.18), который ближе к x_k .

Упражнение 6.3. Обсудите возможность ускорения итерационного метода, имеющего линейную скорость сходимости.

Решение. Для итерационного метода в силу наших предположений о линейной сходимости имеем

$$x^k - x^* \approx aq^k, \quad q \in (0, 1).$$

Числа a, q, x^* — неизвестны и для их нахождения рассмотрим три последовательных итерации x^{k-2}, x^{k-1}, x^k и составим приближенные равенства

$$x^{k-2} - x^* \approx aq^{k-2}, \quad x^{k-1} - x^* \approx aq^{k-1}, \quad x^k - x^* \approx aq^k.$$

Отсюда находим

$$x^* \approx x_k - \frac{(x^k - x^{k-1})^2}{x^k - 2x^{k-1} - x^{k-2}}$$

и поэтому в качестве нового приближения можно взять

$$x^{k+1} = x_k - \frac{(x^k - x^{k-1})^2}{x^k - 2x^{k-1} - x^{k-2}}.$$

Это есть метод Эйткена ускорения сходимости итерационных методов решения нелинейных уравнений.

Упражнение 6.4. *Покажите, что в методе Ньютона (6.9) последовательность приближений либо монотонно убывает ($x^* < x^{k+1} < x^k$ для всех k), либо монотонно возрастает ($x^* > x^{k+1} > x^k$ для всех k), если производная функции $f(x)$ сохраняет знак и монотонна.*

Решение. Будем считать, для определенности, что

$$f'(x) > 0, \quad f''(x) > 0.$$

Случай отрицательных первой и второй производных функции $f(x)$ рассматривается аналогично.

При задании начального приближения x^0 имеем, например, $x^0 > x^*$ (вторая возможность $x^0 < x^*$). Доказательство проводится по индукции. Предположим, что $x^k > x^*$ и докажем, что тогда

$$x^k > x^{k+1} > x^*.$$

Запишем (6.9) в виде

$$x^k - x^{k+1} = \frac{f(x^k) - f(x^*)}{f'(x^k)}. \quad (6.19)$$

Для правой части имеем

$$\frac{f(x^k) - f(x^*)}{f'(x^k)} = \frac{(x^k - x^*)f'(\xi^k)}{f'(x^k)},$$

где $\xi^k \in (x^*, x^k)$. В наших предположениях о производных функции $f(x)$ получим

$$0 < \frac{f'(\xi^k)}{f'(x^k)} < 1$$

и поэтому из (6.19) следует неравенство

$$0 < x^k - x^{k+1} < x^k - x^*,$$

которое обеспечивает монотонность итераций метода Ньютона.

6.4. Задачи

Задача 6.1. Обсудите возможность отделения известного корня x^* уравнения (6.1), применяя итерационный метод к уравнению

$$g(x) = \frac{f(x)}{x - x^*} = 0.$$

Задача 6.2. Покажите, что итерационный метод Ньютона (6.9) можно рассматривать как интерполяционный метод при использовании интерполяционного многочлена Эрмита первой степени.

Задача 6.3. Постройте методы нахождения корней уравнения (6.1) на основе линейной и квадратичной интерполяции функции $x = \varphi(y)$, обратной $y = f(x)$.

Задача 6.4. Покажите, что если корень x^* уравнения (6.1) имеет кратность p , то квадратичную сходимость имеет метод Ньютона в следующей модификации

$$f'(x^k) \frac{x^{k+1} - x^k}{p} + f(x^k) = 0, \quad k = 0, 1, \dots$$

Задача 6.5. Покажите, что итерационный метод (метод Стеффенсена)

$$x^{k+1} = x^k - \frac{f(x^k)}{f(x^k + f(x^k)) - f(x^k)} f(x^k), \quad k = 0, 1, \dots \quad (6.20)$$

обладает квадратичной сходимостью.

Задача 6.6. Примените процесс Эйткена для ускорения сходимости итерационного метода

$$x^{k+1} = x^k - f(x^k), \quad k = 0, 1, \dots$$

и дайте интерпретацию полученного метода как метода Стеффенсена.

Задача 6.7. Исследуйте скорость сходимости итерационного метода Чебышева

$$x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)} - \frac{f^2(x^k) f''(x^k)}{2(f'(x^k))^3}, \quad k = 0, 1, \dots$$

Задача 6.8. Сформулируйте достаточные условия сходимости нелинейного метода Якоби

$$f_i(x_1^k, x_2^k, \dots, x_i^{k+1}, x_{i+1}^k, \dots, x_n) = 0, \quad i = 1, 2, \dots, n$$

для приближенного решения системы уравнений (6.2).

Задача 6.9. Постройте метод Стеффенсена (см. (6.20)) для решения систем нелинейных уравнений.

Задача 6.10. Пусть в системе уравнений (6.3) имеет место разложение $F(x) = F_1(x) + F_2(x)$. Рассмотрите итерационный метод (нелинейный аналог классического итерационного метода переменных направлений)

$$\frac{x^{k+1/2} - x^k}{\tau} + F_1(x^{k+1/2}) + F_2(x^k) = 0,$$

$$\frac{x^{k+1} - x^{k+1/2}}{\tau} + F_1(x^{k+1/2}) + F_2(x^{k+1}) = 0, \quad k = 0, 1, \dots,$$

который основан на последовательном решении двух нелинейных систем уравнений.

Глава 7

Задачи минимизации функций

Среди основных проблем вычислительной математики можно отметить задачи минимизации функций многих переменных (задачи оптимизации). Поиск минимума часто проводится при некоторых дополнительных ограничениях — условная оптимизация. Для численного решения таких задач используются итерационные методы. В задачах с ограничениями применяются методы штрафных функций. Простейшей задачей рассматриваемого класса является поиск минимума одномерной функции.

7.1. Поиск минимума функции многих переменных

Для заданной функции $f(x)$, определенной на допустимом множестве X из евклидова пространства \mathbf{R}^n , ищутся точки минимума (максимума) функции $f(x)$, т. е.

$$f(x) \rightarrow \min, \quad x \in X. \quad (7.1)$$

Точка $x^* \in X$ есть точка глобального минимума функции $f(x)$ на множестве X , если

$$f(x^*) \leq f(x), \quad \forall x \in X, \quad (7.2)$$

и точка локального минимума, если $f(x^*) \leq f(x)$ в окрестности точки $x^* \in X$.

Задача (7.1) называется задачей безусловной оптимизации, если $X = \mathbf{R}^n$, т. е.

$$f(x) \rightarrow \min, \quad x \in \mathbf{R}^n. \quad (7.3)$$

Если X некоторое подмножество пространства \mathbf{R}^n , то мы имеем задачу условной оптимизации. Такие задачи существенно сложнее для численного решения, чем задачи безусловной минимизации. Ограничения могут

Основные обозначения

$x = \{x_i\} = \{x_1, x_2, \dots, x_n\}$ — n -мерный вектор

$f(x)$ — функция одной или
 n переменных

$f(x) \rightarrow \min, x \in \mathbb{R}^n$ — задача безусловной оптимизации

$f(x) \rightarrow \min, x \in X$ — задача условной оптимизации

X — допустимое множество

x^k — приближенное решение на k -ой
итерации

$$(x, y) = \sum_{i=1}^n x_i y_i \text{ — скалярное произведение}$$

формулироваться в виде равенств (например, $g_i(x) = 0, i = 1, 2, \dots, m$) или неравенств ($g_i(x) \leq 0, i = 1, 2, \dots, m$).

7.2. Методы решения задач оптимизации

Вычислительные алгоритмы для приближенного решения задачи оптимизации чаще всего строятся на использовании необходимых и достаточных условий оптимальности, т. е. условий, которые имеют место в точке минимума. Реализация такого подхода связана с решением соответствующих нелинейных уравнений итерационными методами.

7.2.1. Поиск минимума функции одной переменной

Пусть $X = [a, b]$ и кусочно-непрерывная функция $f(x)$ имеет в некоторой точке $x^* \in X$ один минимум. Мы отметим прежде всего простейшие итерационные методы решения задачи минимизации, наиболее полно учитывающие специфику одномерных задач.

Вычислим функцию на концах отрезка и в двух внутренних точках x^1 и $x^2 < x^1$. Будем считать, что эти точки симметричны относительно середины отрезка $[a, b]$. В методе золотого сечения точки x^1 и x^2 выбираются так, чтобы отношение длины всего отрезка $[a, b]$ к длине большей из его

частей $[a, x^1]$ равнялось отношению длины большей части $[a, x^1]$ к длине меньшей части $[x^1, b]$:

$$\frac{b-a}{x^1-a} = \frac{x^1-a}{x^2-a}. \quad (7.4)$$

Далее проводится сравнение значений функции в четырех точках a, x^2, x^1, b и выбирается точка, в которой значение функции наименьшее. Пусть это будет точка x^2 , тогда минимум функции достигается в одном из прилегающих к этой точке отрезков: $[a, x^2]$ или $[x^2, x^1]$ и поэтому в дальнейшем можно рассматривать проблему минимизации на отрезке $[a, x^1]$. После этого процесс повторяется — в соответствии с правилом золотого сечения делится точкой x^3 отрезок $[a, x^1]$.

Для минимизации функции одной переменной широко используются методы полиномиальной интерполяции. В этом случае с использованием ранее найденных точек строится интерполяционный полином, точка минимума которого принимается за очередное приближение. В методе парабол используется интерполяционный многочлен второго порядка.

Пусть, например, известны три приближения $x^{k-2} < x^{k-1}$ и $x^{k-2} < x^k < x^{k-1}$, причем $f(x^k) < f(x^{k-2})$ и $f(x^k) < f(x^{k-1})$. Новое приближение ищется как решение задачи минимизации

$$L_2(x) \rightarrow \min, \quad (7.5)$$

где $L_2(x)$ — интерполяционный многочлен второго порядка, построенный по узлам x^{k-2}, x^{k-1} и x^k .

Интерполяционный многочлен Ньютона имеет вид

$$L_2(x) = f(x^k) + (x - x^k)f(x^k, x^{k-1}) + \\ + (x - x^k)(x - x^{k-1})f(x^k, x^{k-1}, x^{k-2}),$$

где

$$f(x^k, x^{k-1}) = \frac{f(x^k) - f(x^{k-1})}{x^k - x^{k-1}}, \\ f(x^k, x^{k-1}, x^{k-2}) = \frac{f(x^{k-1}, x^{k-2}) - f(x^k, x^{k-1})}{x^{k-2} - x^k}$$

— разделенные разности первого и второго порядка соответственно. Решение задачи (7.5) приводит нас к следующей формуле для нового приближения для точки минимума

$$2x^{k+1} = x^k + x^{k-1} - \frac{f(x^k, x^{k-1})}{f(x^k, x^{k-1}, x^{k-2})}. \quad (7.6)$$

Для дифференцируемой функции $f(x)$ строятся итерационные методы, основанные на решении уравнения (необходимое условие минимума)

$$f'(x) = 0. \quad (7.7)$$

Корень этого уравнения $x^* \in X$ является точкой минимума, если $f''(x) > 0$ (достаточные условия минимума). Для приближенного решения нелинейного уравнения используются итерационные методы.

В итерационном методе Ньютона новое приближение для точки минимума определяется в соответствии с формулой

$$x^{k+1} = x^k - \frac{f'(x^k)}{f''(x^k)}, \quad k = 0, 1, \dots \quad (7.8)$$

Различные модификации метода Ньютона рассматривались в главе 7.

7.2.2. Минимизация функций многих переменных

Для функции многих переменных $f(x)$, $x = \{x_1, x_2, \dots, x_n\}$ определим вектор первых частных производных (градиент)

$$f'(x) = \left\{ \frac{\partial f}{\partial x_i}(x) \right\} \equiv \left\{ \frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right\}.$$

Матрица вторых частных производных (гессиан) в точке x есть

$$f''(x) = \left\{ \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right\}.$$

Будем рассматривать задачу безусловной оптимизации (7.3). Пусть функция $f(x)$ дифференцируема в точке локального минимума $x = x^*$, тогда (необходимые условия оптимальности)

$$f'(x^*) = 0. \quad (7.9)$$

Пусть функция $f(x)$ дважды дифференцируема в точке $x = x^*$ и выполнено (7.9). Если матрица $f''(x)$ положительно определена, т. е.

$$(f''(x^*)y, y) > 0, \quad \forall y \neq 0, \quad (7.10)$$

тогда x^* — точка локального минимума. Условия (7.9), (7.10) есть достаточные условия оптимальности.

Для итерационных методов минимизации будем использовать обозначения

$$x^{k+1} = x^k + \alpha_k h^k, \quad k = 0, 1, \dots, \quad (7.11)$$

где h^k — вектор, который определяет направление $(k+1)$ -го шага минимизации, а коэффициент α_k — длину этого шага.

Вектор h задает направление убывания функции $f(x)$ в точке x , если $f(x + \alpha h) < f(x)$ при достаточно малых $\alpha > 0$. Если вектор h^k задает направление убывания функции $f(x)$ в точке x^k , а $\alpha_k > 0$ такое, что

$$f(x^{k+1}) < f(x^k),$$

то итерационный метод (7.11) называется методом спуска. В градиентном методе $h^k = -f'(x^k)$, т. е.

$$x^{k+1} = x^k - \alpha_k f'(x^k), \quad k = 0, 1, \dots \quad (7.12)$$

Особое внимание уделяется выбору итерационных параметров α_k , $k = 0, 1, \dots$ в методе (7.11). Их можно определять из условия

$$f(x^k + \alpha_k h^k) = \min_{\alpha \geq 0} f(x^k + \alpha h^k),$$

т. е. из решения дополнительной одномерной задачи минимизации.

В вычислительной практике широко используется процедура дробления шага, когда параметр α уменьшается, например, в два раза, до тех пор пока не будет выполнено неравенство

$$f(x^k + \alpha h^k) < f(x^k).$$

При применении метода Ньютона для решения системы нелинейных уравнений (7.9) получим

$$f''(x^k)(x^{k+1} - x^k) + f'(x^k) = 0, \quad k = 0, 1, \dots \quad (7.13)$$

Он записывается в виде (7.11) при

$$\alpha_k = 1, \quad h^k = -(f''(x^k))^{-1} f'(x^k). \quad (7.14)$$

Среди модификаций метода Ньютона отметим метод Ньютона с регуляризацией шага, когда вместо (7.14) используется

$$\alpha_k > 0, \quad h^k = -(f''(x^k))^{-1} f'(x^k).$$

В квазиньютоновских методах

$$h^k = -H_k f'(x^k),$$

где H_k — матрица, которая аппроксимирует матрицу $(f''(x^k))^{-1}$.

7.2.3. Задачи условной минимизации

При минимизации функций с ограничениями широко используются подходы, аналогичные разработанным для задач безусловной минимизации. Наиболее просто это реализуется при переходе от задачи условной минимизации к задаче минимизации без ограничений.

Рассмотрим задачу минимизации с ограничениями типа равенств:

$$f(x) \rightarrow \min, \quad g_i(x) = 0, \quad i = 1, 2, \dots, m. \quad (7.15)$$

Эту задачу можно записать в общем виде (7.1), задав допустимое множество

$$X = \{x \in \mathbf{R}^n \mid g_i(x) = 0, \quad i = 1, 2, \dots, m\}.$$

При некоторых ограничениях задача условной минимизации (7.15) эквивалентна задаче безусловной минимизации функции Лагранжа

$$\Phi(x, y) = f(x) + \sum_{i=1}^m y_i g_i(x),$$

где y_i — неизвестные множители Лагранжа. Тем самым мы приходим к задаче минимизации функции $n + m$ переменных.

Более сложно перейти к задаче безусловной оптимизации при учете ограничений в виде неравенств. Рассмотрим, например, задачу

$$f(x) \rightarrow \min, \quad g_i(x) \leq 0, \quad i = 1, 2, \dots, m, \quad (7.16)$$

т. е. в (7.1)

$$X = \{x \in \mathbf{R}^n \mid g_i(x) \leq 0, \quad i = 1, 2, \dots, m\}.$$

Вместо функции $f(x)$ в методе штрафов минимизируется функция

$$\Phi(x, \varepsilon) = f(x) + \psi(x, \varepsilon), \quad (7.17)$$

где $\psi(x, \varepsilon)$ — штрафная функция, $\varepsilon > 0$ — параметр штрафа. Выбор штрафной функции на допустимом множестве подчинен условиям

$$\psi(x, \varepsilon) \geq 0, \quad \psi(x, \varepsilon) \rightarrow 0, \quad \text{если } \varepsilon \rightarrow 0, \quad x \in X,$$

а вне допустимого множества —

$$\psi(x, \varepsilon) \rightarrow \infty, \quad \text{если } \varepsilon \rightarrow 0, \quad x \notin X.$$

В качестве характерного примера приведем штрафную функцию для задачи условной минимизации (7.16):

$$\psi(x, \varepsilon) = \frac{1}{\varepsilon} \sum_{i=1}^m [\max\{0, g_i(x)\}]^2.$$

После этого рассматривается задача безусловной минимизации

$$\Phi(x, \varepsilon) \rightarrow \min, \quad x \in \mathbb{R}^n.$$

Помимо выбора штрафной функции в методах этого класса очень важен выбор величины параметра штрафа ε .

7.3. Упражнения

Приведены примеры решения задач, связанных с построением и исследованием численных методов приближенного решения задач минимизации функций.

Упражнение 7.1. Пусть точка x^1 проводит золотое сечение отрезка $[a, b]$, причем $x^1 > a + (b - a)/2$, а x^2 — точка золотого сечения отрезка $[a, x^1]$. Покажите, что точки x^1, x^2 расположены симметрично относительно середины отрезка $[a, b]$ и поэтому x^1 есть точка золотого сечения и отрезка $[x^2, b]$.

Решение. Точка x^1 является точкой золотого сечения, если

$$\frac{b - a}{x^1 - a} = \frac{x^1 - a}{b - x^1}.$$

Исходя из этого определения имеем

$$x^1 = a + \frac{\sqrt{5} - 1}{2}(b - a).$$

Аналогично для x^2 имеем

$$\frac{x^1 - a}{x^2 - a} = \frac{x^2 - a}{x^1 - x^2}$$

и поэтому

$$x^2 = a + \frac{3 - \sqrt{5}}{2}(b - a).$$

В силу этого

$$x^1 - \left(a + \frac{b - a}{2}\right) = a + \frac{b - a}{a} - x^2,$$

т. е. точки x^1 и x^2 расположены симметрично относительно центра отрезка $[a, b]$.

Упражнение 7.2. Функция $f(x)$ называется выпуклой на X , если

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

для всех $x, y \in X$ и $\lambda \in [0, 1]$. Пусть функция $f(x)$ выпукла на \mathbb{R}^n и дифференцируема в точке $x^* \in \mathbb{R}^n$. Покажите, что если выполнено

$$f'(x^*) = 0, \tag{7.18}$$

то x^* — точка минимума функции $f(x)$.

Решение. Для любых $x \in \mathbb{R}^n$ и $\lambda \in (0, 1]$ имеем

$$f(\lambda x + (1 - \lambda)x^*) \leq \lambda f(x) + (1 - \lambda)f(x^*).$$

Принимая во внимание дифференцируемость функции $f(x)$ в точке x^* , получим

$$\begin{aligned} f(x) - f(x^*) &\geq \frac{f(x^* + \lambda(x - x^*)) - f(x^*)}{\lambda} = \\ &= \frac{(f'(x^*), \lambda(x - x^*)) + o(\lambda)}{\lambda} = \frac{o(\lambda)}{\lambda}. \end{aligned}$$

Предельный переход $\lambda \rightarrow 0$ дает искомое неравенство $f(x) \geq f(x^*)$.

Важно отметить, что условие (7.18) есть необходимое условие минимума. Мы установили, что это условие является и достаточным при поиске минимума выпуклой функции.

Упражнение 7.3. *Покажите, что итерационный метод Ньютона сходится за одну итерацию при минимизации квадратичной функции*

$$f(x) = \frac{1}{2}(Ax, x) + (b, x),$$

где A — симметричная положительно определенная матрица $n \times n$.

Решение. В нашем случае

$$f'(x) = Ax - b,$$

$$f''(x) = A.$$

При применении метода Ньютона (7.13) при произвольном начальном приближении x^0 имеем

$$f'(x^1) = Ax^1 - b = 0,$$

т. е. $x^1 = x^*$.

Упражнение 7.4. *Для задачи условной минимизации (7.16) приведите несколько примеров штрафных функций.*

Решение. Ограничения в виде неравенств $g_i(x) \leq 0$, $i = 1, 2, \dots, m$ учитываются выбором штрафных функций

$$\begin{aligned} \psi(x, \varepsilon) &= \frac{1}{\varepsilon} \sum_{i=1}^m [\max\{0, g_i(x)\}]^p, \quad p > 1, \\ \psi(x, \varepsilon) &= \sum_{i=1}^m \exp\left(\frac{1}{\varepsilon} g_i(x)\right), \\ \psi(x, \varepsilon) &= \begin{cases} -\varepsilon \sum_{i=1}^m \frac{1}{g_i(x)}, & \text{если } g_i(x) < 0, \quad i = 1, 2, \dots, m, \\ \infty, & \text{в противном случае,} \end{cases} \\ \psi(x, \varepsilon) &= \begin{cases} -\varepsilon \sum_{i=1}^m \ln[-g_i(x)], & \text{если } g_i(x) < 0, \quad i = 1, 2, \dots, m, \\ \infty, & \text{в противном случае.} \end{cases} \end{aligned} \tag{7.19}$$

Наиболее часто используется квадратичная штрафная функция (7.19) с $p = 2$.

7.4. Задачи

Задача 7.1. Найдите длину отрезка локализации минимума одномерной функции после N шагов минимизации на отрезке $[a, b]$ методом золотого сечения.

Задача 7.2. В методе дихотомии первая пара точек есть

$$x^1 = \frac{a+b}{2} + \delta, \quad x^2 = \frac{a+b}{2} - \delta.$$

Каждая последующая пара точек выбирается на расстоянии δ по обе стороны от середины отрезка локализации. Проведите анализ вычислительных аспектов метода дихотомии и сравните его с методом золотого сечения.

Задача 7.3. Пусть A — симметричная положительно определенная матрица $n \times n$ и при минимизации квадратичной функции

$$f(x) = \frac{1}{2}(Ax, x) + (b, x)$$

используется итерационный метод (7.11). Найдите оптимальное значение итерационного параметра из условия

$$f(x^k + \alpha_k h^k) = \min_{\alpha \geq 0} f(x^k + \alpha h^k).$$

Задача 7.4. Рассмотрите метод кубической интерполяции при минимизации одномерной функции $f(x)$, который основан на построении многочлена третьей степени $\varphi(x)$ на отрезке локализации $[x^{k-1}, x^k]$ по условиям

$$\varphi(x^{k-1}) = f(x^{k-1}), \quad \varphi'(x^{k-1}) = f'(x^{k-1}),$$

$$\varphi(x^k) = f(x^k), \quad \varphi'(x^k) = f'(x^k).$$

Задача 7.5. Рассмотрите условия сходимости градиентного итерационного метода (7.12) при минимизации функции $f(x)$, для которой

$$m(y, y) \leq (f''(x)y, y) \leq M(y, y), \quad m > 0.$$

Задача 7.6. Покажите, что если матрица $f''(x^k)$ положительно определена, $f'(x^k) \neq 0$, то направление $h^k = -(f''(x^k))^{-1} f'(x^k)$ является направлением убывания функции $f(x)$ в точке x^k .

Задача 7.7. Получите оценку сходимости метода Ньютона (7.13) при минимизации дважды дифференцируемой функции $f(x)$ при предположениях

$$(f''(x)y, y) \geq m(y, y), \quad m > 0,$$

$$\|f''(x) - f''(y)\| \leq M\|x - y\|.$$

Задача 7.8. Проекцией точки $a \in \mathbb{R}^n$ $X \subset \mathbb{R}^n$ называется точка $P_X(a) \in X$ такая, что

$$\|P_X(a) - a\| \leq \|x - a\|, \quad \forall x \in X,$$

т.е. точка, ближайшая к a среди всех точек X . Рассмотрите метод проекции градиента при решении задачи условной минимизации (7.1), когда

$$x^{k+1} = P_X(x^k - \alpha_k f'(x^k)), \quad k = 0, 1, \dots$$

Задача 7.9. Покажите, что штрафная функция

$$\psi(x, \varepsilon) = \frac{1}{\varepsilon} \sum_{i=1}^m [\max\{0, g_i(x)\}]^p, \quad p > 1$$

является выпуклой непрерывно дифференцируемой функцией, если функции $g_i(x)$, $i = 1, 2, \dots$ — выпуклые непрерывно дифференцируемые функции.

Задача 7.10. Постройте примеры функций штрафа для задачи минимизации с ограничениями типа равенств:

$$f(x) \rightarrow \min, \quad g_i(x) = 0, \quad i = 1, 2, \dots, m.$$

Глава 8

Интегральные уравнения

Среди типичных интегральных уравнений можно выделить интегральные уравнения Фредгольма второго рода. Для их приближенного решения применяется метод квадратур. Второй широко используемый класс методов решения интегральных уравнений — различные варианты проекционных методов. Отдельно необходимо выделить интегральные уравнения с переменными пределами интегрирования — интегральные уравнения Вольтерра. Интегральные уравнения Фредгольма первого рода являются характерным примером некорректных задач, для численного решения которых используются методы регуляризации.

8.1. Задачи для интегральных уравнений

Будем рассматривать одномерные интегральные уравнения, решение которых есть $u(x)$, $x \in [a, b]$. Линейное интегральное уравнение с постоянными пределами интегрирования (уравнение Фредгольма) записывается в виде

$$g(x)u(x) - \lambda \int_a^b K(x, s)u(s) ds = f(x), \quad x \in [a, b], \quad (8.1)$$

где $K(x, s)$ — ядро интегрального уравнения, а $g(x), f(x)$ — заданные функции, а λ заданный или неизвестный числовой параметр.

Наибольшее внимание уделяется нахождению приближенного решения интегрального уравнения Фредгольма второго рода:

$$u(x) - \lambda \int_a^b K(x, s)u(s) ds = f(x), \quad x \in [a, b] \quad (8.2)$$

при заданном λ .

Основные обозначения

$K(x, s)$	— ядро интегрального уравнения
λ	— числовой параметр
$x_i, i = 1, 2, \dots, n$	— узлы квадратурной формулы
α	— параметр регуляризации
y^k	— приближенное решение на k -ой итерации
$\varphi^i(x), i = 1, 2, \dots, n$	— линейно независимые координатные функции

При $f(x) \equiv 0$ уравнение (8.2) есть однородное уравнение Фредгольма

$$u(x) - \lambda \int_a^b K(x, s) u(s) ds = 0, \quad x \in [a, b], \quad (8.3)$$

которое всегда имеет тривиальное решение $u(x) \equiv 0$. Те значения параметра λ , при которых уравнение (8.3) имеет ненулевое решение, называются характеристическими числами, а соответствующие ненулевые решения уравнения — собственными функциями ($1/\lambda$ — собственные значения).

Линейное интегральное уравнение Фредгольма первого рода имеет вид

$$\int_a^b K(x, s) u(s) ds = f(x), \quad x \in [a, b], \quad (8.4)$$

т. е. в общей записи (8.1) $g(x) = 0$ и $\lambda = -1$. Принципиальные трудности приближенного решения этого уравнения порождены тем, что задача нахождения решения интегрального уравнения первого рода является некорректно поставленной. Некорректность обусловлена, прежде всего, отсутствием устойчивости решения по отношению к малым возмущениям правой части уравнения (8.4).

Отдельного рассмотрения заслуживают интегральные уравнения с переменными пределами интегрирования. Интегральное уравнение Воль-

terra второго рода записывается в виде

$$u(x) - \lambda \int_a^x K(x, s) u(s) ds = f(x), \quad x \in [a, b]. \quad (8.5)$$

По аналогии с (8.4), (8.5) для уравнения Вольтерра первого рода имеем

$$\int_a^x K(x, s) u(s) ds = f(x), \quad x \in [a, b]. \quad (8.6)$$

В вычислительной практике рассматриваются и более общие задачи для интегральных уравнений, среди которых отметим прежде всего многомерные интегральные уравнения. Большое внимание уделяется разработке численных методов для специальных классов интегральных уравнений. Отметим, в частности, интегральные уравнения с разностным ядром $K(x - s)$.

8.2. Методы решения интегральных уравнений

Выделены основные классы методов приближенного решения интегральных уравнений. Метод квадратур (механических квадратур) основан на замене интегралов конечными суммами с использованием квадратурных формул. В проекционных методах приближенное решение ищется в виде разложения по системе известных линейно независимых функций. Отмечаются особенности решения интегральных уравнений Вольтерра, кратко обсуждаются методы решения интегральных уравнений первого рода.

8.2.1. Интегральные уравнения Фредгольма второго рода

Будем рассматривать алгоритмы численного решения интегральных уравнений (8.2), считая заданным параметр λ . В основе метода квадратур лежит та или иная квадратурная формула. Пусть $x_1 < x_2 < \dots < x_n$ — узлы, а c_i , $i = 1, 2, \dots, n$ — коэффициенты квадратурной формулы на отрезке интегрирования $[a, b]$. При использовании квадратурной формулы

$$\int_a^b \theta(x) dx \approx \sum_{i=1}^n c_i \theta(x_i)$$

приближенное решение интегрального уравнения (8.2) определим из системы линейных алгебраических уравнений

$$y_i - \lambda \sum_{j=1}^n c_j K(x_i, s_j) y_j = f(x_i), \quad i = 1, 2, \dots, n, \quad (8.7)$$

где y_i — приближенное решение в узле x_i , $i = 1, 2, \dots, n$.

В проекционных методах приближенное решение интегрального уравнения (8.2) ищется в виде

$$y(x) = \sum_{i=1}^n c_i \varphi_i(x), \quad (8.8)$$

где $\varphi_i(x)$, $i = 1, 2, \dots, n$ — заданные линейно независимые функции, которые называются координатными. Часто удобнее ориентироваться на несколько отличное от (8.8) представление приближенного решения:

$$y(x) = f(x) + \sum_{i=1}^n c_i \varphi_i(x), \quad (8.9)$$

Метод проекционного типа характеризуется выбором координатных функций $\varphi_i(x)$, $i = 1, 2, \dots, n$ и способом определения вектора неизвестных коэффициентов $c = \{c_1, c_2, \dots, c_n\}$. Отметим некоторые возможности по нахождению коэффициентов в представлениях (8.8), (8.9).

При использовании представления (8.8) определим невязку

$$r(x, c) = \sum_{i=1}^n c_i \varphi_i(x) - \lambda \int_a^b K(x, s) \sum_{j=1}^n c_j \varphi_j(s) ds - f(x).$$

В методе наименьших квадратов постоянные c_i , $i = 1, 2, \dots, n$ находятся из минимума квадрата нормы невязки в $L_2(a, b)$, т. е.

$$J(c) \equiv \int_a^b r^2(x, c) dx \rightarrow \min, \quad c \in \mathbf{R}^n.$$

Для определения c_i , $i = 1, 2, \dots, n$ получим систему линейных алгебраических уравнений

$$\sum_{j=1}^n a_{ij} c_j = b_i, \quad i = 1, 2, \dots, n, \quad (8.10)$$

где

$$a_{ij} = \int_a^b \left(\varphi_i(x) - \lambda \int_a^b K(x,s)\varphi_i(s) ds \right) \left(\varphi_j(x) - \lambda \int_a^b K(x,s)\varphi_j(s) ds \right) dx,$$

$$b_i = \int_a^b f(x) \left(\varphi_i(x) - \lambda \int_a^b K(x,s)\varphi_i(s) ds \right) dx, \quad i = 1, 2, \dots, n.$$

Тем самым матрица системы (8.10) симметрична.

В методе Галеркина коэффициенты c_i , $i = 1, 2, \dots, n$ определяются из условия ортогональности в $L_2(a, b)$ невязки $r(x, c)$ функциям $\varphi_i(x)$, $i = 1, 2, \dots, n$:

$$\int_a^b r(x, c) \varphi_i(x) dx, \quad i = 1, 2, \dots, n.$$

В этом случае имеем систему линейных уравнений (8.10), в которой

$$a_{ij} = \int_a^b \left(\varphi_j(x) - \lambda \int_a^b K(x,s)\varphi_j(s) ds \right) \varphi_i(x) dx,$$

$$b_i = \int_a^b f(x) \varphi_i(x) dx, \quad i = 1, 2, \dots, n.$$

Отметим среди проекционных методов и метод коллокации. В этом случае на отрезке $[a, b]$ выбирается n точек коллокации x_i , $i = 1, 2, \dots, n$ и коэффициенты c_i , $i = 1, 2, \dots, n$ в представлении (8.8) (или (8.9)) выбираются так, что невязка обращалась в нуль в точках коллокации, т. е.

$$r(x_i, c) = 0, \quad i = 1, 2, \dots, n.$$

Для коэффициентов матрицы и правой части системы (8.10) при использовании представления (8.8) получим

$$a_{ij} = \varphi_j(x_i) - \lambda \int_a^b K(x_i, s)\varphi_j(s) ds,$$

$$b_i = f(x_i), \quad i = 1, 2, \dots, n.$$

Для решения системы линейных алгебраических уравнений (8.10) применяются прямые или итерационные методы.

8.2.2. Интегральные уравнения с переменными пределами интегрирования

При приближенном решении интегрального уравнения Вольтерра второго рода (8.5) используется как метод квадратур, так и проекционные методы. Для определенности, будем считать, что $x_1 = a$, $x_n = b$. Для точек x_i , $i = 1, 2, \dots, n$ из (8.5) получим

$$u(x_i) - \lambda \int_a^{x_i} K(x_i, s) u(s) ds = f(x_i), \quad i = 1, 2, \dots, n. \quad (8.11)$$

Принимая во внимание то, что интегрировать необходимо по отрезку переменной длины, запишем используемую квадратурную формулу в виде

$$\int_a^{x_i} \theta(x) dx \approx \sum_{j=1}^i c_j^{(i)} \theta(x_j), \quad i = 2, 3, \dots, n.$$

Применение к (8.11) дает систему линейных уравнений

$$y_i - \lambda \sum_{j=1}^i c_j^{(i)} K(x_i, s_j) y_j = f(x_i), \quad i = 1, 2, \dots, n. \quad (8.12)$$

Отличительная особенность системы уравнений (8.12) состоит в том, что матрица ее коэффициентов треугольная. Это позволяет найти приближенное решение интегрального уравнения y_1, y_2, \dots, y_n последовательно друг за другом по рекуррентным формулам в предположении, что все диагональные элементы матрицы ненулевые. Наиболее простые расчетные формулы при решении интегрального уравнения Вольтерра второго рода мы получим при использовании квадратурной формулы трапеций.

При численном решении интегрального уравнения первого рода (8.6) можно ориентироваться на использование метода квадратур. Подобно (8.11) из (8.6) будем иметь

$$\int_a^{x_i} K(x_i, s) u(s) ds = f(x_i), \quad i = 1, 2, \dots, n,$$

что дает систему линейных алгебраических уравнений

$$\sum_{j=1}^i c_j^{(i)} K(x_i, s_j) y_j = f(x_i), \quad i = 1, 2, \dots, n.$$

Для того чтобы решение этой системы существовало необходимо потребовать выполнения условия $K(x, x) \neq 0$.

При численном решении интегральных уравнений часто полезно провести предварительное преобразование исходной задачи. Типичным примером является приведение интегрального уравнения Вольтерра первого рода к интегральному уравнению второго рода. Будем считать, что ядро и правая часть дифференцируемы и $K(x, x) \neq 0$. Тогда от уравнения (8.6) можно перейти к уравнению

$$u(x) + \int_a^x \frac{1}{K(x, x)} \frac{\partial K(x, s)}{\partial x} u(s) ds = \frac{1}{K(x, x)} \frac{df}{dx}(x),$$

которое представляет собой интегральное уравнение Вольтерра второго рода.

8.2.3. Интегральное уравнение Фредгольма первого рода

Интегральное уравнение (8.4) есть наиболее характерный пример некорректно поставленной задачи. Некорректность обусловлена тем, что при малых возмущениях правой части $f(x)$ не гарантируется малого возмущения решения.

Помимо (8.4) рассмотрим уравнение с возмущенной правой частью

$$\int_a^b K(x, s) \tilde{u}(s) ds = \tilde{f}(x), \quad x \in [a, b]. \quad (8.13)$$

Ядро $K(x, s)$ есть вещественная непрерывная функция двух аргументов, а $\tilde{f}(x), f(x) \in L_2(a, b)$, причем

$$\|\tilde{f}(x) - f(x)\| \leq \delta,$$

при использовании обозначений

$$\|u(x)\| = \sqrt{(u, u)}, \quad (u, v) = \int_a^b u(x)v(x) dx.$$

При $\delta \rightarrow 0$ норма погрешности решения $\|\tilde{u}(x) - u(x)\|$ не стремится к нулю.

Определим линейный интегральный оператор

$$Ay = \int_a^b K(x, s) y(s) ds, \quad x \in [a, b]. \quad (8.14)$$

Задачу с неточной правой частью (8.13) запишем в виде операторного уравнения первого рода

$$A\tilde{u} = \tilde{f}. \quad (8.15)$$

В методе регуляризации Тихонова приближенное решение задачи (8.15) находится из минимума сглаживающего функционала:

$$J_\alpha(y) \rightarrow \min, \quad y \in L_2(a, b), \quad (8.16)$$

где

$$J_\alpha(y) = \|Ay - \tilde{f}\|^2 + \alpha\|y\|^2,$$

а $\alpha > 0$ — параметр регуляризации.

Обозначим решение задачи (8.16) через y_α . Оно может быть найдено как решение уравнения Эйлера для вариационной задачи (8.16)

$$\alpha y_\alpha + A^* A y_\alpha = A^* \tilde{f},$$

где

$$A^* y = \int_a^b K(s, x) y(s) ds, \quad x \in [a, b].$$

Тем самым приходим к интегральному уравнению Фредгольма

$$\alpha y_\alpha + \int_a^b G(x, s) y_\alpha(s) ds = \psi(x), \quad x \in [a, b]$$

с симметричным ядром

$$G(x, s) = \int_a^b K(t, x)K(t, s) dt$$

и правой частью

$$\psi(x) = \int_a^b K(s, x)f(s) ds.$$

Принципиальный момент в методе регуляризации связан с выбором параметра регуляризации α , его согласованием с погрешностью входных данных. При использовании принципа невязки параметр регуляризации выбирается из условия

$$\|Ay_\alpha - \bar{f}\| = \delta.$$

При таком выборе $\alpha = \alpha(\delta)$ норма погрешности $\|y_\alpha - u\| \rightarrow 0$ при $\delta \rightarrow 0$, т. е. приближенное решение стремится к точному решению задачи.

8.3. Упражнения

Рассмотрены иллюстративные примеры по построению и исследованию вычислительных алгоритмов приближенного решения интегральных уравнений первого и второго рода.

Упражнение 8.1. *Получите условия сходимости метода простой итерации*

$$u^{k+1}(x) = \lambda \int_a^b K(x, s)u^k(s) + f(x), \quad k = 0, 1, \dots \quad (8.17)$$

при, например, $u^0 = f(x)$ для приближенного решения интегрального уравнения (8.2).

Решение. Будем рассматривать сходимость итерационного процесса (8.17) в $L_2(a, b)$. Для погрешности из (8.2), (8.17) имеем

$$u^{k+1}(x) - u(x) = \lambda \int_a^b K(x, s)(u^k(s) - u(s)) ds, \quad k = 0, 1, \dots$$

Пусть существует ограниченная постоянная

$$\chi = \left(\int_a^b \int_a^b K^2(x, s) ds dx \right)^{1/2},$$

тогда с учетом неравенства

$$\left(\int_a^b K(x, s)(u^k(s) - u(s)) ds \right)^2 \leq \int_a^b K^2(x, s) ds \int_a^b (u^k(s) - u(s))^2 ds$$

получим

$$\|u^{k+1}(x) - u(x)\| \leq \lambda \chi \|u^k(x) - u(x)\|. \quad (8.18)$$

Поэтому для сходимости итерационного метода (8.17) достаточно, чтобы

$$\lambda < \frac{1}{\chi}.$$

При таких ограничениях из (8.18) стандартным образом следует оценка для числа итераций, необходимых для достижения искомой относительной точности приближенного решения.

Упражнение 8.2. Рассмотрите алгоритм решения интегрального уравнения (8.2) с вырожденным ядром

$$K(x, s) = \sum_{i=1}^n \psi_i(x) \varphi_i(s), \quad (8.19)$$

где $\psi_i(x), \varphi_i(s), i = 1, 2, \dots, n$ — системы линейно независимых функций.

Решение. Подстановка (8.19) в (8.2) дает

$$u(x) - \lambda \sum_{i=1}^n \psi_i(x) \int_a^b \varphi_i(s) u(s) ds = f(x). \quad (8.20)$$

Используя обозначения

$$\int_a^b \varphi_i(s) u(s) ds = c_i, \quad i = 1, 2, \dots, n,$$

получим следующее представление для решения интегрального уравнения Фредгольма второго рода с ядром (8.19):

$$u(x) = f(x) + \lambda \sum_{i=1}^n c_i \psi_i(x). \quad (8.21)$$

Для коэффициентов c_i , $i = 1, 2, \dots, n$ из (8.20) получим с учетом линейной независимости функций $\psi(x)$, $i = 1, 2, \dots, n$ систему линейных алгебраических уравнений

$$c_i - \lambda \sum_{j=1}^n c_j \int_a^b \psi_j(s) \varphi_i(s) ds = \int_a^b \varphi_i(s) f(s) ds.$$

После нахождения коэффициентов c_i , $i = 1, 2, \dots, n$ в качестве решения интегрального уравнения берется (8.21).

Упражнение 8.3. В методе квадратур находится приближенное решение в точках x_i , $i = 1, 2, \dots, n$. Рассмотрите возможности выполнения приближенного решения на все точки отрезка $[a, b]$ при приближенном решении интегрального уравнения (8.2).

Решение. Простейший подход связан с применением общих интерполяционных формул. Если в методе квадратур используется квадратурная формула интерполяционного типа, то та же интерполяционная формула используется и для восполнения. Например, естественно ориентироваться на применение кусочно-линейной интерполяции, когда по заданным y_i , $i = 1, 2, \dots, n$ строится функция $Y(x)$, такая что

$$Y(x) = \frac{x_{i+1} - x}{x_{i+1} - x_i} y_i + \frac{x - x_i}{x_{i+1} - x_i} y_{i+1},$$

$$x \in [x_i, x_{i+1}], \quad i = 1, 2, \dots, n-1.$$

Для восполнения решений интегрального уравнения можно применять интерполяционные формулы специального типа. Применяя для интегрального члена интегрального уравнения квадратурную формулу, для всех точек отрезка $[a, b]$ получим

$$y(x) = \lambda \sum_{j=1}^n c_j K(x, s_j) y_j + f(x), \quad j = 1, 2, \dots, n.$$

Такая интерполирующая функция более точно учитывает и передает специфику решаемой задачи.

Упражнение 8.4. Получите интегральное уравнение для определения производной порядка m заданной функции $f(x)$.

Решение. Без ограничения общности можем считать, что

$$f(0) = 0, \quad \frac{d^k f}{dx^k}(0) = 0, \quad k = 1, 2, \dots, m-1.$$

Для первой производной имеем

$$\frac{df}{dx}(x) = u(x)$$

и поэтому

$$\int_0^x u(s) ds = f(x).$$

Для производной m -ого порядка имеем

$$\int_0^x \frac{1}{(m-1)!} (x-s)^{m-1} u(s) ds = f(x).$$

Справедливость такого представления устанавливается методом математической индукции. Тем самым мы имеем интегральное уравнение Вольтерра первого рода.

Упражнение 8.5. Приведите расчетные формулы метода квадратур при использовании формулы трапеций для приближенного решения интегрального уравнения Вольтерра второго рода (8.5).

Решение. Для узлов

$$a = x_1 < x_2 < \dots < x_n = b$$

при использовании квадратурной формулы трапеций из (8.11) получим

$$y_i - \lambda \sum_{j=2}^i h_j \frac{K_{ij} y_j + K_{i,j-1} y_{j-1}}{2} = f(x_i), \quad i = 2, 3, \dots, n,$$

где использовались обозначения

$$K_{ij} = K(x_i, s_j), \quad h_i = x_i - x_{i-1}.$$

Отсюда с учетом

$$y_1 = f(x_1)$$

получим рекуррентные формулы для последовательного определения приближенного решения

$$y_i = \frac{2f(x_i) + h_2 K_{i1} y_1 + \sum_{j=2}^{i-1} (h_i + h_{i+1}) K_{ij} y_j}{2 - h_i K_{ii}}$$

для всех $i = 2, 3, \dots, n$. Эти формулы можно использовать при $K_{ii} h_i \neq 2$.

Упражнение 8.6. Пусть норма точного решения операторного уравнения

$$Au = f \tag{8.22}$$

с оператором $A = A^* > 0$ в $L_2(a, b)$ ограничена в $H_{A^{-2}}$, т. е.

$$\|A^{-1}u\| \leq M = \text{const.} \tag{8.23}$$

При решении задачи с приближенной правой частью f_δ такой, что

$$\|f_\delta - f\| \leq \delta, \tag{8.24}$$

используется алгоритм упрощенной регуляризации

$$\alpha y_\alpha + Ay_\alpha = f_\delta. \tag{8.25}$$

Покажите сходимость приближенного решения к точному при выборе параметра регуляризации $\alpha(\delta) = \sqrt{\delta/M}$.

Решение. Необходимо установить $y_\alpha \rightarrow u$ при $\delta \rightarrow 0$. Обозначим через $\lambda_k, \varphi_k, k = 1, 2, \dots$ — собственные значения и собственные функции оператора A соответственно, причем $\lambda_k > 0, k = 1, 2, \dots$. Для точного решения задачи имеем представление

$$u = \sum_{k=1}^{\infty} \frac{(f, \varphi_k)}{\lambda_k} \varphi_k.$$

Аналогичное представление для приближенного решения получим из (8.25):

$$y_\alpha = \sum_{k=1}^{\infty} \frac{(f_\delta, \varphi_k)}{\alpha + \lambda_k} \varphi_k.$$

Пусть y — решение регуляризованной задачи с точной правой частью, т. е.

$$\alpha y + Ay = f. \quad (8.26)$$

Принимая во внимание

$$\|y_\alpha - u\| \leq \|y_\alpha - y\| + \|y - u\|, \quad (8.27)$$

рассмотрим отдельно близость y_α к y и y к u . Равенство Парсеваля с учетом представления решения задач (8.25), (8.26) и оценки (8.24) дает

$$\|y_\alpha - y\|^2 = \sum_{k=1}^{\infty} \frac{(f_\delta - f, \varphi_k)^2}{(\alpha + \lambda_k)^2} \leq \frac{\delta^2}{\alpha^2}.$$

Аналогично с учетом априорных ограничений (8.23) получим

$$\|y - u\|^2 = \alpha^2 \sum_{k=1}^{\infty} \frac{(f, \varphi_k)^2}{(\alpha + \lambda_k)^2 \lambda_k^2} \leq \alpha^2 \sum_{k=1}^{\infty} \frac{(f, \varphi_k)^2}{\lambda_k^4} \leq \alpha^2 M^2.$$

Подстановка в (8.27) дает оценку

$$\|y_\alpha - u\| \leq \frac{\delta}{\alpha} + \alpha M.$$

Минимум правой части достигается при выборе параметра регуляризации $\alpha(\delta) = \sqrt{\delta/M}$, причем

$$\|y_\alpha - u\| \leq 2\sqrt{M\delta}$$

и поэтому $\|y_\alpha - u\| \rightarrow 0$ при $\delta \rightarrow 0$.

8.4. Задачи

Задача 8.1. Получите оценку погрешности метода квадратур Симпсона при решении интегрального уравнения Фредгольма второго рода в равномерной норме.

Задача 8.2. Рассматривается задача нахождения характеристических чисел (собственных значений) интегральных уравнений, т. е.

$$u(x) = \lambda \int_a^b K(x, s)u(s) ds.$$

Ядро интегрального уравнения симметрично и положительно определено, т. е.

$$K(x, s) = K(s, x),$$

$$\int_a^b \int_a^b K(x, s)u(x)u(s) dx ds > 0, \quad \forall u(x) \neq 0.$$

В методе Келлога по заданной функции $\omega^0(x)$ строятся функции

$$\omega^{k+1}(x) = \int_a^b K(x, s)\omega^k(s) ds, \quad k = 0, 1, \dots$$

Покажите, что для наименьшего характеристического числа имеют место приближенные формулы

$$\lambda_1 \approx \frac{\|\omega^{k-1}\|}{\|\omega^k\|}, \quad \lambda_1 \approx \frac{1}{\sqrt[k]{\|\omega^k\|}}.$$

Задача 8.3. Определим k -ое итерированное ядро с помощью рекуррентных соотношений

$$K_1(x, s) = K(x, s),$$

$$K_k(x, s) = \int_a^b K(x, t)K_{k-1}(t, s) dt, \quad k = 2, 3, \dots$$

Число

$$A_k = \int_a^b K_k(x, x) dx$$

называется k -ым следом ядра $K(x, s)$. Покажите, что при достаточно больших k для наименьшего характеристического числа справедливо приближенное выражение (метод следов)

$$|\lambda_k| \approx \sqrt{\frac{A_{2k}}{A_{2k+1}}}.$$

Задача 8.4. Рассмотрите алгоритм решения интегрального уравнения Фредгольма второго рода (8.2) на основе аппроксимации приближенного решения кубическим сплайном.

Задача 8.5. Рассмотрите возможности аппроксимации ядра вырожденным ядром

$$K(x, s) \approx \sum_{i=1}^n \psi_i(x) \varphi_i(s)$$

на основе разложения в ряд Тейлора по одной и двум переменным и построения интерполяционного полинома для функции одной и двух переменных.

Задача 8.6. Получите оценку погрешности в $L_2(a, b)$ приближенного решения интегрального уравнения Фредгольма второго рода при замене ядра вырожденным.

Задача 8.7. Исследуйте сходимость метода простой итерации (8.7) в равномерной норме (в $C[a, b]$).

Задача 8.8. Получите условия сходимости итерационного метода (8.17) для приближенного решения интегрального уравнения (8.2), ядро которого имеет слабую особенность:

$$K(x, s) = \frac{G(x, s)}{|x - s|^\gamma}, \quad |G(x, s)| \leq M = \text{const}, \quad 0 < \gamma < 1.$$

Задача 8.9. Для приближенного решения интегрального уравнения Фредгольма второго рода (8.2) используйте итерационный метод аналогичный методу Зейделя для решения систем линейных алгебраических уравнений. Установите условия сходимости такого метода.

Задача 8.10. Рассмотрите итерационный метод Ньютона для приближенного решения нелинейного интегрального уравнения (уравнения Урысона)

$$u(x) - \int_a^b K(x, s, u(s)) ds = f(x), \quad x \in [a, b].$$

Задача 8.11. Постройте алгоритм приближенного решения интегрального уравнения Вольтерра первого рода (8.6) методом квадратур и рассмотрите условия его применимости.

Задача 8.12. Постройте иллюстративный пример показывающий неустойчивость решения интегрального уравнения Фредгольма первого рода (8.4) к малым (в $L_2(a, b)$) возмущениям правой части.

Задача 8.13. Примените метод регуляризации Тихонова для решения интегрального уравнения Вольтерра первого рода (8.6).

Задача 8.14. Получите уравнение Эйлера в методе регуляризации Тихонова (8.16) для решения задачи (8.14), (8.15) при условии, что сглаживающий функционал имеет вид

$$J_\alpha(y) = \|Ay - \tilde{f}\|^2 + \alpha \left(\|y\|^2 + q \left\| \frac{dy}{dx} \right\|^2 \right).$$

Задача 8.15. Пусть в (8.4) ядро $K(x, s)$ симметрично и положительно определено. Покажите, что итерационный метод

$$\frac{u^{k+1}(x) - u^k(x)}{\tau} + \int_a^b K(x, s) u^k(s) ds = f(x)$$

сходится при

$$0 < \tau < 2\lambda_1,$$

где λ_1 — наименьшее характеристическое число.

Задача 8.16. Покажите, что в условиях предыдущей задачи при решении интегрального уравнения (8.4) с неточной правой частью можно согласовать критерий останова (число итераций) с погрешностью правой части δ для того, чтобы получить приближенное решение, сходящееся при $\delta \rightarrow 0$ к точному.

Глава 9

Задача Коши для обыкновенных дифференциальных уравнений

В вычислительной практике часто приходится иметь дело с задачами с начальными данными для системы дифференциальных уравнений. Для приближенного решения таких задач традиционно широко используются методы Рунге—Кутты, связанные с вычислением правой части системы уравнений в некоторых промежуточных точках. Второй большой класс методов составляют многошаговые методы, когда в вычислениях участвуют три и более расчетных слоев. Отдельно выделяются задачи, для которых решение имеет разномаштабные гармоники (жесткие системы обыкновенных дифференциальных уравнений).

9.1. Задачи с начальными условиями для систем обыкновенных дифференциальных уравнений

Рассматривается задача Коши для системы обыкновенных дифференциальных уравнений

$$\frac{du_i(t)}{dt} = f_i(t, u_1, u_2, \dots, u_m), \quad t > 0, \quad (9.1)$$

$$u_i(0) = u_i^0, \quad i = 1, 2, \dots, m. \quad (9.2)$$

Основные обозначения

- $u = \{u_1, u_2, \dots, u_m\}$ — вектор неизвестных
 $f = \{f_1, f_2, \dots, f_m\}$ — вектор правых частей
 $\omega_\tau = \{t_n = n\tau, n = 0, 1, \dots\}$ — равномерная сетка по t
 $\tau > 0$ — шаг сетки
 $y^n = y(t_n)$ — приближенное решение при $t = t_n$

С использованием векторных обозначений задачу (9.1), (9.2) можем записать как задачу Коши для одного уравнения:

$$\frac{du(t)}{dt} = f(t, u), \quad t > 0, \quad (9.3)$$

$$u(0) = u^0. \quad (9.4)$$

В задаче Коши по известному решению в точке $t = 0$ необходимо найти из уравнения (9.4) решение при других t .

9.2. Численные методы решения задачи Коши

Отмечаются классические методы Рунге—Кутты и многошаговые методы решения задачи Коши для систем обыкновенных дифференциальных уравнений, обсуждается специфика численного решения жестких систем.

9.2.1. Методы Рунге—Кутты

При построении численных алгоритмов будем считать, что решение этой дифференциальной задачи существует, оно единственно и обладает необходимыми свойствами гладкости.

При численном решении задачи (9.3), (9.4) будем использовать равномерную, для простоты, сетку по переменной t с шагом $\tau > 0$:

$$\omega_\tau = \{t_n = n\tau, n = 0, 1, \dots\}.$$

Приближенное решение задачи (9.3), (9.4) в точке $t = t_n$ обозначим y^n . Метод сходится в точке t_n , если $|y^n - u(t_n)| \rightarrow 0$ при $\tau \rightarrow 0$. Метод имеет p -ый порядок точности, если $|y^n - u(t_n)| = O(\tau^p)$, $p > 0$ при $\tau \rightarrow 0$.

Простейшая разностная схема для приближенного решения задачи (9.3), (9.4) есть

$$\frac{y^{n+1} - y^n}{\tau} = \sigma f(t_{n+1}, y^{n+1}) + (1 - \sigma) f(t_n, y^n), \quad n = 0, 1, \dots \quad (9.5)$$

При $\sigma = 0$ имеем явный метод Эйлера и в этом случае разностная схема аппроксимирует уравнение (9.4) с первым порядком.

Симметричная схема ($\sigma = 0,5$ в (9.5)) имеет второй порядок аппроксимации. Эта схема относится к классу неявных — для определения приближенного решения на новом слое необходимо решать нелинейную задачу. Явные схемы второго и более высокого порядка аппроксимации удобно строить, ориентируясь на метод предиктор-корректор. На этапе предиктора (предсказания) используется явная схема

$$\frac{\bar{y}_{n+1} - y_n}{\tau} = f(t_n, y_n),$$

а на этапе корректора (уточнения) — схема

$$\frac{y^{n+1} - y^n}{\tau} = \frac{1}{2} (f(t_{n+1}, \bar{y}^{n+1}) + f(t_n, y^n)), \quad n = 0, 1, \dots$$

В одношаговых методах Рунге—Кутта идеи предиктора-корректора реализуются наиболее полно. Этот метод записывается в общем виде

$$\frac{y^{n+1} - y^n}{\tau} = \sum_{i=1}^s b_i k_i, \quad (9.6)$$

где

$$k_i = f \left(t_n + c_i \tau, y^n + \tau \sum_{j=1}^s a_{ij} k_j \right), \quad i = 1, 2, \dots, s. \quad (9.7)$$

Формула (9.6) основана на s вычислениях функции f и называется s -стадийной. Если $a_{ij} = 0$ при $j \geq i$ имеем явный метод Рунге—Кутта. Если $a_{ij} = 0$ при $j > i$ и $a_{ii} \neq 0$, то k_i определяется неявно из уравнения

$$k_i = f \left(t_n + c_i \tau, y^n + \tau \sum_{j=1}^{i-1} a_{ij} k_j + \tau a_{ii} k_i \right).$$

О таком методе Рунге—Кутта говорят как о диагонально-неявном.

Параметры b_i, c_i, a_{ij} определяют вариант метода Рунге—Кутты. Используется следующее представление метода (таблица Бутчера):

$$\frac{c}{b^s} \left| \begin{array}{c|cccc} A & a_{11} & a_{12} & \cdots & a_{1s} \\ & a_{21} & a_{22} & \cdots & a_{2s} \\ & \vdots & \vdots & \ddots & \vdots \\ & a_{s1} & a_{s2} & \cdots & a_{ss} \\ \hline b^s & b_1 & b_2 & \cdots & b_s \end{array} \right. \quad (9.8)$$

Одним из наиболее распространенных является явный метод Рунге—Кутта четвертого порядка:

$$\begin{aligned} k_1 &= f(t_n, y^n), & k_2 &= f\left(t_n + \frac{\tau}{2}, y^n + \tau \frac{k_1}{2}\right), \\ k_3 &= f\left(t_n + \frac{\tau}{2}, y^n + \tau \frac{k_2}{2}\right), & k_4 &= f(t_n + \tau, y^n + \tau k_3), \\ \frac{y^{n+1} - y^n}{\tau} &= \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4). \end{aligned}$$

В компактном представлении (9.8) этого метода имеем

$$\frac{c}{b^s} \left| \begin{array}{c|cccc} A & 0 & 0 & 0 & 0 \\ & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ & 1 & 0 & 0 & 1 & 0 \\ \hline b^s & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{array} \right. .$$

Применяя метод Рунге—Кутта (9.6), (9.7) к решению задачи Коши для уравнения

$$\frac{du(t)}{dt} = f(t), \quad t > 0,$$

получим

$$y^{n+1} - y^n = \sum_{i=1}^s \tau b_i f(t_n + c_i \tau).$$

Правую часть можно рассматривать как квадратурную формулу для правой части равенства

$$u(t_{n+1}) - u(t_n) = \int_{t_n}^{t_{n+1}} f(t) dt.$$

Исследование устойчивости используемых разностных схем при решении задачи Коши для систем обыкновенных дифференциальных уравнений проводится чаще всего на модельном одномерном уравнении

$$\frac{du(t)}{dt} = \lambda u, \quad t > 0, \quad (9.9)$$

где λ — комплексное число. Для конкретного численного метода рассматривается множество всех точек комплексной плоскости $\mu = \tau\lambda$, для которых имеет место устойчивость. Для явного метода Эйлера область устойчивости представляет круг единичного радиуса с центром в точке $(-1, 0)$. Метод называется A -устойчивым, если область его устойчивости содержит полуплоскость $\operatorname{Re} \mu < 0$. При $\operatorname{Re} \lambda < 0$ устойчиво решение уравнения (9.9) и поэтому для этой задачи условие A -устойчивости означает абсолютную устойчивость (устойчивость при всех $\tau > 0$).

9.2.2. Многошаговые методы

В методах Рунге—Кутты в вычислениях участвуют значения приближенного решения только в двух соседних узлах y^n и y^{n+1} — один шаг по переменной t . Линейный m -шаговый разностный метод записывается в виде

$$\frac{1}{\tau} \sum_{i=0}^m a_i y^{n+1-i} = \sum_{i=0}^m b_i f(t_{n+1-i}, y^{n+1-i}), \quad n = m-1, m, \dots \quad (9.10)$$

Вариант численного метода определяется заданием коэффициентов a_i, b_i , $i = 0, 1, \dots, m$, причем $a_0 \neq 0$. Для начала расчетов по рекуррентной формуле (9.10) необходимо задать m начальных значений y^0, y^1, \dots, y^{m-1} .

Различные варианты многошаговых методов (методы Адамса) решения задачи с начальными условиями для систем обыкновенных дифференциальных уравнений могут быть получены на основе использования

квадратурных формул для правой части равенства

$$u(t_{n+1}) - u(t_n) = \int_{t_n}^{t_{n+1}} f(t, u) dt. \quad (9.11)$$

Для получения неявного многошагового метода используем для подынтегральной функции интерполяционную формулу по значениям функции $f^{n+1} = f(t_{n+1}, y^{n+1}), f^n, \dots, f^{n+1-m}$, т. е.

$$\frac{y^{n+1} - y^n}{\tau} = \sum_{i=0}^m b_i f(t_{n+1-i}, y^{n+1-i}). \quad (9.12)$$

Для интерполяционного метода Адамса (9.12) наивысший порядок аппроксимации равен $m + 1$.

Для построения явных многошаговых методов можно использовать процедуру экстраполяции подынтегральной функции в правой части (9.11). В этом случае приближение осуществляется по значениям $f^n, f^{n-1}, \dots, f^{n+1-m}$ и поэтому

$$\frac{y^{n+1} - y^n}{\tau} = \sum_{i=1}^m b_i f(t_{n+1-i}, y^{n+1-i}). \quad (9.13)$$

Для экстраполяционного метода Адамса (9.13) погрешность аппроксимации имеет m -ый порядок.

Примерами методов Адамса (9.12), (9.13) при $m = 3$ являются

$$\frac{y^{n+1} - y^n}{\tau} = \frac{1}{24} (9f^{n+1} + 19f^n - 5f^{n-1} + f^{n-2}), \quad (9.14)$$

$$\frac{y^{n+1} - y^n}{\tau} = \frac{1}{12} (23f^n - 16f^{n-1} + 5f^{n-2}) \quad (9.15)$$

соответственно.

На основе методов Адамса строятся и схемы предиктор-корректор. На этапе предиктора используется явный метод Адамса, на этапе корректора — аналог неявного метода Адамса. Например, при использовании методов третьего порядка аппроксимации в соответствии с (9.15) для предсказания решения положим

$$\bar{y}^{n+1} - y^n = \frac{1}{12} (23f^n - 16f^{n-1} + 5f^{n-2}).$$

Для уточнения решения (см. (9.14)) используется схема

$$\frac{y^{n+1} - y^n}{\tau} = \frac{1}{24} (9f(t_{n+1}, \bar{y}^{n+1}) + 19f^n - 5f^{n-1} + f^{n-2}).$$

Аналогично строятся и другие классы многошаговых методов.

9.2.3. Жесткие системы обыкновенных дифференциальных уравнений

При численном решении задачи Коши для систем обыкновенных дифференциальных уравнений (9.1), (9.2) могут возникнуть дополнительные трудности, порожденные жесткостью системы. Локальные особенности поведения решения в точке $u = w$ передаются линейной системой

$$\frac{dv_i(t)}{dt} = \sum_{j=1}^m \frac{\partial f_i}{\partial u_j}(t, w)v_j + \bar{f}_i(t), \quad t > 0.$$

Пусть $\lambda_i(t)$, $i = 1, 2, \dots, m$ — собственные числа матрицы

$$A(t) = \{a_{ij}(t)\}, \quad a_{ij}(t) = \frac{\partial f_i}{\partial u_j}(t, w).$$

Система уравнений (9.1) является жесткой, если число

$$S(t) = \frac{\max_{1 \leq i \leq m} |\operatorname{Re} \lambda_i(t)|}{\min_{1 \leq i \leq m} |\operatorname{Re} \lambda_i(t)|}$$

велико. Это означает, что в решении присутствуют составляющие с сильно различающимися масштабами изменения по переменной t .

Для численного решения жестких задач используются вычислительные алгоритмы, которые имеют повышенный запас устойчивости. Необходимо ориентироваться на использование A -устойчивых или $A(\alpha)$ -устойчивых методов.

Метод называется A -устойчивым, если при решении задачи Коши для уравнения (9.9) область его устойчивости содержит угол

$$|\arg(-\mu)| < \alpha, \quad \mu = \lambda\tau.$$

Среди A -устойчивых методов можно выделить чисто неявные многошаговые методы (методы Гира), когда

$$\frac{1}{\tau} \sum_{i=0}^m a_i y^{n+1-i} = f(t_{n+1}, y^{n+1}).$$

В частности, при $m = 3$ имеем схему

$$\frac{11y^{n+1} - 18y^n + 9y^{n-1} - 2y^{n-2}}{6\tau} = f(t_{n+1}, y^{n+1}),$$

которая имеет третий порядок аппроксимации.

9.3. Упражнения

В приведенных ниже примерах основное внимание уделено построению численных методов решения задачи Коши для обыкновенных дифференциальных уравнений и исследованию свойств аппроксимации и устойчивости.

Упражнение 9.1. *Покажите возможность построения метода приближенного решения задачи Коши (9.3), (9.4) на основе разложения решения в ряд Тейлора.*

Решение. Будем считать, что решение задачи (9.3), (9.4) при $m = 1$ и правая часть достаточно гладкие функции своих аргументов. Разлагая $u(t)$ при $t = t_n$, получим

$$u(t_{n+1}) = u(t_n) + (t_{n+1} - t_n) \frac{du}{dt}(t_n) + \\ + \frac{(t_{n+1} - t_n)^2}{2} \frac{d^2u}{dt^2}(t_n) + \frac{(t_{n+1} - t_n)^3}{3!} \frac{d^3u}{dt^3}(t_n) + \dots$$

Для нахождения производных решения используется уравнение (9.3), так что, например,

$$\frac{du}{dt}(t) = f(t, u), \\ \frac{d^2u}{dt^2}(t) = \frac{\partial f}{\partial t}(t, u) + \frac{\partial f}{\partial u}(t, u) f(t, u),$$

$$\begin{aligned} \frac{d^3 u}{dt^3}(t) &= \frac{\partial^2 f}{\partial t^2}(t, u) + 2 \frac{\partial^2 f}{\partial t \partial u}(t, u) f(t, u) + \\ &+ \frac{\partial^2 f}{\partial t^2}(t, u) f^2(t, u) + \frac{\partial f}{\partial u}(t, u) \frac{d^2 u}{dt^2}(t). \end{aligned}$$

Ограничиваясь несколькими первыми членами разложения для приближенного решения задачи (9.3), (9.4) получим

$$\begin{aligned} \frac{y^{n+1} - y^n}{\tau} &= f(t_n, y^n), \quad n = 0, 1, \dots, \\ \frac{y^{n+1} - y^n}{\tau} &= f(t_n, y^n) + \frac{t_{n+1} - t_n}{2} \left(\frac{\partial f}{\partial t}(t_n, y^n) + \frac{\partial f}{\partial u}(t_n, y^n) f(t_n, y^n) \right), \\ n &= 0, 1, \dots \end{aligned}$$

Для получения производных правой части по переменным t и u можно использовать современные компьютерные системы аналитических вычислений.

Упражнение 9.2. Постройте явный метод Рунге–Кутты при $s = 2$ и дополнительном ограничении $c_1 = 0$.

Решение. В силу общего представления (9.6), (9.7) в рассматриваемом случае имеем

$$\frac{y^{n+1} - y^n}{\tau} = b_1 k_1 + b_2 k_2, \quad (9.16)$$

$$k_1 = f(t_n, y^n), \quad k_2 = f(t_n + c_2 \tau, y_n + a_{21} \tau k_1). \quad (9.17)$$

Параметры метода a_{21}, b_1, b_2, c_2 найдем из условия, чтобы погрешность аппроксимации имела наибольший порядок.

Для погрешности аппроксимации из (9.16), (9.17) получим

$$\psi_n = -\frac{u^{n+1} - u^n}{\tau} + b_1 f(t_n, u^n) + b_2 f(t_n + c_2 \tau, u^n + a_{21} \tau f(t_n, u^n)).$$

Разложение в ряд Тейлора дает

$$\begin{aligned} \frac{u^{n+1} - u^n}{\tau} &= \frac{du}{dt}(t_n) + \frac{\tau}{2} \frac{d^2 u}{dt^2}(t_n) + O(\tau^2), \\ f(t_n + c_2 \tau, u^n + a_{21} \tau f(t_n, u^n)) &= \\ &= f(t_n, u^n) + c_2 \tau \frac{\partial f}{\partial t}(t_n, u^n) + a_{21} \tau f(t_n, u^n) \frac{\partial f}{\partial u}(t_n, u^n) + O(\tau^2). \end{aligned}$$

На решениях уравнения (9.3)

$$\frac{d^2 u}{dt^2}(t) = \frac{\partial f}{\partial t}(t, u) + \frac{\partial f}{\partial u}(t, u)f(t, u)$$

и поэтому

$$\begin{aligned} \psi_n = & -\frac{du}{dt}(t_n) + (b_1 + b_2)f(t_n, u^n) + \\ & + \tau \left(b_2 a_{21} - \frac{1}{2} \right) f(t_n, u^n) \frac{\partial f}{\partial u}(t_n, u^n) + \\ & + \tau \left(b_2 c_2 - \frac{1}{2} \right) \frac{\partial f}{\partial t}(t_n, u^n) + O(\tau^2). \end{aligned}$$

Следовательно, метод (9.16), (9.17) имеет первый порядок аппроксимации, если $b_1 + b_2 = 1$. При

$$b_2 a_{21} = \frac{1}{2}, \quad b_2 c_2 = \frac{1}{2}$$

имеем однопараметрическое семейство методов Рунге—Кутты второго порядка аппроксимации

$$\frac{y^{n+1} - y^n}{\tau} = (1 - \sigma) f(t_n, u^n) + \sigma f(t_n + c\tau, u^n + c\tau f(t_n, u^n)), \quad (9.18)$$

где $b_2 = 1 - b_1 = \sigma$, $c_2 = a_{21} = c$, причем $c\sigma = 0,5$.

В классе (9.18) наиболее известными методами являются методы с $\sigma = 1$ и $\sigma = 0,5$.

Упражнение 9.3. Получите условия устойчивости явного метода Рунге—Кутты четвертого порядка точности

$$\begin{aligned} \frac{y^{n+1} - y^n}{\tau} &= \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \\ k_1 &= f(t_n, y^n), \quad k_2 = f\left(t_n + \frac{\tau}{2}, y^n + \tau \frac{k_1}{2}\right), \\ k_3 &= f\left(t_n + \frac{\tau}{2}, y^n + \tau \frac{k_2}{2}\right), \quad k_4 = f(t_n + \tau, y^n + \tau k_3). \end{aligned}$$

Решение. Для модельного уравнения с правой частью $f(t, u) = -\lambda u$ имеем

$$y^{n+1} = q(\mu)y^n, \quad \mu = \lambda\tau,$$

где

$$q(\mu) = 1 - \mu + \frac{1}{2} \mu^2 - \frac{1}{6} \mu^3 + \frac{1}{24} \mu^4.$$

Метод устойчив, если выполнено $|q| \leq 1$. Это неравенство (см. рис. 9.1) справедливо при

$$\mu = \lambda\tau \leq 2,78.$$

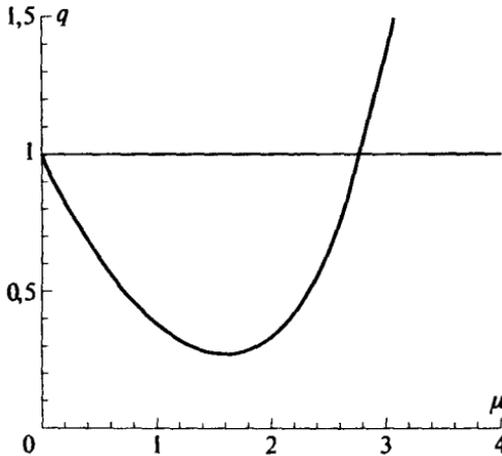


Рис. 9.1. График функции $q(\mu)$

Упражнение 9.4. Постройте трехслойный метод решения задачи Коши (9.3), (9.4) с использованием квадратурной формулы Симпсона.

Решение. Интегрирование уравнения (9.3) по t от t_{n-1} до t_{n+1} дает

$$\frac{u^{n+1} - u^{n-1}}{2\tau} = \frac{1}{2\tau} \int_{t_{n-1}}^{t_{n+1}} f(t, u) dt.$$

На правой части использование формулы Симпсона дает

$$\begin{aligned} \frac{1}{2\tau} \int_{t_{n-1}}^{t_{n+1}} f(t, u) dt &= \\ &= \frac{1}{6} (f(t_{n+1}, u^{n+1}) + 4f(t_n, u^n) + f(t_{n-1}, u^{n-1})) + O(\tau^4). \end{aligned}$$

Это приводит нас к методу

$$\frac{y^{n+1} - y^{n-1}}{2\tau} = \frac{1}{6}(f(t_{n+1}, y^{n+1}) + 4f(t_n, y^n) + f(t_{n-1}, y^{n-1})),$$

который имеет четвертый порядок аппроксимации.

Упражнение 9.5. Приведите общую схему построения m -шагового метода Гира и исследуйте погрешность аппроксимации.

Решение. В данном классе методов в уравнении

$$\frac{du}{dt}(t_{n+1}) = f(t_{n+1}, u(t_{n+1}))$$

используются аппроксимации производной с помощью направленных разностных производных. Для m -шагового метода аппроксимации строится по узлам t^{n+1-i} , $i = 0, 1, \dots, m$, т. е. используется метод

$$\frac{1}{\tau} \sum_{i=0}^m a_i y^{n+1-i} = f(t_{n+1}, y^{n+1}).$$

Для определения коэффициентов a_i , $i = 0, 1, \dots, m$ можно применять метод неопределенных коэффициентов, когда используются разложения функций y^{n+1-i} , $i = 0, 1, \dots, m$ в ряд Тейлора в точке $t = t_{n+1}$. Более конструктивным представляется подход с использованием формул численного дифференцирования, построенными с привлечением обычных интерполяционных формул.

По точкам (t_{n+1-i}, y^{n+1-i}) , $i = 0, 1, \dots, m$ построим интерполяционный полином $L_m(t)$ степени m и используем для численного решения уравнения (9.3) схему

$$\frac{dL_m}{dt}(t_{n+1}) = f(t_{n+1}, y^{n+1}).$$

При использовании интерполяционной формулы Ньютона получим

$$L_m(t) = y(t_{n+1}) + (t - t_{n+1})y(t_{n+1}, t_n) + \\ + (t - t_{n+1})(t - t_n)y(t_{n+1}, t_n, t_{n-1}) + \dots,$$

где $y(t_{n+1}, t_n)$, $y(t_{n+1}, t_n, t_{n-1})$, \dots — разделенные разности соответствующего порядка. Тем самым

$$\frac{dL_m}{dt}(t_{n+1}) = y(t_{n+1}, t_n) + (t_{n+1} - t_n)y(t_{n+1}, t_n, t_{n-1}) + \dots$$

Например, при $m = 2$ получим

$$\frac{dL_2}{dt}(t_{n+1}) = \frac{1}{2\tau}(3y^{n+1} - 4y^n + y^{n-1})$$

при использовании равномерной сетки с шагом τ .

Упражнение 9.6. Для задачи Коши

$$\begin{aligned} \frac{d^2 u}{dt^2} &= f(t, u), \quad t > 0, \\ u(0) &= u^0, \quad \frac{du}{dt}(0) = v^0 \end{aligned} \quad (9.19)$$

используются методы Штермера, когда

$$\frac{y^{n+1} - 2y^n + y^{n-1}}{\tau^2} = \sum_{i=0}^s b_i f(t_{n+1-i}, y^{n+1-i}). \quad (9.20)$$

Укажите возможность построения таких схем на основе построения специальных квадратурных формул для правой части уравнения (9.19).

Решение. Домножим уравнение (9.19) на финитную функцию

$$\varphi(t) = \begin{cases} 1 - \frac{|t - t_n|}{\tau}, & t \in [t_{n-1}, t_{n+1}], \\ 0, & t \notin [t_{n-1}, t_{n+1}]. \end{cases}$$

Непосредственные выкладки дают

$$\int_{t_{n-1}}^{t_{n+1}} \frac{d^2 u}{dt^2} \varphi(t) dt = y^{n+1} - 2y^n + y^{n-1},$$

поэтому для уравнения (9.19) имеет место равенство

$$\frac{u^{n+1} - 2u^n + u^{n-1}}{\tau^2} = \int_{t_{n-1}}^{t_{n+1}} f(t, u) \varphi(t) dt. \quad (9.21)$$

Формулы типа (9.20) мы можем получить на основе использования тех или иных (интерполяционных и экстраполяционных) квадратурных формул для правой части (9.21).

9.4. Задачи

Задача 9.1. Покажите, что в классе двухстадийных явных методов Рунге—Кутта (9.18) нет методов третьего порядка аппроксимации.

Задача 9.2. На примере системы двух обыкновенных уравнений рассмотрите особенности построения методов Рунге—Кутта для систем.

Задача 9.3. Получите оценки погрешности при использовании метода Рунге—Кутта для приближенного решения задачи Коши (9.3), (9.4) при условии, что функция $f(t, u)$ удовлетворяет условию Липшица по второму аргументу.

Задача 9.4. Рассмотрите метод Рунге—Кутта

$$\frac{c}{b^s} \Big| A = \begin{array}{c|cc} \frac{3 - \sqrt{3}}{6} & \frac{1}{4} & \frac{3 - 2\sqrt{3}}{12} \\ \hline \frac{3 + \sqrt{3}}{6} & \frac{3 + 2\sqrt{3}}{12} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array},$$

связав его с квадратурной формулой Гаусса.

Задача 9.5. Постройте трехстадийный явный метод Рунге—Кутта ($s = 3$ (9.6), (9.7)) при $c_1 = 0$.

Задача 9.6. Методом неопределенных коэффициентов постройте явную двухшаговую схему третьего порядка аппроксимации.

Задача 9.7. Получите схемы Адамса (9.14), (9.15) и исследуйте погрешность аппроксимации.

Задача 9.8. Покажите, что наивысший достижимый порядок аппроксимации неявных многошаговых методов (9.10) равен $2m$, а явных — $2m - 1$.

Задача 9.9. Сформулируйте условия сходимости многошагового метода (9.10) при решении задачи (9.3), (9.4), когда функция $f(t, u)$ удовлетворяет условию Липшица по второму аргументу.

Задача 9.10. Получите условия устойчивости неявной схемы Адамса третьего порядка точности

$$\frac{y^{n+1} - y^n}{\tau} = \frac{1}{12} (5f^{n+1} + 8f^n - f^{n-1}).$$

Задача 9.11. Докажите, что среди явных многошаговых методов ($b_0 = 0$ в (9.10)) нет A -устойчивых.

Задача 9.12. Докажите, что среди неявных многошаговых методов (9.10) нет A -устойчивых методов, имеющих порядок аппроксимации выше второго.

Задача 9.13. Получите условия $A(\alpha)$ -устойчивости метода Гира

$$\frac{25y^{n+1} - 48y^n + 36y^{n-1} - 16y^{n-2} + 3y^{n-3}}{12\tau} = f(t_{n+1}, y^{n+1}),$$

который имеет четвертый порядок точности.

Задача 9.14. На примере модельной задачи Коши для уравнения

$$\frac{du}{dt} = \lambda(t)u$$

покажите, что результаты по устойчивости, установленные для задачи с постоянными коэффициентами, не всегда верны для задач с переменными коэффициентами.

Задача 9.15. Исследуйте схему Штермера

$$\frac{y^{n+1} - 2y^n + y^{n-1}}{\tau^2} = \frac{1}{12} (f(t_{n+1}, y^{n+1}) + 10f(t_n, y^n) + f(t_{n-1}, y^{n-1}))$$

для приближенного решения задачи Коши для уравнения (9.19).

Задача 9.16. Рассмотрите схему Рунге—Кутта

$$\frac{y^{n+1} - y^n}{\tau} = \frac{\tau}{6} (k_1 + k_2 + k_3),$$

$$\frac{v^{n+1} - v^n}{\tau} = \frac{1}{6} (k_1 + 2k_2 + 2k_3 + k_4),$$

где

$$k_1 = f(t_n, y^n, v^n),$$

$$k_2 = f\left(t_n + \frac{\tau}{2}, y^n + \frac{\tau}{2}v^n, v^n + \frac{1}{2}k_1\right),$$

$$k_3 = f\left(t_n + \frac{\tau}{2}, y^n + \frac{\tau}{2}v^n + \frac{\tau}{4}k_1, v^n + \frac{1}{2}k_2\right),$$

$$k_4 = f\left(t_n + \tau, y^n + \tau v^n + \frac{\tau}{2}k_2, v^n + k_3\right),$$

для решения задачи Коши для уравнения второго порядка

$$\frac{d^2u}{dt^2} = f\left(t, u, \frac{du}{dt}\right), \quad t > 0,$$

$$u(0) = u^0, \quad \frac{du}{dt}(0) = v^0.$$

Глава 10

Краевые задачи для обыкновенных дифференциальных уравнений

Наиболее важным классом краевых задач для обыкновенных дифференциальных уравнений являются задачи для уравнения второго порядка. Отмечены основные подходы к построению дискретных аналогов краевых задач с различными граничными условиями. Рассмотрены вопросы сходимости приближенного решения к точному и вычислительной реализации на основе использования прямых методов линейной алгебры. Помимо уравнений второго порядка кратко обсуждаются краевые задачи для модельного обыкновенного дифференциального уравнения четвертого порядка. Основное внимание уделяется разностным методам приближенного решения краевых задач.

10.1. Краевые задачи

В качестве базового рассматривается обыкновенное дифференциальное уравнение второго порядка

$$-\frac{d}{dx} \left(k(x) \frac{du}{dx} \right) + q(x) u = f(x), \quad 0 < x < l \quad (10.1)$$

с переменными коэффициентами

$$k(x) \geq \kappa > 0, \quad q(x) \geq 0.$$

Для однозначного определения неизвестной функции $u(x)$ уравнение (10.1) дополняется двумя граничными условиями на концах отрезка

Основные обозначения

$u = u(x), x \in [0, l]$	— неизвестная функция
$0 = x_0, x_1, \dots, x_N = l$	— узлы сетки
h	— шаг равномерной сетки
ω	— множество внутренних узлов
$\partial\omega$	— множество граничных узлов
H	— гильбертово пространство сеточных функций
(\cdot, \cdot)	— скалярное произведение в H
$\ \cdot \ $	— норма в H
$y_x = (y(x+h) - y(x))/h$	— правая разностная производная в точке x
$y_x = (y(x) - y(x-h))/h$	— левая разностная производная в точке x
$y_x^c = \frac{1}{2}(y_x + y_x)$	— центральная разностная производная в точке x
$y_{\bar{x}\bar{x}} = (y_x - y_x)/h$	— вторая разностная производная в точке x

ка $[0, l]$. Задаваться может функция $u(x)$ (граничное условие первого рода), поток $w(x) = -k(x) \frac{du}{dx}(x)$ (граничное условие второго рода) или же их линейная комбинация (граничное условие третьего рода):

$$u(0) = \mu_1, \quad u(l) = \mu_2, \quad (10.2)$$

$$-k(0) \frac{du}{dx}(0) = \mu_1, \quad k(l) \frac{du}{dx}(l) = \mu_2, \quad (10.3)$$

$$-k(0) \frac{du}{dx}(0) + \sigma_1 u(0) = \mu_1, \quad k(l) \frac{du}{dx}(l) + \sigma_2 u(l) = \mu_2. \quad (10.4)$$

Эллиптические уравнения второго порядка, прототипом которых является уравнение (10.1), используются при моделировании многих физико-механических процессов.

В задачах с разрывными коэффициентами (контакт двух сред) формулируются дополнительные условия. Простейшие из них (условие идеального контакта) для уравнения (10.1) связывается с непрерывностью решения и потока в точке контакта $x = x^*$:

$$[u(x)] = 0, \quad \left[k(x) \frac{du}{dx} \right] = 0, \quad x = x^*,$$

где использованы обозначения

$$[g(x)] = g(x + 0) - g(x - 0).$$

Отдельного рассмотрения заслуживают задачи с несамосопряженным оператором, когда, например,

$$- \frac{d}{dx} \left(k(x) \frac{du}{dx} \right) + v(x) \frac{du}{dx} + q(x) u = f(x), \quad 0 < x < l. \quad (10.5)$$

Уравнение конвекции-диффузии (10.5) является модельным при исследовании процессов в механике сплошной среды.

При описании деформаций пластин и оболочек, задач гидродинамики математические модели включают эллиптические уравнения четвертого порядка. Их рассмотрение необходимо начать с краевой задачи для обыкновенного дифференциального уравнения четвертого порядка. Простейшим такой задачей является задача для уравнения

$$\frac{d^4 u}{dx^4}(x) = f(x), \quad 0 < x < l. \quad (10.6)$$

В этом случае задаются по два граничных условия на концах отрезка. Например, уравнение (10.6) дополняется условиями первого рода:

$$u(0) = \mu_1, \quad u(l) = \mu_2, \quad (10.7)$$

$$\frac{du}{dx}(0) = \nu_1, \quad \frac{du}{dx}(l) = \nu_2. \quad (10.8)$$

При формулировке других типов краевых задач для уравнения (10.6) в граничных точках могут участвовать вторая и третья производные.

10.2. Численные методы решения краевых задач

При построении вычислительных алгоритмов для приближенного решения краевых задач для обыкновенных дифференциальных уравнений основное внимание уделяется вопросам аппроксимации уравнений, краевых условий и условий сопряжения для задач с разрывными коэффициентами. Проводится исследование точности приближенного решения в различных нормах, обсуждаются особенности прямых методов решения сеточных уравнений для рассматриваемого класса задач.

10.2.1. Аппроксимация краевых задач

Обозначим через $\bar{\omega}$ равномерную, для простоты, сетку с шагом h на интервале $[0, l]$:

$$\bar{\omega} = \{x \mid x = x_i = ih, \quad i = 0, 1, \dots, N, \quad Nh = l\},$$

причем ω — множество внутренних узлов, а $\partial\omega$ — множество граничных узлов.

Будем использовать безындексные обозначения, когда $u = u_i = u(x_i)$. Для левой разностной производной имеем

$$u_x \equiv \frac{u_i - u_{i-1}}{h} = \frac{du}{dx}(x_i) - \frac{h}{2} \frac{d^2u}{dx^2}(x_i) + O(h^2).$$

Тем самым левая разностная производная u_x аппроксимирует первую производную $\frac{du}{dx}$ с первым порядком (погрешность аппроксимации $O(h)$ в каждом внутреннем узле) при $u(x) \in C^{(2)}(\Omega)$. Аналогично для правой разностной производной получим

$$u_x \equiv \frac{u_{i+1} - u_i}{h} = \frac{du}{dx}(x_i) + \frac{h}{2} \frac{d^2u}{dx^2}(x_i) + O(h^2).$$

Для трехточечного шаблона (узлы x_{i-1}, x_i, x_{i+1}) можно использовать центральную разностную производную:

$$u_x \equiv \frac{u_{i+1} - u_{i-1}}{2h} = \frac{du}{dx}(x_i) + \frac{h^2}{3} \frac{d^3u}{dx^3}(x_i) + O(h^3),$$

которая аппроксимирует производную $\frac{du}{dx}$ со вторым порядком при $u(x) \in C^{(3)}(\Omega)$.

Для второй производной $\frac{d^2u}{dx^2}$ получим

$$u_{\bar{x}x} = \frac{u_x - u_{\bar{x}}}{h} = \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}.$$

Этот разностный оператор аппроксимирует в узле $x = x_i$ вторую производную со вторым порядком при $u(x) \in C^{(4)}(0, l)$.

Для внутренних узлов сетки аппроксимируем дифференциальный оператор

$$\mathcal{L}u = -\frac{d}{dx} \left(k(x) \frac{du}{dx} \right) + q(x)u, \quad x \in (0, l) \quad (10.9)$$

с достаточно гладкими коэффициентами и решением разностным оператором

$$Ly = -(ay_{\bar{x}})_x + cy, \quad x \in \omega. \quad (10.10)$$

Для аппроксимации со вторым порядком необходимо выбрать коэффициенты разностного оператора так, чтобы

$$\frac{a_{i+1} - a_i}{h} = \frac{dk}{dx}(x_i) + O(h^2), \quad (10.11)$$

$$\frac{a_{i+1} + a_i}{2} = k(x_i) + O(h^2), \quad (10.12)$$

$$c_i = q(x_i) + O(h^2). \quad (10.13)$$

В соответствии с (10.13) положим, например, $c_i = q(x_i)$, а условия (10.11), (10.12) удовлетворяют, в частности, следующие формулы для определения a_i :

$$a_i = k_{i-1/2} = k(x_i - 0,5h),$$

$$a_i = \frac{k_{i-1} + k_i}{2},$$

$$a_i = 2 \left(\frac{1}{k_{i-1}} + \frac{1}{k_i} \right)^{-1}.$$

Метод формальной замены дифференциальных операторов разностными может использоваться и при аппроксимации граничных условий. Для построения разностных схем в задачах с разрывными коэффициентами необходимо ориентироваться на использование интегро-интерполяционного метода (метода баланса).

При построении разностных схем естественно исходить из законов сохранения (балансов) для отдельных ячеек разностной сетки. В уравнении (10.1) выделим контрольные объемы в виде отрезков $x_{i-1/2} \leq x \leq x_{i+1/2}$, где $x_{i-1/2} = (i - 1/2)h$. Интегрирование уравнения (10.1) по контрольному объему дает

$$q_{i+1/2} - q_{i-1/2} + \int_{x_{i-1/2}}^{x_{i+1/2}} q(x)u(x) dx = \int_{x_{i-1/2}}^{x_{i+1/2}} f(x) dx.$$

Для получения разностного уравнения из этого балансного соотношения необходимо использовать те или иные восполнения сеточных функций. Само решение будем искать в целых узлах ($y(x)$, $x = x_i$), а потоки — в полуцелых ($q(x)$, $x = x_{i+1/2}$). Это приводит нас к разностному уравнению

$$Ly = \varphi, \quad x \in \omega, \quad (10.14)$$

в котором оператор L определен согласно (10.10) с коэффициентами

$$a_i = \left(\frac{1}{h} \int_{x_{i-1}}^{x_i} \frac{dx}{k(x)} \right)^{-1}, \quad (10.15)$$

$$c_i = \frac{1}{h} \int_{x_{i-1/2}}^{x_{i+1/2}} q(x) dx.$$

Правая часть уравнения (10.14) есть

$$\varphi_i = \frac{1}{h} \int_{x_{i-1/2}}^{x_{i+1/2}} f(x) dx.$$

Аналогично проводятся аппроксимации уравнения (10.1) и на неравномерных сетках.

Построение дискретных аналогов краевых задач для уравнения (10.1) может осуществляться на основе метода конечных элементов. Используя простейшие кусочно-линейные элементы, представим приближенное решение в виде

$$y(x) = \sum_{i=1}^{N-1} y_i w_i(x), \quad (10.16)$$

где пробные функции $w_i(x)$ имеют вид

$$w_i(x) = \begin{cases} 0, & x < x_{i-1}; \\ \frac{(x - x_{i-1})}{h}, & x_{i-1} \leq x \leq x_i; \\ \frac{(x_{i+1} - x)}{h}, & x_i \leq x \leq x_{i+1}; \\ 0, & x > x_{i+1}. \end{cases}$$

Коэффициенты разложения в методе Галеркина определяются из системы линейных уравнений, которую мы получаем после умножения исходного уравнения (10.1) на функцию $w_i(x)$ и интегрирования по всей области. С учетом финитности пробных функций получим

$$\int_{x_{i-1}}^{x_{i+1}} k(x) \frac{dy}{dx} \frac{dw_i}{dx} dx + \int_{x_{i-1}}^{x_{i+1}} q(x) y(x) w_i(x) dx = \int_{x_{i-1}}^{x_{i+1}} f(x) w_i(x) dx.$$

Подстановка представления приближенного решения (10.16) приводит к трехточечному разностному уравнению (10.14), в котором

$$a_i = \frac{1}{h} \int_{x_{i-1}}^{x_i} k(x) dx - \frac{1}{h} \int_{x_{i-1}}^{x_i} q(x)(x - x_{i-1})(x_i - x) dx,$$

$$c_i = \frac{1}{h^2} \left(\int_{x_{i-1}}^{x_i} q(x)(x - x_{i-1}) dx + \int_{x_{i-1}}^{x_i} q(x)(x_{i+1} - x) dx \right),$$

$$\varphi_i = \frac{1}{h^2} \left(\int_{x_{i-1}}^{x_i} f(x)(x - x_{i-1}) dx + \int_{x_{i-1}}^{x_i} f(x)(x_{i+1} - x) dx \right).$$

Наиболее просто аппроксимируются граничные условия (10.2):

$$y_0 = \mu_1, \quad y_N = \mu_2. \quad (10.17)$$

Для аппроксимации граничных условий второго и третьего рода со вторым порядком в граничных узлах $x = x_0 = 0$ и $x = x_N = l$ привлекается уравнение (10.1) — аппроксимация на решениях задачи. В случае уравнения (10.1) краевые условия (10.4) аппроксимируются разностными соотношениями

$$\begin{aligned} -a_1 y_{\bar{x},1} + \left(\sigma_1 + \frac{h}{2} q_0 \right) y_0 &= \mu_1 + \frac{h}{2} f_0, \\ a_N y_{x,N} + \left(\sigma_2 + \frac{h}{2} q_N \right) y_N &= \mu_2 + \frac{h}{2} f_N. \end{aligned}$$

К подобным аппроксимациям мы приходим при использовании интегроинтерполяционного метода и при построении схем конечных элементов.

10.2.2. Сходимость разностных схем

Исследование сходимости приближенного решения к точному при численном решении краевых задач базируется на основе априорных оценок в сеточном гильбертовом пространстве. При исследовании сходимости в равномерной норме привлекается принцип максимума и разностная функция Грина.

На множестве внутренних узлов ω и на сетке

$$\omega^+ = \{x \mid x = x_i = ih, \quad i = 1, 2, \dots, N, \quad Nh = l\}$$

определим скалярные произведения

$$\begin{aligned} (y, w) &\equiv \sum_{x \in \omega} y(x)w(x)h, \\ (y, w)^+ &\equiv \sum_{x \in \omega^+} y(x)w(x)h. \end{aligned}$$

В сеточных гильбертовых пространствах H и H^+ норму введем соотношением

$$\|y\| = (y, y)^{1/2}, \quad \|y\|^+ = ((y, y)^+)^{1/2}.$$

Рассмотрим разностное уравнение (10.14) при однородных краевых условиях первого рода:

$$y_0 = 0, \quad y_N = 0. \quad (10.18)$$

Для любых сеточных функций, обращающихся в нуль на $\partial\omega$, верно неравенство (разностное неравенство Фридрикса)

$$\|y\|^2 \leq M_0 (\|y_{\bar{x}}\|^+)^2, \quad M_0 = \frac{l^2}{8}. \quad (10.19)$$

С учетом этого на множестве сеточных функций, удовлетворяющих (10.18), разностный оператор L , определяемый согласно (10.10), является самосопряженным и положительно определенным:

$$L = L^* \geq \frac{\kappa}{M_0} E. \quad (10.20)$$

Для исследования точности разностной схемы (10.14), (10.17) рассмотрим задачу для погрешности приближенного решения

$$z(x) = y(x) - u(x), \quad x \in \bar{\omega}.$$

Для погрешности приближенного решения задачи (10.1), (10.2) получим разностную задачу

$$\begin{aligned} Lz &= \psi(x), \quad x \in \omega, \\ z_0 &= 0, \quad z_N = 0, \end{aligned}$$

где $\psi(x)$ — погрешность аппроксимации:

$$\psi(x) = \varphi(x) - Lu, \quad x \in \omega.$$

В случае достаточно гладких коэффициентов и решения для погрешности аппроксимации получим

$$\psi(x) = O(h^2), \quad x \in \omega.$$

Для погрешности рассматриваемой разностной схемы справедлива априорная оценка

$$\|z_{\bar{x}}\|^+ \leq \frac{M_0^{1/2}}{\kappa} \|\psi\|,$$

которая обеспечивает сходимость разностного решения к точному решению дифференциальной задачи со вторым порядком.

При рассмотрении одномерных задач конвекции-диффузии мы ориентируемся на использование трехточечных разностных схем, которые запишем для внутренних узлов в виде

$$-\alpha_i y_{i-1} + \gamma_i y_i - \beta_i y_{i+1} = \varphi_i, \quad i = 1, 2, \dots, N-1. \quad (10.21)$$

Для граничных узлов считаем выполненными условия (10.18).

Будем рассматривать разностные схемы (10.18), (10.21), в которых

$$\alpha_i > 0, \quad \beta_i > 0, \quad \gamma_i > 0, \quad i = 1, 2, \dots, N-1.$$

Сформулируем критерий монотонности разностной схемы, т.е. сформулируем условия, при которых разностная схема удовлетворяет разностному принципу максимума.

Пусть в разностной схеме (10.18), (10.21) $\varphi_i \geq 0$ для всех $i = 1, 2, \dots, N-1$ (или же $\varphi_i \leq 0$ для $i = 1, 2, \dots, N-1$). Тогда при выполнении условий

$$\gamma_i \geq \alpha_i + \beta_i, \quad i = 1, 2, \dots, N-1 \quad (10.22)$$

имеет место $y_i \geq 0$, $i = 1, 2, \dots, N-1$ ($y_i \leq 0$, $i = 1, 2, \dots, N-1$).

Для разностных схем (10.18), (10.21), для которых выполнены условия монотонности (10.22), доказывается сходимость в равномерной норме. Исследование базируется на применении соответствующих теорем сравнения и построении мажорантных функций.

Пусть для разностной схемы (10.18), (10.21) выполнены условия (10.22) и $w(x)$ — решение задачи

$$-\alpha_i w_{i-1} + \gamma_i w_i - \beta_i w_{i+1} = \phi_i, \quad i = 1, 2, \dots, N-1,$$

$$w_0 = 0, \quad w_N = 0.$$

Тогда при

$$|\varphi_i| \leq \phi_i, \quad i = 1, 2, \dots, N-1$$

справедлива оценка

$$|y_i| \leq w_i, \quad i = 1, 2, \dots, N-1.$$

Функция $w(x)$ называется мажорантной функцией для решения задачи (10.18), (10.21). Если удастся построить мажоранту, то это значит, что получена априорная оценка для решения задачи в $L_\infty(\omega)$:

$$\|y(x)\|_\infty \leq \|w(x)\|_\infty, \quad (10.23)$$

где на множестве сеточных функций, обращающихся в нуль на $\partial\omega$,

$$\|y(x)\|_{\infty} \equiv \max_{x \in \omega} |y(x)|.$$

На основе рассмотрения задачи для погрешности с использованием оценки (10.23) устанавливается сходимость исследуемой разностной схемы.

10.2.3. Другие задачи

Среди более общих, чем (10.1), (10.4) краевых задач отметим задачи для уравнения (10.5). Простейшая центральноразностная аппроксимация члена с первой производной дает разностное уравнение

$$-(ay_{\bar{x}})_x + by_{\bar{x}} + cy = \varphi, \quad x \in \omega, \quad (10.24)$$

где, например, $b_x = v(x_x)$. Разностная схема (10.17), (10.24) аппроксимирует краевую задачу (10.2), (10.5) со вторым порядком. Ее основной недостаток связан с тем, что эта схема монотонна только при достаточно малых шагах сетки h .

Безусловно монотонные разностные схемы для уравнения (10.5) можно построить при использовании для конвективного слагаемого аппроксимаций первого порядка направленными разностями. Вместо (10.24) рассмотрим разностное уравнение

$$-(ay_{\bar{x}})_x + b^+ y_{\bar{x}} + b^- y_x + cy = \varphi, \quad x \in \omega, \quad (10.25)$$

где

$$\begin{aligned} b(x) &= b^+(x) + b^-(x), \\ b^+(x) &= \frac{1}{2}(b(x) + |b(x)|) \geq 0, \\ b^-(x) &= \frac{1}{2}(b(x) - |b(x)|) \leq 0. \end{aligned}$$

К сожалению, схема (10.17), (10.25) имеет только первый порядок аппроксимации.

При разностной аппроксимации краевой задачи для обыкновенного дифференциального уравнения четвертого порядка (10.8)–(10.11) удобно

использовать расширенную сетку с дополнительными (фиктивными) узлами $x_{-1} = -h$, $x_{N+1} = l+h$. Тогда дифференциальному уравнению (10.8) можно сопоставить разностное уравнение

$$y_{\bar{x}\bar{x}\bar{x}} = \varphi(x), \quad x \in \omega. \quad (10.26)$$

Аппроксимация краевых условий (10.9) и (10.10) дает

$$y_0 = \mu_1, \quad y_N = \mu_2, \quad (10.27)$$

$$\frac{y_1 - y_{-1}}{2h} = \nu_1, \quad \frac{y_{N+1} - y_{N-1}}{2h} = \nu_2. \quad (10.28)$$

При вычислительной реализации значения в фиктивных и граничных узлах находятся из (10.27), (10.28) непосредственно, а для определения приближенного решения в узлах $x \in \omega$ из (10.26) получим пятидиагональную систему линейных алгебраических уравнений.

10.2.4. Решение сеточных уравнений

Для нахождения приближенного решения краевой задачи для обыкновенного дифференциального уравнения необходимо решить соответствующую систему линейных алгебраических уравнений. Для нахождения разностного решения используются традиционные прямые методы линейной алгебры. Излагаемый метод прогонки (алгоритм Томаса), как хорошо известно, является классическим методом Гаусса для матриц специальной ленточной структуры.

Для примера рассмотрим разностное трехточечное уравнение (10.21) с однородными условиями (10.18). В подобном виде записываются и разностные схемы для задачи с краевыми условиями третьего рода (на расширенной сетке с $y_{-1} = 0$, $y_{N+1} = l+h$). В матричном виде рассматриваемая разностная задача имеет вид

$$Ay = \varphi, \quad x \in \omega,$$

где

$$A = \begin{bmatrix} \gamma_1 & -\beta_1 & 0 & \dots & 0 \\ -\alpha_2 & \gamma_2 & -\beta_2 & \dots & 0 \\ 0 & -\alpha_3 & \gamma_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \gamma_{N-1} \end{bmatrix}.$$

Для нахождения решения сеточной задачи используются следующие рекуррентные формулы для вычисления прогоночных коэффициентов (прямая прогонка):

$$\xi_{i+1} = \frac{\beta_i}{\gamma_i - \alpha_i \xi_i}, \quad i = 1, 2, \dots, N-1,$$

$$\vartheta_{i+1} = \frac{\varphi_i + \alpha_i \vartheta_i}{\gamma_i - \alpha_i \xi_i}, \quad i = 1, 2, \dots, N-1$$

при

$$\xi_1 = 0, \quad \vartheta_1 = 0.$$

Для решения имеем (обратная прогонка)

$$y_i = \xi_{i+1} y_{i+1} + \vartheta_{i+1}, \quad i = 0, 1, \dots, N-1, \quad y_N = 0.$$

Пусть для системы уравнений (10.21), (10.18) выполнены условия

$$|\alpha_i| > 0, \quad |\beta_i| > 0, \quad i = 1, 2, \dots, N-1,$$

$$|\gamma_i| \geq |\alpha_i| + |\beta_i|, \quad i = 1, 2, \dots, N-1.$$

Тогда алгоритм прогонки корректен, т.е. в расчетных формулах $\gamma_i - \alpha_i \xi_i \neq 0$.

В настоящее время существует ряд вариантов метода прогонки, ориентированных на определенный класс сеточных задач. Среди них отметим прогонку для задач с периодическими граничными условиями, метод прогонки для пятиточечных разностных уравнений.

10.3. Упражнения

Приведем некоторые примеры построения и исследования вычислительных алгоритмов для приближенного решения краевых задач для обыкновенных дифференциальных уравнений.

Упражнение 10.1. Найдите решение краевой задачи (10.1), (10.4) на основе решения задач Коши.

Решение. Среди возможных подходов к сведению граничной задачи к задаче Коши отметим метод вариации постоянных. В этом случае решение представляется в виде

$$u(x) = y(x) + c_1 v(x) + c_2 w(x),$$

где c_1, c_2 произвольные постоянные, $y(x), v(x)$ и $w(x)$ — решения следующих задач Коши:

$$-\frac{d}{dx} \left(k(x) \frac{dy}{dx} \right) + q(x)y = f(x), \quad 0 < x < l,$$

$$y(0) = 0, \quad k(0) \frac{dy}{dx}(0) = 0,$$

$$-\frac{d}{dx} \left(k(x) \frac{dv}{dx} \right) + q(x)v = 0, \quad 0 < x < l,$$

$$v(0) = 0, \quad k(0) \frac{dv}{dx}(0) = 1,$$

$$-\frac{d}{dx} \left(k(x) \frac{dw}{dx} \right) + q(x)w = 0, \quad 0 < x < l,$$

$$w(0) = 1, \quad k(0) \frac{dw}{dx}(0) = 0.$$

Граничные условия (10.4) приводят к системе уравнений

$$\begin{aligned} -c_1 + \sigma_1 c_2 &= \mu_1, \\ c_1 \left(k(l) \frac{dv}{dx}(l) + \sigma_2 v(l) \right) + c_2 \left(k(l) \frac{dw}{dx}(l) + \sigma_2 w(l) \right) &= \\ &= \mu_2 - k(l) \frac{dy}{dx}(l) - \sigma_2 v(l) \end{aligned}$$

для определения постоянных c_1 и c_2 .

Упражнение 10.2. *Аппроксимируйте граничные условия третьего рода (10.4) со вторым порядком для решения уравнения (10.1) при использовании расширенной на полшага сетки.*

Решение. Введем сетку с узлами

$$x_i = x_0 + ih, \quad i = 0, 1, \dots, N, \quad x_0 = -\frac{h}{2}, \quad x_N = l + \frac{h}{2},$$

т.е. сетка сдвинута на полшага. Для внутренних узлов ($x_i, i = 1, 2, \dots, N-1$) применяется обычная аппроксимация (10.10), (10.14). Для достаточно гладкой функции $u(x)$ имеем

$$\begin{aligned} u(x) &= \frac{1}{2} (u(x+0,5h) + u(x-0,5h)) + O(h^2), \\ \frac{du}{dx}(x) &= \frac{1}{h} (u(x+0,5h) - u(x-0,5h)) + O(h^2). \end{aligned}$$

В силу этого граничные условия (10.4) аппроксимируются условиями

$$\begin{aligned} -k(0)y_{x,0} + \sigma_1 \frac{y_0 + y_1}{2} &= \mu_1, \\ k(l)y_{\bar{x},N} + \sigma_2 \frac{y_N + y_{N-1}}{2} &= \mu_2 \end{aligned}$$

со вторым порядком.

Упражнение 10.3. Постройте схему четвертого порядка аппроксимации для уравнения

$$-\frac{d^2 u}{dx^2} + q(x)u = f(x) \quad (10.29)$$

на равномерной сетке при использовании трехточечного шаблона.

Решение. Используем стандартную аппроксимацию

$$-y_{\bar{x}x} + c y = \varphi \quad (10.30)$$

и подберем сеточные функции c и φ . Погрешность аппроксимации в нашем случае есть

$$\begin{aligned} \psi &= \varphi + u_{\bar{x}x} - cu = \varphi - f + \left(u_{\bar{x}x} - \frac{d^2 u}{dx^2} \right) - (c - q)u = \\ &= \frac{h^2}{12} \frac{d^4 u}{dx^4} + \varphi - f - (c - q)u + O(h^4). \end{aligned}$$

На решениях уравнения (10.29) имеем

$$\frac{d^4 u}{dx^4} = q^2 u - qf - \frac{d^2 f}{dx^2}$$

и поэтому

$$\psi = \varphi - \left(f + \frac{h^2}{12} qf + \frac{h^2}{12} \frac{d^2 f}{dx^2} \right) - \left(c - q - \frac{h^2}{12} q^2 \right) u + O(h^4).$$

Поэтому схема (10.30), в которой

$$c = q + \frac{h^2}{12} q^2, \quad \varphi = f + \frac{h^2}{12} qf + \frac{h^2}{12} \frac{d^2 f}{dx^2}$$

будет искомой схемой четвертого порядка.

Упражнение 10.4. Пусть погрешность разностной схемы (10.10), (10.14), (10.17) имеет следующий вид

$$\psi(x) = \psi^0(x) + \psi^*(x), \quad \psi^0 = \eta_x, \quad x \in \omega, \quad (10.31)$$

причем

$$\eta(x) = O(h^2), \quad \psi^*(x) = O(h^2).$$

Покажите, что разностная схема сходится со вторым порядком.

Решение. Для погрешности разностного решения $z(x)$ справедлива априорная оценка

$$\|z_{\bar{x}}\|^+ \leq \frac{1}{\kappa} (\|\eta\|^+ + M_0^{1/2} \|\psi^*\|). \quad (10.32)$$

Для ее доказательства скалярно умножим уравнение для погрешности

$$Lz = \psi(x), \quad x \in \omega$$

на $z(x)$ и получим равенство

$$-((az_{\bar{x}})_x, z) + (cz, z) = (\psi, z).$$

Принимая во внимание, что $k(x) \geq \kappa > 0$, будем иметь

$$-((az_{\bar{x}})_x, z) = (az_{\bar{x}}, z_{\bar{x}})^+ \geq \kappa (\|z_{\bar{x}}\|^+)^2.$$

Левая часть с учетом представления (10.31) для погрешности аппроксимации и неравенства Фридрикса преобразуется следующим образом:

$$\begin{aligned} (\psi, z) &= (\eta_{\bar{x}}, z) + (\psi^*, z) \leq -(\eta, z_{\bar{x}})^+ + \|\psi^*\| \|z\| \leq \\ &\leq (\|\eta\|^+ + M_0^{1/2} \|\psi^*\|) \|z_{\bar{x}}\|^+. \end{aligned}$$

С учетом неотрицательности κ это дает априорную оценку для погрешности (10.32). Из этой оценки следует сходимость разностной схемы (10.10), (10.14), (10.17) со вторым порядком. Заметим, что при этом локальная погрешность аппроксимации имеет только первый порядок.

Упражнение 10.5. Постройте абсолютно монотонную разностную схему второго порядка аппроксимации для краевой задачи (10.2), (10.5).

Решение. При построении дискретных аналогов для одномерных задач полезно использовать возможность перехода от исходной несамосопряженной дифференциальной задачи к самосопряженной, для которой уже потом строить те или иные аппроксимации. От исследуемого уравнения (10.5) можно перейти к уравнению

$$-\frac{d}{dx} \left(\bar{k}(x) \frac{du}{dx} \right) + \bar{q}(x) u = \bar{f}(x), \quad 0 < x < l, \quad (10.33)$$

в котором

$$\bar{k}(x) = k(x)\chi(x), \quad \bar{f}(x) = f(x)\chi(x), \quad \bar{q}(x) = q(x)\chi(x),$$

где

$$\chi(x) = \exp \left(- \int_0^x \frac{v(s)}{k(s)} ds \right).$$

Далее строятся обычные разностные схемы второго порядка точности для уравнения (10.33). Например, можно использовать разностную схему

$$-(\bar{a}y_{\bar{x}})_x + \bar{c}y = \bar{\varphi}(x), \quad x \in \omega, \quad (10.34)$$

в которой, например,

$$\bar{a}(x) = \bar{k}(x - 0,5h) = k(x - 0,5h)\chi(x - 0,5h),$$

$$\chi(x - 0,5h) = \exp \left(- \int_0^{x-0,5h} \frac{v(s)}{k(s)} ds \right).$$

С точностью до $O(h^2)$ положим

$$\chi(x - 0,5h) = \chi(x) \exp(\theta(x)), \quad \theta(x) = \frac{v(x)h}{2k(x)}.$$

Левая часть (10.34) преобразуется следующим образом

$$\begin{aligned} -(\bar{a}y_{\bar{x}})_x &= \frac{\chi(x)}{h} k(x - 0,5h) \exp(\theta(x)) y_{\bar{x}} - \\ &- \frac{\chi(x)}{h} k(x + 0,5h) \exp(-\theta(x)) y_x. \end{aligned}$$

Приходим к разностной схеме

$$-\frac{1}{h} \left(k(x + 0,5h) \exp(-\theta(x)) y_x - k(x - 0,5h) \exp(\theta(x)) y_x \right) + c(x)y = \varphi(x), \quad x \in \omega,$$

которая является монотонной и аппроксимирует уравнение (10.5) со вторым порядком.

Упражнение 10.6. Постройте вариант метода прогонки для решения системы линейных алгебраических уравнений

$$\begin{aligned} -\alpha_0 y_N + \gamma_0 y_0 - \beta_0 y_1 &= \varphi_0, \\ -\alpha_i y_{i-1} + \gamma_i y_i - \beta_i y_{i+1} &= \varphi_i, \quad i = 1, 2, \dots, N-1, \\ -\alpha_N y_{N-1} + \gamma_N y_N - \beta_N y_0 &= \varphi_N, \end{aligned} \quad (10.35)$$

которая возникает, например, при рассмотрении краевой задачи для уравнения (10.1) с периодическими краевыми условиями.

Решение. Будем использовать представление решения сеточной задачи (10.35) в виде

$$y_i = v_i + y_N w_i, \quad i = 0, 1, \dots, N. \quad (10.36)$$

В силу этого для новых сеточных функций v и w получим условия

$$v_N = 0, \quad w_N = 1.$$

Далее подставляем представление (10.36) в (10.35). Из первого уравнения получим

$$\begin{aligned} \gamma_0 v_0 - \beta_0 v_1 &= \varphi_0, \\ \gamma_0 w_0 - \beta_0 w_1 &= \alpha_0. \end{aligned}$$

Для внутренних узлов получим

$$\begin{aligned} -\alpha_i v_{i-1} + \gamma_i v_i - \beta_i v_{i+1} &= \varphi_i, \quad i = 1, 2, \dots, N-1, \\ -\alpha_i w_{i-1} + \gamma_i w_i - \beta_i w_{i+1} &= \alpha_0, \quad i = 1, 2, \dots, N-1. \end{aligned}$$

Тем самым для определения сеточных функций v и w мы имеем стандартные задачи с трехдиагональной матрицей. Для их нахождения используется метод прогонки.

Для определения y_N используется последнее из уравнений (10.35):

$$y_N = \frac{\varphi_N + \beta_N v_0 + \alpha_N w_{N-1}}{\gamma_N - \beta_N w_0 - \alpha_N w_{N-1}}.$$

После этого в соответствии с (10.36) рассчитываются другие значения сеточной функции y .

10.4. Задачи

Задача 10.1. Рассмотрите возможность сведения краевой задачи (10.1), (10.4) к решению задач Коши на основе использования представления решения в виде (метод дифференциальной прогонки):

$$u(x) = v(x) \frac{du}{dx}(x) + w(x).$$

Задача 10.2. Покажите, что на произвольной неравномерной сетке $x_{i+1} = x_i + h_{i+1}$ разностное уравнение

$$-\frac{2}{h_{i+1} + h_i} \left(\frac{y_{i+1} - y_i}{h_{i+1}} - \frac{y_i - y_{i-1}}{h_i} \right) + q(\bar{x}) \frac{y_{i+1} + y_i + y_{i-1}}{3} = f(\bar{x}),$$

$$\bar{x} = \frac{x_{i+1} + x_i + x_{i-1}}{3}$$

аппроксимирует уравнение

$$-\frac{d^2 u}{dx^2} + q(x)u = f(x)$$

со вторым порядком в неузловой точке $x = \bar{x}$.

Задача 10.3. С помощью интегро-интерполяционного метода постройте разностную схему второго порядка аппроксимации на равномерной сетке

$$\bar{\omega} = \{x \mid x = x_i = ih, \quad i = 0, 1, \dots, N, \quad Nh = l\}$$

для краевой задачи

$$-\frac{1}{x} \frac{d}{dx} \left(xk(x) \frac{du}{dx} \right) + q(x)u = f(x), \quad 0 < x < l, \quad (10.37)$$

$$\lim_{x \rightarrow 0} xk(x) \frac{du}{dx}(x) = 0, \quad u(l) = \mu. \quad (10.38)$$

Задача 10.4. Для краевой задачи (10.1) с однородными граничными условиями постройте схему конечных элементов с представлением решения в виде (10.16) на основе минимизации функционала (метод Рунта)

$$J(v) = \frac{1}{2} \int_0^l \left(k(x) \left(\frac{dv}{dx} \right)^2 + q(x)v^2(x) \right) dx - \int_0^l f(x)v(x) dx.$$

Задача 10.5. Для решения задачи (10.1), (10.4) справедливо интегральное тождество

$$\int_0^l \left(k(x) \frac{du}{dx} \frac{dv}{dx} + q(x)uv - f(x)v \right) dx + \sigma_1 u(0)v(0) + \sigma_2 u(l)v(l) - \mu_1 v(0) - \mu_2 v(l) = 0,$$

где $v = v(x)$ — произвольная дифференцируемая функция. На основе использования квадратурной формулы трапеций аппроксимируйте это интегральное тождество и постройте разностную схему для приближенного решения краевой задачи (10.1), (10.4) (метод сумматорных тождеств).

Задача 10.6. При решении краевой задачи

$$-\frac{d}{dx} \left(k(x) \frac{du}{dx} \right) = 0, \\ u(0) = 1, \quad u(1) = 0$$

используется разностная схема

$$-k(x)y_{xx} - k_x y_x = 0.$$

Покажите, что эта схема расходится в классе разрывных коэффициентов.

Задача 10.7. Покажите справедливость следующих разностных аналогов первой и второй формул Грина

$$\begin{aligned} ((ay_{\bar{x}})_x, w) &= -(ay_{\bar{x}}, w_{\bar{x}})^+ + a_N y_{\bar{x}, N} w_N - a_1 y_{x, 0} w_0, \\ ((ay_{\bar{x}})_x, w) - (y, (aw_{\bar{x}})_x) &= a_N (y_{\bar{x}, N} w_N - y_N w_{\bar{x}, N}) - a_1 (y_{x, 0} w_0 - y_0 w_{x, 0}) \end{aligned}$$

на равномерной сетке.

Задача 10.8. Покажите, что разностное уравнение

$$\begin{aligned} -\alpha_i y_{i-1} + \gamma_i y_i - \beta_i y_{i+1} &= \varphi_i, \quad \alpha_i \neq 0, \quad \beta_i \neq 0, \quad i = 1, 2, \dots, N-1, \\ y_0 &= \mu_1, \quad y_N = \mu_2 \end{aligned}$$

можно привести к симметричному виду

$$\bar{\alpha}_{i+1} (y_{i+1} - y_i) - \bar{\alpha}_i (y_i - y_{i-1}) - d_i y_i = \bar{\varphi}_i, \quad i = 1, 2, \dots, N-1.$$

Задача 10.9. Докажите, что для всякой сеточной функции $y(x)$, заданной на сетке

$$\bar{\omega} = \{x \mid x = x_i = ih, \quad i = 0, 1, \dots, N, \quad Nh = l\}$$

и обращающейся в нуль при $x = 0$ и $x = l$, справедливо неравенство (теорема вложения для сеточных функций)

$$\|y(x)\|_{\infty} \leq \frac{\sqrt{l}}{2} \|y_{\bar{x}}\|^+.$$

Задача 10.10. Найдите собственные функции и собственные значения разностной задачи

$$y_{\bar{x}x} + \lambda y = 0, \quad x \in \omega,$$

$$y_0 = 0, \quad y_N = 0.$$

Задача 10.11. Для уравнения (10.1), в котором

$$k(x) \geq \kappa > 0, \quad q(x) \geq \delta > 0,$$

рассмотрите краевую задачу с условиями периодичности

$$u(x + l) = u(x).$$

Постройте разностную схему второго порядка аппроксимации на равномерной сетке и исследуйте ее сходимость на основе принципа максимума.

Задача 10.12. Для задачи (10.1), (10.2) с разрывными коэффициентами постройте интегро-интерполяционным методом разностную схему (10.10), (10.14), (10.17) и на основе представления погрешности аппроксимации в виде (10.31) докажите сходимость разностного решения к точному со вторым порядком.

Задача 10.13. Покажите, что при решении задачи (10.1), (10.2) на неравномерной сетке $x_{i+1} = x_i + h_{i+1}$ с использованием разностной схемы

$$-\frac{2}{h_{i+1} + h_i} \left(a_{i+1} \frac{y_{i+1} - y_i}{h_{i+1}} - a_i \frac{y_i - y_{i-1}}{h_i} \right) + c_i y_i = \varphi_i, \\ i = 1, 2, \dots, N - 1$$

приближенное решение сходится к точному со вторым порядком.

Задача 10.14. Для решения краевой задачи (10.37), (10.38) постройте разностную схему на сетке

$$x_0 = 0, \quad x_1 = 0,5h, \quad x_{i+1} = x_i + h, \quad i = 1, 2, \dots, N, \quad (N - 0,5)h = l$$

и исследуйте ее сходимость.

Задача 10.15. Покажите, что разностная схема

$$-(ay_{\bar{x}})_x = \varphi(x), \quad x \in \omega,$$

в которой

$$a_i = \left(\frac{1}{h} \int_{x_{i-1}}^{x_i} \frac{dx}{k(x)} \right)^{-1}, \\ \varphi_i = \frac{1}{h^2} \left(a_{i+1} \int_{x_i}^{x_{i-1}} \frac{dx}{k(x)} \int_{x_i}^x f(s) ds + a_i \int_{x_{i-1}}^{x_i} \frac{dx}{k(x)} \int_x^{x_i} f(s) ds \right),$$

для приближенного решения уравнения

$$-\frac{d}{dx} \left(k(x) \frac{du}{dx} \right) = f(x), \quad 0 < x < l$$

с граничными условиями (10.2) является точной.

Задача 10.16. Пусть выполнены условия

$$|\alpha_i| > 0, \quad |\beta_i| > 0, \quad \delta_i = |\gamma_i| - |\alpha_i| - |\beta_i| > 0 \quad i = 1, 2, \dots, N-1.$$

Докажите, что в этих условиях для решения задачи

$$-\alpha_i y_{i-1} + \gamma_i y_i - \beta_i y_{i+1} = \varphi_i, \quad i = 1, 2, \dots, N-1,$$

$$y_0 = 0, \quad y_N = 0$$

справедлива оценка

$$\|y\|_\infty \leq \left\| \left\| \frac{\varphi}{\delta} \right\| \right\|_\infty.$$

Задача 10.17. Рассматривается задача

$$-\frac{d^2 u}{dx^2} + v(x) \frac{du}{dx} + q(x)u = f(x), \quad 0 < x < l,$$

$$u(0) = \mu_1, \quad u(l) = \mu_2.$$

Покажите, что разностная схема

$$-\frac{1}{1+\tau} y_{\bar{x}x} + v^+ y_x + v^- y_x + cy = \varphi, \quad x \in \omega,$$

где

$$v(x) = v^+(x) + v^-(x),$$

$$v^+(x) = \frac{1}{2}(v(x) + |v(x)|) \geq 0,$$

$$v^-(x) = \frac{1}{2}(v(x) - |v(x)|) \leq 0,$$

$$\tau(x) = \frac{1}{2}|v(x)|h,$$

является безусловно монотонной и имеет второй порядок аппроксимации.

Задача 10.18. На основе формул прогонки получите априорную оценку

$$\|y\|_{\infty} \leq \sum_{i=1}^{N-1} \frac{h}{|a_{i+1}|} \sum_{k=1}^i |\varphi_k|$$

для решения разностной задачи (10.10), (10.14), (10.18).

Задача 10.19. Получите расчетные формулы пятидиагональной прогонки для решения системы уравнений

$$\begin{aligned} \gamma_0 y_0 - \delta_0 y_1 + \varepsilon_0 y_2 &= \varphi_0, \\ -\beta_1 y_0 + \gamma_1 y_1 - \delta_1 y_2 + \varepsilon_1 y_3 &= \varphi_1, \\ \alpha_i y_{i-2} - \beta_i y_{i-1} + \gamma_i y_i - \delta_i y_{i+1} + \varepsilon_i y_{i+2} &= \varphi_i, \\ i &= 2, 3, \dots, N-2, \\ \alpha_{N-1} y_{N-3} - \beta_{N-1} y_{N-2} + \gamma_{N-1} y_{N-1} - \delta_{N-1} y_N &= \varphi_{N-1}, \\ \alpha_N y_{N-2} - \beta_N y_{N-1} + \gamma_N y_N &= \varphi_N. \end{aligned}$$

Задача 10.20. На множестве сеточных функций, обращающихся в нуль в точках $x = 0$ и $x = l$ рассматриваются схемы

$$\begin{aligned} -(ay_{\bar{x}})_x + cy &= \varphi, \quad x \in \omega, \\ -(\bar{a}\bar{y}_{\bar{x}})_x + \bar{c}\bar{y} &= \bar{\varphi}, \quad x \in \omega \end{aligned}$$

с коэффициентами

$$a(x) \geq \kappa > 0, \quad \bar{a}(x) \geq \kappa > 0, \quad c(x) \geq 0, \quad \bar{c} \geq 0.$$

Получите оценки разности $z(x) = \bar{y}(x) - y(x)$ через величины возмущения коэффициентов и правой части (коэффициентная устойчивость).

Глава 11

Краевые задачи для эллиптических уравнений

Среди стационарных задач математической физики наибольшее внимание уделяется краевым задачам для эллиптических уравнений второго порядка. Рассматриваются вопросы аппроксимации таких уравнений и краевых условий, формулируется принцип максимума для сеточных уравнений. Проводится исследование сходимости приближенного решения к точному в различных нормах. Отмечаются некоторые основные итерационные методы решения сеточных уравнений.

11.1. Двумерные краевые задачи

Будем рассматривать двумерные краевые задачи, когда расчетная область есть прямоугольник

$$\Omega = \{x \mid x = (x_1, x_2), 0 < x_\alpha < l_\alpha, \alpha = 1, 2\}.$$

Основным объектом нашего исследования будет эллиптическое уравнение второго порядка

$$-\sum_{\alpha=1}^2 \frac{\partial}{\partial x_\alpha} \left(k(x) \frac{\partial u}{\partial x_\alpha} \right) + q(x) u = f(x), \quad x \in \Omega. \quad (11.1)$$

На коэффициенты уравнения накладываются ограничения

$$k(x) \geq \kappa > 0, \quad q(x) \geq 0, \quad x \in \Omega.$$

Основные обозначения

$u = u(\mathbf{x})$, $\mathbf{x} = (x_1, x_2)$ — неизвестная функция

h_1, h_2 — шаги равномерной сетки

ω — множество внутренних узлов

$\partial\omega$ — множество граничных узлов

H — гильбертово пространство сеточных функций

(\cdot, \cdot) — скалярное произведение в H

$\|\cdot\|$ — норма в H

$y_{x_1} = (y(x_1 + h_1, x_2) - y(x)) / h_1$ — правая разностная производная в точке \mathbf{x} по переменной x_1

$y_{\bar{x}_1} = (y(\mathbf{x}) - y(x_1 - h_1, x_2)) / h_1$ — левая разностная производная в точке \mathbf{x} по переменной x_1

$y_{x_1}^0 = \frac{1}{2}(y_{x_1} + y_{\bar{x}_1})$ — центральная разностная производная в точке \mathbf{x} по переменной x_1

$y_{\bar{x}_1 x_1} = (y_{x_1} - y_{\bar{x}_1}) / h_1$ — вторая разностная производная в точке \mathbf{x} по переменной x_1

Характерным примером является уравнение Пуассона

$$-\Delta u \equiv -\sum_{\alpha=1}^2 \frac{\partial^2 u}{\partial x_\alpha^2} = f(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad (11.2)$$

т. е. в уравнении (11.1) $k(\mathbf{x}) = 1$, $q(\mathbf{x}) = 0$.

Для уравнения (11.1) будем рассматривать граничные условия первого рода

$$u(\mathbf{x}) = \mu(\mathbf{x}), \quad \mathbf{x} \in \partial\Omega. \quad (11.3)$$

На границе области или ее части могут задаваться и граничные условия второго и третьего рода, например,

$$k(\mathbf{x}) \frac{\partial u}{\partial n} + \sigma(\mathbf{x})u = \mu(\mathbf{x}), \quad \mathbf{x} \in \partial\Omega,$$

где \mathbf{n} — внешняя по отношению к Ω нормаль.

11.2. Численное решение краевых задач

Приведем некоторые факты по аппроксимации краевых задач для эллиптических уравнений, сформулируем достаточные условия для выполнения принципа максимума для сеточных функций, рассмотрим вопросы оценки точности приближенного решения и проблемы решения сеточных уравнений.

11.2.1. Аппроксимация краевых задач для эллиптических уравнений

Будем использовать равномерную по каждому направлению сетку. Для сеток по отдельным направлениям x_α , $\alpha = 1, 2$ используем обозначения

$$\bar{\omega}_\alpha = \{x_\alpha \mid x_\alpha = i_\alpha h_\alpha, \quad i_\alpha = 0, 1, \dots, N_\alpha, \quad N_\alpha h_\alpha = l_\alpha\},$$

где

$$\omega_\alpha = \{x_\alpha \mid x_\alpha = i_\alpha h_\alpha, \quad i_\alpha = 1, 2, \dots, N_\alpha - 1, \quad N_\alpha h_\alpha = l_\alpha\},$$

$$\omega_\alpha^+ = \{x_\alpha \mid x_\alpha = i_\alpha h_\alpha, \quad i_\alpha = 1, 2, \dots, N_\alpha, \quad N_\alpha h_\alpha = l_\alpha\}.$$

Для сетки в Ω положим

$$\bar{\omega} = \bar{\omega}_1 \times \bar{\omega}_2 = \{\mathbf{x} \mid \mathbf{x} = (x_1, x_2), \quad x_\alpha \in \bar{\omega}_\alpha, \quad \alpha = 1, 2\},$$

$$\omega = \omega_1 \times \omega_2.$$

Для гладких коэффициентов уравнения (11.1) разностная схема строится на основе непосредственного перехода от дифференциальных операторов к разностным. Подобно одномерному случаю для краевой задачи (11.1), (11.3) поставим в соответствие разностное уравнение

$$Ly = \sum_{\alpha=1}^2 L^{(\alpha)} y = \varphi(\mathbf{x}), \quad \mathbf{x} \in \omega, \quad (11.4)$$

где

$$L^{(\alpha)} y = - (a^{(\alpha)} y_{i_\alpha})_{x_\alpha} + \theta_\alpha c(\mathbf{x}) y, \quad \alpha = 1, 2, \quad \mathbf{x} \in \omega, \quad (11.5)$$

где $\theta_1 + \theta_2 = 1$.

Для коэффициентов при старших производных можно положить

$$a^{(1)}(x) = k(x_1 - 0,5h_1, x_2), \quad x_1 \in \omega_1^+, \quad x_2 \in \omega_2,$$

$$a^{(2)}(x) = k(x_1, x_2 - 0,5h_2), \quad x_1 \in \omega_1, \quad x_2 \in \omega_2^+.$$

Для младшего коэффициента и правой части (11.4), (11.5) имеем

$$c(x) = q(x), \quad \varphi(x) = f(x), \quad x \in \omega.$$

В общем случае применяется интегро-интерполяционный метод. Интегрирование по контрольному объему для отдельного узла x сетки ω

$$\Omega_x = \{s \mid s = (s_1, s_2), \quad x_1 - 0,5h_1 \leq s_1 \leq x_1 + 0,5h_1, \\ x_2 - 0,5h_2 \leq s_2 \leq x_2 + 0,5h_2\}$$

дает, например,

$$a^{(1)}(x) = \frac{1}{h_2} \int_{x_2-0,5h_2}^{x_2+0,5h_2} \left(\frac{1}{h_1} \int_{x_1-h_1}^{x_1} \frac{ds_1}{k(s)} \right)^{-1} ds_2,$$

$$a^{(2)}(x) = \frac{1}{h_1} \int_{x_1-0,5h_1}^{x_1+0,5h_1} \left(\frac{1}{h_2} \int_{x_2-h_2}^{x_2} \frac{ds_2}{k(s)} \right)^{-1} ds_1.$$

Для граничных узлов $\partial\omega$ ($\bar{\omega} = \omega \cup \partial\omega$) используется аппроксимация

$$y(x) = \mu(x), \quad x \in \partial\omega \quad (11.6)$$

краевых условий (11.3).

11.2.2. Принцип максимума

Разностное уравнение (11.4), (11.5) запишем в виде

$$Sy(x) = \varphi(x), \quad x \in \omega, \quad (11.7)$$

где линейный оператор S определяется формулой

$$Sv(x) = A(x)v(x) - \sum_{\xi \in \mathcal{W}'(x)} B(x, \xi)v(\xi). \quad (11.8)$$

Здесь $\mathcal{W}(x)$ — шаблон, а $\mathcal{W}' \equiv \mathcal{W} \setminus \{x\}$ — окрестность узла $x \in \omega$.

Будем считать, что для рассматриваемых эллиптических уравнений второго порядка шаблон \mathcal{W} содержит узлы $(x_1 \pm h_1, x_2)$, $(x_1, x_2 \pm h_2)$ (шаблон, как минимум, пятиточечный), а коэффициенты удовлетворяют условиям

$$\begin{aligned} A(\mathbf{x}) > 0, \quad B(\mathbf{x}, \xi) > 0, \quad \xi \in \mathcal{W}'(\mathbf{x}), \\ D(\mathbf{x}) = A(\mathbf{x}) - \sum_{\xi \in \mathcal{W}'(\mathbf{x})} B(\mathbf{x}, \xi) > 0, \quad \mathbf{x} \in \omega. \end{aligned} \quad (11.9)$$

Для разностного уравнения (11.7), (11.8) при выполнении (11.9) справедлив принцип максимума. В частности, если сеточная функция $y(\mathbf{x})$, удовлетворяет граничным условиям

$$y(\mathbf{x}) = 0, \quad \mathbf{x} \in \partial\omega, \quad (11.10)$$

а правая часть

$$\varphi(\mathbf{x}) \leq 0 \quad (\varphi(\mathbf{x}) \geq 0), \quad \mathbf{x} \in \omega,$$

то $y(\mathbf{x}) \leq 0$ ($y(\mathbf{x}) \geq 0$).

На основе принципа максимума устанавливаются теоремы сравнения для решений сеточных эллиптических уравнений. Рассмотрим, например, задачу

$$\begin{aligned} Sw(\mathbf{x}) &= \phi(\mathbf{x}), \quad \mathbf{x} \in \omega, \\ w(\mathbf{x}) &= \nu(\mathbf{x}), \quad \mathbf{x} \in \partial\omega \end{aligned}$$

и пусть

$$\begin{aligned} |\varphi(\mathbf{x})| &\leq \phi(\mathbf{x}), \quad \mathbf{x} \in \omega, \\ |\mu(\mathbf{x})| &\leq \nu(\mathbf{x}), \quad \mathbf{x} \in \partial\omega. \end{aligned}$$

Тогда для решения задачи (11.6), (11.7) справедлива оценка

$$|y(\mathbf{x})| \leq w(\mathbf{x}), \quad \mathbf{x} \in \bar{\omega}.$$

Отсюда следует, что для решения однородного уравнения (11.6) ($\varphi(\mathbf{x}) = 0$, $\mathbf{x} \in \omega$) с граничными условиями (11.7) имеет место априорная оценка устойчивости

$$\max_{\mathbf{x} \in \omega} |y(\mathbf{x})| \leq \max_{\mathbf{x} \in \partial\omega} |\mu(\mathbf{x})|.$$

С привлечением подобных априорных оценок доказывается сходимость разностных схем в равномерной норме. Будем использовать для приближенного решения задачи Дирихле для уравнения Пуассона (11.2), (11.3) разностное уравнение

$$-y_{\bar{x}_1 z_1} - y_{\bar{x}_2 z_2} = \varphi(\mathbf{x}), \quad \mathbf{x} \in \omega, \quad (11.11)$$

дополнив его граничными условиями (11.6). Для погрешности $z(\mathbf{x}) = y(\mathbf{x}) - u(\mathbf{x})$, $\mathbf{x} \in \bar{\omega}$ получим задачу

$$\begin{aligned} -z_{\bar{x}_1 z_1} - z_{\bar{x}_2 z_2} &= \psi(\mathbf{x}), \quad \mathbf{x} \in \omega, \\ z(\mathbf{x}) &= 0, \quad \mathbf{x} \in \partial\omega, \end{aligned}$$

где $\psi(\mathbf{x}) = O(h_1^2 + h_2^2)$ — погрешность аппроксимации. Выбрав в качестве мажорантной функции

$$w(\mathbf{x}) = \frac{1}{4}(l_1^2 + l_2^2 - x_1^2 - x_2^2) \|\psi(\mathbf{x})\|_\infty,$$

где

$$\|v(\mathbf{x})\|_\infty = \max_{\mathbf{x} \in \omega} |v(\mathbf{x})|,$$

для погрешности получим оценку

$$\|y(\mathbf{x}) - u(\mathbf{x})\|_\infty \leq \frac{1}{4}(l_1^2 + l_2^2) \|\psi(\mathbf{x})\|_\infty.$$

Тем самым разностная схема (11.6), (11.11) сходится в $L_\infty(\omega)$ со вторым порядком.

11.2.3. Разностные уравнения в гильбертовом пространстве

Остановимся на решении уравнения (11.1) с однородными граничными условиями первого рода ($\mu(\mathbf{x}) = 0$ в (11.2)), которому ставится в соответствие разностная схема (11.4), (11.5), (11.10). Для сеточных функций, обращающихся в нуль на множестве граничных узлов $\partial\omega$, определим гильбертово пространство $H = L_2(\omega)$, в котором скалярное произведение и норма задаются следующим образом:

$$(y, w) \equiv \sum_{\mathbf{x} \in \omega} y(\mathbf{x})w(\mathbf{x})h_1 h_2, \quad \|y\| \equiv \sqrt{(y, y)}.$$

Определим для двумерных разностных функций, обращающихся в нуль на $\partial\omega$ сеточный аналог нормы в $W_2^1(\omega)$:

$$\|\nabla y\|^2 \equiv \sum_{x_1 \in \omega_1^+} \sum_{x_2 \in \omega_2} (y_{\bar{x}_1})^2 h_1 h_2 + \sum_{x_1 \in \omega_1} \sum_{x_2 \in \omega_2^+} (y_{\bar{x}_2})^2 h_1 h_2.$$

На H разностный оператор L самосопряжен и справедлива при наших предположениях о коэффициентах уравнения (11.1) оценка

$$(Ly, y) \geq \kappa \|\nabla y\|^2. \quad (11.12)$$

Для сеточных функций $y(\mathbf{x})$, обращающихся в нуль на $\partial\omega$, верно неравенство (неравенство Фридрикса для двумерных сеточных функций)

$$\|y\|^2 \leq M_0 \|\nabla y\|^2, \quad M_0^{-1} = \frac{8}{l_1^2} + \frac{8}{l_2^2}. \quad (11.13)$$

Из (11.12), (11.13) следует оценка оператора L снизу

$$L \geq \kappa \lambda_0 E, \quad \lambda_0 = M_0^{-1}. \quad (11.14)$$

Приведем также оценку оператора L сверху:

$$L \leq M_1 E \quad (11.15)$$

с постоянной

$$M_1 = \frac{4}{h_1^2} \max_{\mathbf{x} \in \omega} \frac{a^{(1)}(\mathbf{x}) + a^{(1)}(x_1 + h_1, x_2)}{2} + \\ + \frac{4}{h_2^2} \max_{\mathbf{x} \in \omega} \frac{a^{(2)}(\mathbf{x}) + a^{(2)}(x_1, x_2 + h_2)}{2} + \max_{\mathbf{x} \in \omega} |c(\mathbf{x})|.$$

Задача для погрешности разностного решения

$$z(\mathbf{x}) = y(\mathbf{x}) - u(\mathbf{x}), \quad \mathbf{x} \in \bar{\omega}$$

имеет вид

$$Lz = \psi(\mathbf{x}), \quad \mathbf{x} \in \omega,$$

где $\psi(\mathbf{x})$, как обычно, погрешность аппроксимации:

$$\psi(\mathbf{x}) = \varphi(\mathbf{x}) - Lu, \quad \mathbf{x} \in \omega.$$

Будем считать, что решение краевой задачи имеет достаточно гладкое классическое решение. На равномерной прямоугольной сетке погрешность аппроксимации в этих условиях при использовании разностного оператора (11.4), (11.5) имеет второй порядок:

$$\psi(x) = O(|h|^2), \quad |h|^2 \equiv h_1^2 + h_2^2, \quad x \in \omega.$$

Для рассматриваемой разностной схемы (11.4)–(11.6) справедлива априорная оценка для погрешности

$$\|\nabla z\| \leq \frac{M_0^{1/2}}{\kappa} \|\psi\|.$$

В силу этого разностная схема сходится в $W_2^1(\omega)$ со вторым порядком.

11.2.4. Решение сеточных уравнений

Исходная дифференциальная задача при аппроксимации заменяется сеточной. Соответствующие разностные (сеточные) уравнения есть система линейных алгебраических уравнений для неизвестных значений сеточной функции. Для их нахождения используются методы линейной алгебры, которые максимально учитывают специфику сеточных задач. Особенности сеточных задач проявляются в том, что соответствующая матрица системы алгебраических уравнений является разреженной, т. е. содержит много нулевых элементов, имеет ленточную структуру. При решении многомерных задач матрица имеет очень большой порядок, равный общему числу узлов сетки.

Классический подход к решению простейших линейных задач математической физики связан с использованием метода разделения переменных. Естественно ожидать, что аналогичная идея получит свое развитие и применительно к сеточным уравнениям. Рассмотрим сеточную задачу для уравнения Пуассона (11.10) с однородными краевыми условиями (11.11).

Для применения метода Фурье для решения этой двумерной задачи рассмотрим задачу на собственные значения для разностного оператора второй производной по переменной x_1 :

$$\begin{aligned} -v_{\bar{x}_1 x_1} + \lambda v &= 0, & x_1 \in \omega_1, \\ v_0 &= 0, & v_{N_1} = 0. \end{aligned}$$

Соответствующие собственные значения и собственные функции обозначим $\lambda_k, v^{(k)}(x_1)$, $k = 1, 2, \dots, N_1 - 1$:

$$\lambda_k = \frac{4}{h_1^2} \sin^2 \frac{k\pi h_1}{2l_1},$$

$$v^{(k)}(x_1) = \sqrt{\frac{2}{l_1}} \sin \frac{k\pi x_1}{l_1}, \quad k = 1, 2, \dots, N_1 - 1.$$

Будем искать приближенное решение задачи (11.10), (11.11) в виде разложения:

$$y(x) = \sum_{k=1}^{N_1-1} c^{(k)}(x_2) v^{(k)}(x_1), \quad x \in \omega. \quad (11.16)$$

Пусть $\varphi^{(k)}(x_2)$ — коэффициенты Фурье правой части:

$$\varphi^{(k)}(x_2) = \sum_{k=1}^{N_1-1} \varphi(x) v^{(k)}(x_1) h_1. \quad (11.17)$$

Для определения $c^{(k)}(x_2)$ получим трехточечные задачи:

$$-c_{\bar{x}_2 x_2}^{(k)} - \lambda c^{(k)} = \varphi^{(k)}(x_2), \quad x_2 \in \omega_2, \quad (11.18)$$

$$c_0^{(k)} = 0, \quad c_{N_2}^{(k)} = 0. \quad (11.19)$$

Разностная задача (11.18), (11.19) при каждом $k = 1, 2, \dots, N_1 - 1$ решается методом прогонки.

Таким образом метод Фурье основан на определении собственных функций и собственных значений одномерной сеточной задачи, вычислении коэффициентов Фурье правой части согласно (11.17), решении задач (11.18), (11.19) для коэффициентов разложения и, наконец, решение задачи определяется по формулам суммирования (11.16).

Эффективные вычислительные алгоритмы метода разделения переменных связаны с быстрым преобразованием Фурье (FFT). В этом случае можно вычислить коэффициенты Фурье правой части и восстановить решение при затратах $Q = O(N_1 N_1 \log N_1)$. Для задач с постоянными коэффициентами можно использовать преобразование Фурье по обоим переменным (разложение по собственным функциям двумерного сеточного оператора L).

Для приближенного решения многомерных сеточных эллиптических задач с переменными коэффициентами используются итерационные методы. Основные понятия теории итерационных методов решения систем линейных уравнений обсуждались выше. Здесь мы отметим только наиболее важные особенности итерационного решения краевых задач для эллиптических уравнений, которые касаются выбора оператора B (перепреобуславливателя) при переходе на новое итерационное приближение.

Для разностной задачи (11.4), (11.5), (11.10) запишем соответствующую систему линейных уравнений

$$Ay = \varphi \quad (11.20)$$

для нахождения сеточного решения $y(x)$, $x \in \omega$. Здесь A рассматривается как линейный оператор, действующий в конечномерном гильбертовом пространстве $H = L_2(\omega)$, а $\varphi(x)$ — заданный элемент H .

Для приближенного решения уравнения (11.20) с

$$A = A^* > 0$$

будем использовать двухслойный итерационный метод

$$B \frac{y^{k+1} - y^k}{\tau_{k+1}} + Ay^k = \varphi, \quad k = 0, 1, \dots \quad (11.21)$$

Особенности итерационных методов для решения эллиптических задач проявляются при построении оператора B .

Пусть априорная информация об операторах B и A задана в виде двухстороннего операторного неравенства

$$\gamma_1 B \leq A \leq \gamma_2 B, \quad \gamma_1 > 0, \quad (11.22)$$

т. е. операторы B и A энергетически эквивалентны с постоянными энергетической эквивалентности γ_α , $\alpha = 1, 2$. В итерационном методе (11.21) с оптимальным значением итерационного параметра

$$\tau = \tau_0 = \frac{2}{\gamma_1 + \gamma_2}$$

для числа итераций K , необходимых для достижения точности ϵ , справедлива оценка

$$K \geq K_0(\epsilon) = \frac{\ln \epsilon}{\ln \rho_0}, \quad (11.23)$$

где

$$\varrho_0 = \frac{1 - \xi}{1 + \xi}, \quad \xi = \frac{\gamma_1}{\gamma_2}.$$

При использовании чебышевского набора итерационных параметров и для метода сопряженных градиентов имеем

$$K \geq K_0(\varepsilon) = \frac{\ln(2\varepsilon^{-1})}{\ln \varrho_1^{-1}}, \quad (11.24)$$

где

$$\varrho_1 = \frac{1 - \xi^{1/2}}{1 + \xi^{1/2}}, \quad \xi = \frac{\gamma_1}{\gamma_2}.$$

Для явного итерационного метода $B = E$ и в силу (11.14), (11.15) ($A = L$) для постоянных энергетической эквивалентности получим

$$\gamma_1 = \kappa M_0^{-1} = O(1), \quad \gamma_2 = M_1 = O(|h|^{-2}).$$

В методе простой итерации при оптимальном значении итерационного параметра из (11.23) получим

$$K_0(\varepsilon) = O\left(\frac{1}{|h|} \ln \frac{1}{\varepsilon}\right).$$

Для метода сопряженных градиентов оценка (11.24) дает

$$K_0(\varepsilon) = O\left(\frac{1}{|h|} \ln \frac{1}{\varepsilon}\right). \quad (11.25)$$

При применении попеременно-треугольного итерационного метода используется разложение

$$A = A_1 + A_2 = A^* > 0, \quad A_2^* = A_1$$

и оператор B задается в виде

$$B = (G + \nu A_1) G^{-1} (G + \nu A_2), \quad (11.26)$$

где $G = G^* > 0$. При априорной информации

$$A \geq \delta G, \quad \delta > 0, \quad A_1 G^{-1} A_2 \leq \frac{\Delta}{4} A \quad (11.27)$$

оптимальным является выбор параметра

$$\nu = \nu_0 = \frac{2}{\sqrt{\Delta\delta}},$$

причем для числа итераций верна оценка

$$K \geq K_0(\varepsilon) = \frac{1}{2\sqrt{2}\sqrt[4]{\eta}} \ln \frac{2}{\varepsilon}, \quad \eta = \frac{\delta}{\Delta} \quad (11.28)$$

при использовании чебышевского набора итерационных параметров или метода сопряженных градиентов.

Для эллиптических уравнений второго порядка имеет место следующая зависимость от шагов сетки

$$\delta = O(1), \quad \Delta = O(|h|^{-2}).$$

Поэтому для числа итераций попеременно-треугольного итерационного метода получим

$$K_0(\varepsilon) = O\left(\frac{1}{\sqrt{|h|}} \ln \frac{1}{\varepsilon}\right).$$

Оптимизация метода достигается за счет выбора оператора $D = D^* > 0$.

11.3. Упражнения

Приведены примеры аппроксимации эллиптического уравнения второго порядка, исследуются свойства разностной задачи и обсуждаются вопросы итерационного решения сеточных эллиптических задач.

Упражнение 11.1. Постройте схему повышенного порядка аппроксимации для решения задачи Дирихле для уравнения Пуассона.

Решение. Для разностного оператора второй производной по отдельному направлению имеем

$$u_{\bar{x}_\alpha \bar{x}_\alpha} = \frac{\partial^2 u}{\partial x_\alpha^2} + \frac{h_\alpha^2}{12} \frac{\partial^4 u}{\partial x_\alpha^4} + O(h_\alpha^4), \quad \alpha = 1, 2$$

и поэтому

$$-u_{\bar{x}_1 x_1} - u_{\bar{x}_2 x_2} = -\frac{\partial^2 u}{\partial x_1^2} - \frac{\partial^2 u}{\partial x_2^2} - \frac{h_1^2}{12} \frac{\partial^4 u}{\partial x_1^4} - \frac{h_2^2}{12} \frac{\partial^4 u}{\partial x_2^4} + O(|h|^4).$$

На решениях уравнения (11.2) получим

$$-\frac{\partial^4 u}{\partial x_1^4} = \frac{\partial^4 u}{\partial x_1^2 \partial x_2^2} + \frac{\partial^2 f}{\partial x_1^2}, \quad -\frac{\partial^4 u}{\partial x_2^4} = \frac{\partial^4 u}{\partial x_1^2 \partial x_2^2} + \frac{\partial^2 f}{\partial x_2^2}$$

и тем самым

$$\begin{aligned} -u_{\bar{x}_1 x_1} - u_{\bar{x}_2 x_2} &= -\frac{\partial^2 u}{\partial x_1^2} - \frac{\partial^2 u}{\partial x_2^2} + \frac{h_1^2 + h_2^2}{12} \frac{\partial^4 u}{\partial x_1^2 \partial x_2^2} + \\ &+ \frac{h_1^2}{12} \frac{\partial^2 f}{\partial x_1^2} + \frac{h_2^2}{12} \frac{\partial^2 f}{\partial x_2^2} + O(|h|^4). \end{aligned}$$

Используем для смешанной производной аппроксимацию

$$\frac{\partial^4 u}{\partial x_1^2 \partial x_2^2} = u_{\bar{x}_1 x_1 \bar{x}_2 x_2}$$

и приходим к разностному уравнению

$$\begin{aligned} -y_{\bar{x}_1 x_1} - y_{\bar{x}_2 x_2} - \frac{h_1^2 + h_2^2}{12} y_{\bar{x}_1 x_1 \bar{x}_2 x_2} &= \varphi(\mathbf{x}), \quad \mathbf{x} \in \omega, \\ \varphi(\mathbf{x}) &= f(\mathbf{x}) + \frac{h_1^2}{12} f_{\bar{x}_1 x_1} + \frac{h_2^2}{12} f_{\bar{x}_2 x_2}. \end{aligned} \tag{11.29}$$

Схема (11.6), (11.29) аппроксимирует краевую задачу (11.2), (11.3) с четвертым порядком.

Упражнение 11.2. Рассмотрите аппроксимацию эллиптического уравнения второго порядка со смешанными производными

$$-\sum_{\alpha, \beta=1}^2 \frac{\partial}{\partial x_\alpha} \left(k_{\alpha\beta}(\mathbf{x}) \frac{\partial u}{\partial x_\beta} \right) = f(\mathbf{x}), \quad \mathbf{x} \in \Omega,$$

в котором

$$k_{\alpha\beta}(\mathbf{x}) = k_{\beta\alpha}(\mathbf{x}), \quad \alpha, \beta = 1, 2.$$

Решение. Рассмотрим разностную схему

$$Ly = \sum_{\alpha, \beta=1}^2 L^{(\alpha, \beta)} y = \varphi(x), \quad x \in \omega,$$

в которой

$$L^{(\alpha, \beta)} = L_-^{(\alpha, \beta)} + L_+^{(\alpha, \beta)},$$

$$L_-^{(\alpha, \beta)} y = -\frac{1}{2}(k_{\alpha\beta} y_{\bar{x}\beta})_{x_\alpha}, \quad L_+^{(\alpha, \beta)} y = -\frac{1}{2}(k_{\alpha\beta} y_{x_\beta})_{\bar{x}_\alpha}.$$

При $\alpha = \beta$ имеем

$$L^{(\alpha, \alpha)} u = \frac{\partial}{\partial x_\alpha} \left(k_{\alpha\alpha} \frac{\partial u}{\partial x_\alpha} \right) + O(|h|^2), \quad \alpha = 1, 2.$$

Формула Тейлора дает нам

$$u_{\bar{x}_2} = \frac{\partial u}{\partial x_2}(x) - \frac{h_2}{2} \frac{\partial^2 u}{\partial x_2^2}(x) + O(h_2^2),$$

$$v_{x_1} = \frac{\partial u}{\partial x_1}(x) + \frac{h_1}{2} \frac{\partial^2 u}{\partial x_1^2}(x) + O(h_1^2).$$

Подстановкой $v = k_{12} u_{\bar{x}_2}$ получаем

$$\begin{aligned} L_-^{(12)} u &= -\frac{\partial}{\partial x_1} \left(k_{12} \frac{\partial u}{\partial x_2} \right) - \frac{h_1}{2} \frac{\partial^2}{\partial x_1^2} \left(k_{12} \frac{\partial u}{\partial x_2} \right) + \\ &+ \frac{h_2}{2} \frac{\partial^2}{\partial x_1 x_2} \left(k_{12} \frac{\partial u}{\partial x_2} \right) + O(|h|^2). \end{aligned}$$

Аналогичные выкладки приводят к

$$\begin{aligned} L_+^{(12)} u &= -\frac{\partial}{\partial x_1} \left(k_{12} \frac{\partial u}{\partial x_2} \right) + \frac{h_1}{2} \frac{\partial^2}{\partial x_1^2} \left(k_{12} \frac{\partial u}{\partial x_2} \right) - \\ &- \frac{h_2}{2} \frac{\partial^2}{\partial x_1 x_2} \left(k_{12} \frac{\partial u}{\partial x_2} \right) + O(|h|^2) \end{aligned}$$

и поэтому

$$L^{(12)} u = -\frac{\partial}{\partial x_1} \left(k_{12} \frac{\partial u}{\partial x_2} \right) + O(|h|^2).$$

Тем самым рассматриваемая разностная схема имеет второй порядок аппроксимации.

Упражнение 11.3. *Покажите, что для решения задачи (11.7)–(11.10) при $D(\mathbf{x}) > 0$, $\mathbf{x} \in \omega$ справедлива оценка*

$$\|y(\mathbf{x})\|_{\infty} \leq \left\| \frac{\varphi(\mathbf{x})}{D(\mathbf{x})} \right\|_{\infty}.$$

Решение. Рассмотрим сеточную функцию (мажоранту) $w(\mathbf{x})$, которая определяется как решение задачи

$$\begin{aligned} Sw(\mathbf{x}) &= |\varphi(\mathbf{x})|, \quad \mathbf{x} \in \omega, \\ w(\mathbf{x}) &= 0, \quad \mathbf{x} \in \partial\omega. \end{aligned}$$

Пусть $w(\mathbf{x}) \geq 0$ принимает максимальное значение в некотором узле $\mathbf{x}^* \in \omega$. В этом узле в силу (11.8), (11.9) имеем

$$D(\mathbf{x}^*)w(\mathbf{x}^*) - \sum_{\xi \in W'(\mathbf{x}^*)} B(\mathbf{x}^*, \xi)(w(\xi) - w(\mathbf{x}^*)) = |\varphi(\mathbf{x}^*)|.$$

Отсюда следует

$$D(\mathbf{x}^*)w(\mathbf{x}^*) \leq |\varphi(\mathbf{x}^*)|$$

и тем самым

$$\|w(\mathbf{x})\|_{\infty} = w(\mathbf{x}^*) \leq \frac{|\varphi(\mathbf{x}^*)|}{D(\mathbf{x}^*)} \leq \left\| \frac{\varphi(\mathbf{x})}{D(\mathbf{x})} \right\|_{\infty}.$$

Но функция $w(\mathbf{x})$ является мажорантой для рассматриваемой задачи и поэтому приходим к доказываемому утверждению.

Упражнение 11.4. *Пусть расчетная область G лежит в прямоугольнике Ω , в котором введена сетка $\bar{\omega}$. Граница G состоит из отрезков, соединяющих узлы этой сетки. Получите оценки снизу и сверху для разностного оператора Лапласа в области G для функций, обращающихся в нуль на ∂G .*

Решение. Подмножество узлов, лежащих в G обозначим $\omega(G)$ и рассмотрим разностный оператор

$$Ly = -y_{\bar{x}_1 x_1} - y_{\bar{x}_2 x_2} = \varphi(\mathbf{x}), \quad \mathbf{x} \in \omega(G),$$

на множестве сеточных функций

$$y(x) = 0, \quad x \in \partial\omega(G).$$

Доопределим функцию y , положим

$$\tilde{y}(x) = \begin{cases} y(x), & x \in \omega(G), \\ 0, & x \in \omega \setminus \omega(G). \end{cases}$$

В силу такого продолжения сеточной функции при использовании обозначений

$$(v, y)_{\omega(G)} = \sum_{x \in \omega(G)} v(x)y(x)h_1h_2$$

получим

$$\|\tilde{y}\| = \|y\|_{\omega(G)}, \quad (L\tilde{y}, \tilde{y}) = (Ly, y)_{\omega(G)}.$$

Для сеточной функции заданной в Ω имеет место

$$\delta\|\tilde{y}\|^2 \leq (L\tilde{y}, \tilde{y}) \leq \Delta\|\tilde{y}\|^2,$$

где

$$\delta = \frac{8}{l_1^2} + \frac{8}{l_2^2}, \quad \Delta = \frac{4}{h_1^2} + \frac{4}{h_2^2}.$$

Аналогичная оценка

$$\delta\|y\|_{\omega(G)}^2 \leq (Ly, y)_{\omega(G)} \leq \Delta\|y\|_{\omega(G)}^2$$

имеет место и для разностного оператора в области G .

Упражнение 11.5. Рассмотрите попеременно-треугольный метод для итерационного решения разностной задачи Дирихле для уравнения Пуассона (11.10), (11.11).

Решение. Разностный оператор Лапласа представляется в виде

$$A = A_1 + A_2 = A^* > 0, \quad A_2^* = A_1,$$

где

$$A_1y = \frac{1}{h_1}y_{\bar{x}_1} + \frac{1}{h_2}y_{\bar{x}_2},$$

$$A_2y = -\frac{1}{h_1}y_{x_1} - \frac{1}{h_2}y_{x_2}.$$

Будем использовать вариант попеременно-треугольного метода, когда (см. (11.26))

$$B = (E + \nu A_1)(E + \nu A_2).$$

Оптимальное значение параметра

$$\nu = \nu_0 = \frac{2}{\sqrt{\Delta \delta}},$$

где в соответствии с (11.27)

$$A \geq \delta E, \quad \delta > 0, \quad A_1 A_2 \leq \frac{\Delta}{4} A.$$

Здесь δ есть минимальное собственное значение оператора Лапласа, которое равно

$$\delta = \frac{4}{h_1^2} \sin^2 \frac{\pi h_1}{2l_1} + \frac{4}{h_2^2} \sin^2 \frac{\pi h_2}{2l_2}.$$

Принимая во внимание, что

$$(Ay, y) = \|\nabla y\|^2 \geq \|y_{x_1}\|^2 + \|y_{x_2}\|^2,$$

получим

$$\begin{aligned} (A_1 A_2 y, y) &= \|A_2 y\|^2 = \left\| \frac{y_{x_1}}{h_1} + \frac{y_{x_2}}{h_2} \right\|^2 = \left(\left(\frac{y_{x_1}}{h_1} + \frac{y_{x_2}}{h_2} \right), 1 \right) \leq \\ &\leq \left(\left(\frac{1}{h_1^2} + \frac{1}{h_2^2} \right) (y_{x_1}^2 + y_{x_2}^2), 1 \right) \leq \left(\frac{1}{h_1^2} + \frac{1}{h_2^2} \right) (Ay, y), \end{aligned}$$

т. е.

$$\Delta = \frac{4}{h_1^2} + \frac{4}{h_2^2}.$$

Подстановка в (11.28) дает оценку

$$K \geq K_0(\varepsilon) \approx \frac{0,29}{\sqrt{h}} \ln \frac{2}{\varepsilon}$$

при решении задачи в единичном квадрате на квадратной сетке $h_1 = h_2 = h$.

Упражнение 11.6. Пусть в итерационном методе (11.21)

$$A = A_1 + A_2, \quad A_1 A_2 = A_2 A_1,$$

$$\delta_\alpha E \leq A_\alpha \leq \Delta_\alpha E, \quad A_\alpha = A_\alpha^*, \quad \delta_\alpha > 0, \quad \alpha = 1, 2,$$

а оператор B представлен в факторизованном виде

$$B = (E + \nu A_1)(E + \nu A_2).$$

Найдите оптимальное значение параметра ν .

Решение. Скорость сходимости определяется постоянными γ_1, γ_2 в неравенстве (11.22). Поэтому сначала найдем их и выберем ν из условия максимума отношения γ_1/γ_2 .

В силу перестановочности операторов $A_\alpha, \alpha = 1, 2$ имеем

$$\begin{aligned} B &= E - \nu(A_1 + A_2) + \nu^2 A_1 A_2 + 2\nu(A_1 + A_2) = \\ &= (E - \nu A_1)(E - \nu A_2) + 2\nu A \end{aligned}$$

и поэтому $\gamma_2 = 1/(2\nu)$. Пусть

$$\delta = \min_{\alpha} \delta_{\alpha}, \quad \Delta = \max_{\alpha} \Delta_{\alpha},$$

тогда

$$B = E + \nu(A_1 + A_2) + \nu^2 A_1 A_2 \leq \left(\frac{1}{\delta} + \nu + \nu^2 \Delta \right) A.$$

Следовательно

$$\gamma_1 = \frac{\delta}{1 + \nu\delta + \nu^2\delta\Delta}$$

и максимум γ_1/γ_2 достигается при

$$\nu = \nu_0 = \frac{1}{\sqrt{\Delta\delta}}.$$

Рассматриваемый вариант соответствует использованию метода переменных направлений в сеточной эллиптической задаче с разделяющимися переменными.

11.4. Задачи

Задача 11.1. Уравнения Пуассона в круговом цилиндре при использовании цилиндрических координат записывается в виде

$$-\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u}{\partial r} \right) - \frac{1}{r^2} \frac{\partial^2 u}{\partial \varphi^2} - \frac{\partial^2 u}{\partial z^2} = f(r, \varphi, z).$$

Постройте разностную схему для этого уравнения с граничными условиями первого рода на поверхности цилиндра.

Задача 11.2. Постройте разностную схему для решения краевой задачи (11.1), (11.3) с условиями сопряжения при $x_1 = x_1^*$:

$$u(x_1^* + 0, x_2) - u(x_1^* - 0, x_2) = 0,$$

$$k \frac{\partial u}{\partial x_1}(x_1^* + 0, x_2) - k \frac{\partial u}{\partial x_1}(x_1^* - 0, x_2) = \chi(x_2).$$

Задача 11.3. Пусть в прямоугольнике $\bar{\Omega}$ с равномерной сеткой $\bar{\omega}$ проведено разбиение на треугольники диагоналями от узла (i, j) к узлу $(i + 1, j + 1)$. Постройте для задачи (11.2), (11.3) схему конечных элементов с использованием непрерывных, линейных на каждом треугольнике функций.

Задача 11.4. Рассмотрите аппроксимацию уравнения Пуассона (11.2) с условиями (11.3) в области с криволинейной границей с использованием неравномерной вблизи границы сетки.

Задача 11.5. Сформулируйте достаточные условия выполнения принципа максимума для схемы повышенного порядка аппроксимации для задачи (11.2), (11.3).

Задача 11.6. На основе принципа максимума проведите исследования скорости сходимости для разностной схемы для уравнения (уравнение Пуассона в круге)

$$-\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u}{\partial r} \right) - \frac{1}{r^2} \frac{\partial^2 u}{\partial \varphi^2} = f(r, \varphi), \quad 0 < r < R, \quad 0 \leq \varphi \leq 2\pi$$

с граничными условиями

$$\lim_{r \rightarrow 0} r \frac{\partial u}{\partial r}(r, \varphi) = 0, \quad u(R, \varphi) = \mu(\varphi)$$

и условиями периодичности

$$u(r, \varphi) = u(r, 2\pi + \varphi), \quad 0 < r < R, \quad 0 \leq \varphi < 2\pi.$$

Задача 11.7. Получите оценки снизу и сверху для разностного оператора (см. (11.29))

$$Ly = -y_{\bar{x}_1 x_1} - y_{\bar{x}_2 x_2} - \frac{h_1^2 + h_2^2}{12} y_{\bar{x}_1 x_1 \bar{x}_2 x_2}, \quad x \in \omega,$$

определенного на множестве функций $y(x) = 0$, $x \notin \omega$.

Задача 11.8. Постройте схему повышенного порядка аппроксимации для задачи Неймана для уравнения Пуассона:

$$\begin{aligned} -\sum_{\alpha=1}^2 \frac{\partial^2 u}{\partial x_\alpha^2} &= f(x), \quad x \in \Omega, \\ -\frac{\partial u}{\partial x_1} &= \mu_-^{(1)}(x_2), \quad \frac{\partial u}{\partial x_1} = \mu_+^{(1)}(x_2), \quad 0 < x_2 < l_2, \\ -\frac{\partial u}{\partial x_2} &= \mu_-^{(2)}(x_1), \quad \frac{\partial u}{\partial x_2} = \mu_+^{(2)}(x_1), \quad 0 < x_1 < l_1. \end{aligned}$$

Задача 11.9. Аппроксимируйте граничные условия третьего рода

$$-k(0, x_2) \frac{\partial u}{\partial x_1} + \sigma(x_2)u(0, x_2) = \mu(x_2),$$

заданное на части границы прямоугольника Ω (на других участках границы — граничные условия первого рода), при численном решении уравнения (11.1).

Задача 11.10. Постройте разностную схему для приближенного решения уравнения четвертого порядка

$$\frac{\partial^4 u}{\partial x_1^4} + 2 \frac{\partial^4 u}{\partial x_1^2 \partial x_2^2} + \frac{\partial^4 u}{\partial x_2^4} = f(x), \quad x \in \Omega$$

в прямоугольнике $\bar{\Omega}$, когда на границе задана сама функция $u(x)$ и ее нормальная производная (первая краевая задача).

Задача 11.11. Рассмотрите аппроксимации на равномерной прямоугольной сетке системы уравнений теории упругости (уравнения Ламэ)

$$\begin{aligned} -\frac{\partial}{\partial x_1} \left((2\mu + \lambda) \frac{\partial u_1}{\partial x_1} \right) - \frac{\partial}{\partial x_2} \left(\mu \frac{\partial u_1}{\partial x_2} \right) - \\ - \frac{\partial}{\partial x_1} \left(\lambda \frac{\partial u_2}{\partial x_2} \right) - \frac{\partial}{\partial x_2} \left(\mu \frac{\partial u_2}{\partial x_1} \right) = f_1(x), \end{aligned}$$

$$\begin{aligned}
 & -\frac{\partial}{\partial x_2} \left(\lambda \frac{\partial u_1}{\partial x_1} \right) - \frac{\partial}{\partial x_1} \left(\mu \frac{\partial u_1}{\partial x_2} \right) - \frac{\partial}{\partial x_1} \left(\mu \frac{\partial u_2}{\partial x_1} \right) - \\
 & - \frac{\partial}{\partial x_2} \left((2\mu + \lambda) \frac{\partial u_2}{\partial x_2} \right) = f_2(\mathbf{x}), \quad \mathbf{x} \in \Omega.
 \end{aligned}$$

Задача 11.12. Для решения сеточной задачи

$$Ay = \varphi, \quad A = A^* > 0$$

используется треугольный итерационный метод

$$(D + \tau A_1) \frac{y^{k+1} - y^k}{\tau} + Ay^k = \varphi,$$

где D — произвольный самосопряженный оператор, а

$$A = A_1 + A_2, \quad A_1 = A_2^*.$$

Найдите оптимальное значение итерационного параметра τ , когда априорная информация задана в виде

$$\delta D \leq A, \quad A_1 D^{-1} A_2 \leq \frac{\Delta}{4} A.$$

Задача 11.13. Получите оценки скорости сходимости метода верхней релаксации при решении сеточной задачи Дирихле для уравнения Пуассона (11.6), (11.11).

Задача 11.14. Пусть A — разностный оператор, который определен на $H = L_2(\omega)$ и соответствует разностной задаче (11.2)–(11.6). Пусть B есть диагональная часть оператора A . Покажите, что для постоянных энергетической эквивалентности (см. неравенство (11.22)) операторов A и B верно равенство

$$\gamma_1 + \gamma_2 = 2.$$

Задача 11.15. Рассмотрите итерационный метод (11.21) для приближенного решения сеточной эллиптической задачи с переменными коэффициентами (11.4)–(11.6), когда в качестве переобуславливателя B берется сеточный оператор, соответствующий решению разностного уравнения

$$-y_{\bar{x}_1 x_1} - y_{\bar{x}_2 x_2} + \chi y = \varphi(\mathbf{x}), \quad \mathbf{x} \in \omega,$$

т. е. разностной задачи с постоянными коэффициентами.

Задача 11.16. Пусть решается сеточная задача

$$Ay = \varphi,$$

с несамосопряженным положительным оператором

$$A = A_0 + A_1 > 0, \quad A_0 = A_0^*, \quad A_1 = -A_1^*.$$

Проводится симметризация задачи следующим образом:

$$\bar{A}y = \bar{\varphi}, \quad \bar{A} = A^* A_0^{-1} A, \quad \bar{\varphi} = A^* A_0^{-1} \varphi.$$

Исследуйте скорость сходимости итерационного метода

$$A_0 \frac{y^{k+1} - y^k}{\tau_{k+1}} + \bar{A}y^k = \bar{\varphi}, \quad k = 0, 1, \dots,$$

когда

$$\|A_y\|^2 \leq M(A_0 y, y),$$

т. е. в условиях подчиненности кососимметричной части оператора A .

Глава 12

Нестационарные задачи математической физики

Рассматриваются разностные методы приближенного решения краевых задач для нестационарных уравнений с частными производными. Основное внимание уделяется построению и исследованию разностных схем для параболических уравнений второго порядка. Теоретической основой при исследовании сходимости разностных схем является общая теория устойчивости операторно-разностных схем. Приведены основные результаты об устойчивости двух- и трехслойных разностных схем по начальным данным и правой части. Отмечаются особенности исследования схем для гиперболических уравнений второго порядка. Строятся экономичные разностные схемы для приближенного решения многомерных нестационарных задач математической физики.

12.1. Нестационарные краевые задачи

В качестве базового нестационарного уравнения математической физики выступает одномерное параболическое уравнение второго порядка. В прямоугольнике

$$\bar{Q}_T = \bar{\Omega} \times [0, T], \quad \bar{\Omega} = \{x \mid 0 \leq x \leq l\}, \quad 0 \leq t \leq T$$

рассматривается уравнение

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(k(x) \frac{\partial u}{\partial x} \right) + f(x, t), \quad 0 < x < l, \quad 0 < t \leq T. \quad (12.1)$$

Оно дополняется (первая краевая задача) граничными

$$u(0, t) = 0, \quad u(l, t) = 0, \quad 0 < t \leq T \quad (12.2)$$

Основные обозначения

$u = u(x, t)$	— неизвестная функция
h	— шаг равномерной сетки по переменной x
τ	— шаг сетки по времени t
ω_h	— множество внутренних узлов по пространству
$\partial\omega_h$	— множество граничных узлов
H	— гильбертово пространство сеточных функций
(\cdot, \cdot)	— скалярное произведение в H
$\ \cdot\ $	— норма в H
$y_x = (y(x+h, t) - y(x, t))/h$	— правая разностная производная в точке (x, t) по переменной x
$y_{\bar{x}} = (y(x, t) - y(x-h, t))/h$	— левая разностная производная в точке (x, t) по переменной x
y^n	— решение на момент времени $t = t_n$
$y_t = (y^{n+1} - y^n)/\tau$	— производная вперед по переменной t
$y_{\bar{t}} = (y^n - y^{n-1})/\tau$	— производная назад по переменной t

и начальными

$$u(x, 0) = v^0(x), \quad 0 \leq x \leq l \quad (12.3)$$

условиями.

Для простоты изложения мы ограничились однородными граничными условиями и зависимостью коэффициента k только от пространственной переменной, причем $k(x) \geq \kappa > 0$.

Вместо условий первого рода (12.2) могут задаваться другие граничные условия. Например, во многих прикладных задачах необходимо

ориентироваться на формулирование граничных условий третьего рода:

$$\begin{aligned} -k(0) \frac{du}{dx}(0, t) + \sigma_1(t) u(0, t) &= \mu_1(t), \\ k(l) \frac{du}{dx}(l, t) + \sigma_2(t) u(l, t) &= \mu_2(t), \quad 0 < t \leq T. \end{aligned} \quad (12.4)$$

Среди других нестационарных краевых задач необходимо выделить гиперболическое уравнение второго порядка. В одномерном по пространству случае ищется решение уравнения

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x} \left(k(x) \frac{\partial u}{\partial x} \right) + f(x, t), \quad 0 < x < l, \quad 0 < t \leq T. \quad (12.5)$$

Для однозначного определения решения этого уравнения помимо граничных условий (12.2) задаются два начальных условия

$$u(x, 0) = u_0(x), \quad \frac{\partial u}{\partial t}(0, t) = u_1(x), \quad 0 \leq x \leq l. \quad (12.6)$$

Особое внимание необходимо уделять методам численного решения многомерных нестационарных задач математической физики. Примером может служить двумерное параболическое уравнение. Будем искать в прямоугольнике

$$\Omega = \{x \mid x = (x_1, x_2), \quad 0 < x_\alpha < l_\alpha, \quad \alpha = 1, 2\}$$

функцию $u(x, t)$, удовлетворяющую уравнению

$$\begin{aligned} \frac{\partial u}{\partial t} = \sum_{\alpha=1}^2 \frac{\partial}{\partial x_\alpha} \left(k(x) \frac{\partial u}{\partial x_\alpha} \right) - q(x, t) u + f(x, t), \\ x \in \Omega, \quad 0 \leq t \leq T \end{aligned} \quad (12.7)$$

и условиям

$$u(x, t) = 0, \quad x \in \partial\Omega, \quad 0 < t \leq T, \quad (12.8)$$

$$u(x, 0) = u_0(x), \quad x \in \Omega. \quad (12.9)$$

Аналогично формулируются и другие нестационарные многомерные краевые задачи для уравнений с частными производными.

12.2. Разностные методы решения нестационарных задач

Прежде чем проводить исследование конкретных разностных схем для нестационарных уравнений математической физики введем базовые понятия теории устойчивости операторно-разностных схем, рассматриваемых в конечномерных гильбертовых пространствах. На основе приведенных оценок устойчивости двух- и трехслойных разностных схем по начальным данным и правой части проводится исследование разностных схем для уравнений параболического и гиперболического типов.

12.2.1. Устойчивость двухслойных операторно-разностных схем

Дадим некоторые основные понятия и определения общей теории устойчивости операторно-разностных схем. Пусть задано вещественное конечномерное гильбертово пространство H и сетка по времени

$$\bar{\omega}_\tau = \omega_\tau \cup \{T\} = \{t_n = n\tau, \quad n = 0, 1, \dots, N_0; \quad \tau N_0 = T\}.$$

Обозначим через $A, B: H \rightarrow H$ линейные операторы в H и пусть они, для простоты, не зависят от τ, t_n . Рассмотрим задачу Коши для операторно-разностного уравнения

$$B \frac{y^{n+1} - y^n}{\tau} + Ay^n = \varphi^n, \quad t_n \in \omega_\tau, \quad (12.10)$$

$$y^0 = u_0, \quad (12.11)$$

где $y^n = y(t_n) \in H$ — искомая функция, а $\varphi^n, u_0 \in H$ — заданы.

Определим двухслойную разностную схему как множество задач Коши (12.10), (12.11), зависящих от параметра τ , а запись (12.10), (12.11) будем называть канонической формой двухслойных схем.

Двухслойная схема называется устойчивой, если существуют такие положительные постоянные m_1 и m_2 , не зависящие от τ и выбора u_0, φ ,

что при любых $u_0 \in H$, $\varphi \in H$, $t \in \bar{\omega}_\tau$ для решения задачи (12.10), (12.11) справедлива оценка

$$\|y^{n+1}\| \leq m_1 \|u_0\| + m_2 \max_{0 \leq k \leq n} \|\varphi^k\|_*, \quad t_n \in \omega_\tau, \quad (12.12)$$

где $\|\cdot\|$ и $\|\cdot\|_*$ — некоторые нормы в пространстве H .

Неравенство (12.12) отражает свойство непрерывной зависимости решения задачи (12.10), (12.11) от входных данных. Обычно разделяют понятия устойчивости по начальным данным и устойчивости по правой части.

Разностная схема

$$B \frac{y^{n+1} - y^n}{\tau} + Ay^n = 0, \quad t_n \in \omega_\tau, \quad (12.13)$$

$$y^0 = u_0 \quad (12.14)$$

называется устойчивой по начальным данным, если для решения задачи (12.13), (12.14) выполняется оценка

$$\|y^{n+1}\| \leq m_1 \|u_0\|, \quad t_n \in \omega_\tau. \quad (12.15)$$

Двухслойная разностная схема

$$B \frac{y^{n+1} - y^n}{\tau} + Ay^n = \varphi^n, \quad t_n \in \omega_\tau, \quad (12.16)$$

$$y^0 = 0 \quad (12.17)$$

устойчива по правой части, если для решения выполняется неравенство

$$\|y^{n+1}\| \leq m_2 \max_{0 \leq k \leq n} \|\varphi^k\|_*, \quad t_n \in \omega_\tau. \quad (12.18)$$

Получение оценок устойчивости чаще всего базируется на априорных оценках разностного решения при переходе с одного временного слоя на другой. Для самосопряженного положительного оператора R через H_R обозначим гильбертово пространство, состоящее из элементов пространства H и снабженное скалярным произведением и нормой

$$(y, w)_R = (Ry, w), \quad \|y\|_R = \sqrt{(Ry, y)}.$$

Разностная схема (12.13), (12.14) называется ϱ -устойчивой (равномерно устойчивой) по начальным данным в H_R , если существуют постоянная $\varrho > 0$ и постоянная m_1 , не зависящие от τ , n , такие, что при любых n и при всех $y^n \in H$ для решения y^{n+1} разностного уравнения (12.13) справедлива оценка

$$\|y^{n+1}\|_R \leq \varrho \|y^n\|_R, \quad t_n \in \omega_\tau, \quad (12.19)$$

причем $\varrho^n \leq m_1$.

В теории разностных схем в качестве константы ϱ выбирается обычно одна из величин

$$\begin{aligned} \varrho &= 1, \\ \varrho &= 1 + c\tau, \quad c > 0, \\ \varrho &= \exp(c\tau), \end{aligned}$$

где постоянная c не зависит от τ , n .

Из оценки разностного решения на слое

$$\|y^{n+1}\| \leq \varrho \|y^n\| + \tau \|\varphi^n\|_*$$

следует априорная оценка (разностный аналог леммы Гронуолла)

$$\|y_{n+1}\| \leq \varrho^{n+1} \|y_0\| + \sum_{k=0}^n \tau \varrho^{n-k} \|\varphi_k\|_*.$$

Сформулируем основные критерии устойчивости двухслойных операторно-разностных схем по начальным данным. Основным является следующий результат о точных (совпадающих необходимых и достаточных) условиях устойчивости в H_A .

Пусть в уравнении (12.13) оператор A является самосопряженным положительным оператором. Условие

$$B \geq \frac{\tau}{2} A, \quad t \in \omega_\tau \quad (12.20)$$

необходимо и достаточно для устойчивости в H_A , т.е. для выполнения оценки

$$\|y_{n+1}\|_A \leq \|u_0\|_A, \quad t \in \omega_\tau. \quad (12.21)$$

При рассмотрении общих нестационарных задач необходимо ориентироваться на условия ϱ -устойчивости.

Пусть

$$A = A^*, \quad B = B^* > 0,$$

тогда условия

$$\frac{1 - \varrho}{\tau} B \leq A \leq \frac{1 + \varrho}{\tau} B \quad (12.22)$$

необходимы и достаточны для ϱ -устойчивости в H_B схемы (12.13), (12.14), т. е. для выполнения

$$\|y_{n+1}\|_B \leq \varrho \|y_n\|_B.$$

Из устойчивости разностной схемы по начальным данным в H_R , $R = R^* > 0$ следует и устойчивость схемы по правой части. Более точно это утверждение формулируется следующим образом.

Пусть разностная схема (12.10), (12.11) ϱ -устойчива в H_R по начальным данным, т. е. имеет место оценка (12.19) при $\varphi^n = 0$. Тогда разностная схема (12.10), (12.11) устойчива по правой части и для решения справедлива априорная оценка

$$\|y^{n+1}\|_R \leq \varrho^{n+1} \|u_0\|_R + \sum_{k=0}^n \tau \varrho^{n-k} \|B^{-1} \varphi^k\|_R. \quad (12.23)$$

Приведем оценку устойчивости по начальным данным и правой части при загрузлении критерия устойчивости (12.20).

Пусть A — самосопряженный и положительный оператор, а B удовлетворяет условию

$$B \geq \frac{1 + \varepsilon}{2} \tau A \quad (12.24)$$

с некоторой постоянной $\varepsilon > 0$, не зависящей от τ . Тогда для разностной схемы (12.10), (12.11) справедлива априорная оценка

$$\|y^{n+1}\|_A^2 \leq \|u_0\|_A^2 + \frac{1 + \varepsilon}{2\varepsilon} \sum_{k=0}^n \tau \|\varphi^k\|_{B^{-1}}^2. \quad (12.25)$$

Оценки устойчивости по правой части используются при исследовании точности разностных схем для нестационарных задач.

12.2.2. Устойчивость трехслойных разностных схем

При приближенном решении нестационарных задач математической физики наряду с двухслойными разностными схемами часто используют и трехслойные. Здесь мы формулируем некоторые основные условия устойчивости трехслойных операторно-разностных схем.

Используется следующая каноническая форма трехслойных разностных схем:

$$B \frac{y^{n+1} - y^{n-1}}{2\tau} + R(y^{n+1} - 2y^n + y^{n-1}) + Ay^n = \varphi^n, \quad (12.26)$$

$$n = 1, 2, \dots$$

при заданных

$$y^0 = u_0, \quad y^1 = u_1. \quad (12.27)$$

Сформулируем условия устойчивости по начальным данным при постоянных, не зависящих от n , самосопряженных операторах A , B , R , т. е. вместо (12.26) будем рассматривать

$$B \frac{y^{n+1} - y^{n-1}}{2\tau} + R(y^{n+1} - 2y^n + y^{n-1}) + Ay^n = 0, \quad (12.28)$$

$$n = 1, 2, \dots$$

При выполнении условий

$$B \geq 0, \quad A > 0, \quad R > \frac{1}{4}A \quad (12.29)$$

для разностной схемы (12.27), (12.29) имеет место априорная оценка

$$\begin{aligned} \frac{1}{4} \|y_{n+1} + y_n\|_A^2 + \|y_{n+1} - y_n\|_R^2 - \frac{1}{4} \|y_{n+1} - y_n\|_A^2 &\leq \\ &\leq \frac{1}{4} \|y_n + y_{n-1}\|_A^2 + \|y_n - y_{n-1}\|_R^2 - \frac{1}{4} \|y_n - y_{n-1}\|_A^2, \end{aligned} \quad (12.30)$$

т. е. операторно-разностная схема (12.27), (12.29) устойчива по начальным данным.

Устойчивость рассматриваемых трехслойных операторно-разностных схем установлена в гильбертовых пространствах со сложной составной

нормой (см. (12.30)). Можно получить оценки устойчивости в более простых нормах за счет несколько более жестких условий устойчивости.

Пусть в операторно-разностной схеме (12.28) операторы R и A являются самосопряженными. Тогда при выполнении условий

$$B \geq 0, \quad A > 0, \quad R > \frac{1+\varepsilon}{4}A \quad (12.31)$$

с $\varepsilon > 0$ имеют место априорные оценки

$$\|y_{n+1}\|_A^2 \leq 2 \frac{1+\varepsilon}{\varepsilon} \left(\|y_0\|_A^2 + \|y_1 - y_0\|_R^2 \right), \quad (12.32)$$

$$\|y_{n+1}\|_A^2 + \|y_n - y_{n-1}\|_R^2 \leq \frac{4+3\varepsilon}{\varepsilon} \left(\|y_0\|_A^2 + \|y_1 - y_0\|_R^2 \right). \quad (12.33)$$

Для разностной схемы (12.26) при тех же предположениях об операторах R и A при выполнении операторных неравенств

$$B \geq \varepsilon E, \quad A > 0, \quad R > \frac{1}{4}A \quad (12.34)$$

с постоянной $\varepsilon > 0$ для разностного решения справедливы априорные оценки

$$\mathcal{E}_{n+1} \leq \mathcal{E}_1 + \frac{1}{2\varepsilon} \sum_{k=1}^n \tau \|\varphi_k\|^2, \quad (12.35)$$

$$\mathcal{E}_{n+1} \leq \mathcal{E}_1 + \frac{1}{2} \sum_{k=1}^n \tau \|\varphi_k\|_{B^{-1}}^2. \quad (12.36)$$

Здесь

$$\begin{aligned} \mathcal{E}_{n+1} = & \frac{1}{4} (A(y_{n+1} + y_n), y_{n+1} + y_n) + (R(y_{n+1} - y_n), y_{n+1} - y_n) - \\ & - \frac{1}{4} (A(y_{n+1} - y_n), y_{n+1} - y_n). \end{aligned}$$

При сформулированных ограничениях величина \mathcal{E}_n задает норму.

12.2.3. Разностные схемы для параболического уравнения

Рассмотрим разностные схемы для одномерного параболического уравнения (12.1). По пространству будем использовать равномерную сетку

$$\bar{\omega}_h = \{x \mid x = x_i = ih, \quad i = 0, 1, \dots, N, \quad Nh = l\},$$

и пусть ω_h — множество внутренних узлов ($i = 1, 2, \dots, N - 1$), а $\partial\omega_h$ — множество граничных узлов.

При приближенном решении задачи (12.1)–(12.3) определим сеточный оператор

$$Ay = -(ay_{\bar{x}})_x, \quad x \in \omega_h, \quad (12.37)$$

для сеточных функций $y = 0$, $x \notin \omega_h$. Для задания коэффициента можно использовать выражения

$$a_i = k(x_{i-1/2}), \quad x_{i-1/2} = x_i - \frac{h}{2},$$

$$a_i = 0,5(k(x_{i-1}) + k(x_i)),$$

$$a_i = \left[\frac{1}{h} \int_{x_{i-1}}^{x_i} \frac{dx}{k(x)} \right]^{-1}.$$

В $H = L_2(\omega_h)$ скалярное произведение и норму введем соотношениями

$$(y, v) = \sum_{i=1}^{N-1} y_i v_i h, \quad \|y\| = \sqrt{(y, y)}.$$

Оператор A является самосопряженным и положительным:

$$A^* = A > 0. \quad (12.38)$$

Приведем также оценки оператора A снизу и сверху:

$$\delta E \leq A \leq \Delta E, \quad (12.39)$$

где

$$\delta = \frac{8}{l^2} \min_{0 \leq x \leq l} k(x), \quad \Delta = \frac{4}{h^2} \max_{0 \leq x \leq l} k(x).$$

Исходной дифференциальной задаче (12.1)–(12.3) поставим в соответствие задачу Коши для дифференциально-разностного уравнения:

$$\frac{dv}{dt} + Av = \varphi(t),$$

$$v(0) = u_0.$$

Для ее решения используем схему с весами

$$\frac{y^{n+1} - y^n}{\tau} + A(\sigma y^{n+1} + (1 - \sigma)y^n) = \varphi^n, \quad n = 0, 1, \dots, \quad (12.40)$$

$$y^0 = u_0. \quad (12.41)$$

Схема с весами будет устойчивой в H_A при

$$\sigma \geq \sigma_0, \quad \sigma_0 = \frac{1}{2} - \frac{1}{\tau \|A\|}. \quad (12.42)$$

В частности, схема с $\sigma \geq 0,5$ абсолютно (при всех $\tau > 0$) устойчива.

Рассмотрим вопрос о точности разностной схемы с весами (12.40), (12.41). Сформулируем соответствующую задачу для погрешности приближенного решения

$$z^n(x) = y^n(x) - u(x, t_n), \quad x \in \omega_h, \quad t_n \in \omega_\tau$$

с учетом

$$z^n(x) = 0, \quad x \in \partial\omega_h, \quad t_n \in \omega_\tau.$$

Начальное условие задается точно и поэтому положим

$$z_0(x) = 0, \quad x \in \omega_h.$$

Для погрешности из (12.41) следует

$$\frac{z^{n+1} - z^n}{\tau} + A(\sigma z^{n+1} + (1 - \sigma)z^n) = \psi^n, \quad n = 0, 1, \dots$$

Предполагая достаточную гладкость точного решения и коэффициентов уравнения (12.1), для погрешности аппроксимации будем иметь

$$\psi^n(x) = O(|h|^2 + \tau^\nu), \quad x \in \omega, \quad t_n \in \omega_\tau,$$

где $\nu = \nu(\sigma) = 2$ при $\sigma = 0,5$ и $\nu = 1$, если $\sigma \neq 0,5$.

Для погрешности верна априорная оценка

$$\|z^{n+1}\|_A \leq \|\psi^n\|_{A^{-1}} + \sum_{k=1}^n \tau \|\psi_i^k\|_{A^{-1}},$$

где использованы обозначения

$$y_i = \frac{y^n - y^{n-1}}{\tau}.$$

Следовательно разностная схема с весами сходится в H_A со скоростью $O(h^2 + \tau^\nu)$.

На основе использования оценок устойчивости по правой части устанавливается сходимость и в других нормах.

12.2.4. Гиперболические уравнения

Рассмотрим теперь разностные схемы для решения краевой задачи для одномерного гиперболического уравнения второго порядка (12.2), (12.5), (12.6). После дискретизации по пространству приходим к дифференциально-разностной задаче

$$\begin{aligned} \frac{d^2 v}{dt^2} + Av &= \varphi(t), \\ v(0) &= u_0, \quad \frac{dv}{dt}(0) = u_1 \end{aligned}$$

с ранее рассмотренным разностным оператором A .

Будем использовать разностное уравнение

$$\frac{y^{n+1} - 2y^n + y^{n-1}}{\tau^2} + A(\sigma y^{n+1} + (1 - 2\sigma)y^n + \sigma y^{n-1}) = \varphi^n, \quad (12.43)$$

$$n = 1, 2, \dots,$$

которое аппроксимирует (12.5) со вторым порядком по времени и по пространству. Схема (12.43) записывается в каноническом виде (12.28) при

$$B = 0, \quad R = \frac{1}{\tau^2} E + \sigma A.$$

Условия устойчивости (12.29) дают следующие ограничения на вес:

$$\sigma \geq \sigma_0, \quad \sigma_0 = \frac{1}{4} - \frac{1}{\tau^2 \|A\|}.$$

С привлечением априорных оценок устойчивости по правой части исследуется задача для погрешности и устанавливается сходимость разностной схемы (12.43).

12.2.5. Многомерные задачи

Будем рассматривать краевую задачу для параболического уравнения второго порядка (12.7)–(12.9) в прямоугольнике Ω . Введем равномерную прямоугольную сетку с шагами h_1 и h_2 , так что

$$\omega_h = \{x \mid x = (x_1, x_2), x_\alpha = i_\alpha h_\alpha, \\ i_\alpha = 1, 2, \dots, N_\alpha, N_\alpha h_\alpha = l_\alpha, \alpha = 1, 2\}.$$

Определим разностный оператор

$$A = \sum_{\alpha=1}^2 A^{(\alpha)}, \quad (12.44)$$

где $A^{(\alpha)}$, $\alpha = 1, 2$ — одномерные разностные операторы

$$A^{(\alpha)} y = -(a^{(\alpha)} y_{\bar{x}_\alpha})_{x_\alpha}, \quad \alpha = 1, 2, \quad x \in \omega_h, \quad (12.45)$$

определенные для сеточных функций $y(x) = 0$, $x \notin \omega_h$. Для коэффициентов положим, например,

$$a^{(1)}(x) = k(x_1 - 0,5h_1, x_2), \quad a^{(2)}(x) = k(x_1, x_2 - 0,5h_2).$$

Вычислительная реализация неявных схем ($\sigma \neq 0$) (12.40), (12.41) для численного решения задачи (12.7)–(12.9) связана с решением сеточной эллиптической задачи. В экономичных разностных схемах переход на новый временной слой осуществляется с вычислительными затратами на один узел, не зависящими от общего числа узлов дискретизации по пространству. Экономичные схемы строятся на основе аддитивного представления (12.44) с переходом к последовательности более простых

задач с операторами $A^{(\alpha)}$, $\alpha = 1, 2$. Приведем примеры некоторых схем расщепления.

Для правой части уравнения используется аддитивное представление

$$\varphi^n = \varphi_1^n + \varphi_2^n.$$

Классическая разностная схема переменных направлений (схема Писмена—Рэкфорда) при расщеплении (12.44), (12.45) состоит из двух шагов. Сначала по известному y^n находится вспомогательная сеточная функция, которую мы обозначим $y^{n+1/2}$, из уравнения

$$\frac{y^{n+1/2} - y^n}{0,5\tau} + A^{(1)}y^{n+1/2} + A^{(2)}y^n = 2\varphi_1^n. \quad (12.46)$$

Интерпретируя $y^{n+1/2}$ как решение на момент времени $t = t_{n+1/2}$, можем заметить, что (12.46) при $2\varphi_1^n = \varphi_n$ соответствует определению решения по чисто неявной схеме по переменной x_1 (оператор $A^{(1)}$) и по явной схеме по переменной x_2 (оператор $A^{(2)}$).

$$\frac{y^{n+1} - y^{n+1/2}}{0,5\tau} + A^{(1)}y^{n+1/2} + A^{(2)}y^{n+1} = 2\varphi_2^n. \quad (12.47)$$

Тем самым второй шаг связывается с использованием явной схемы по первой переменной и чисто неявной — по второй переменной.

Сформулируем условия устойчивости схемы переменных направлений. Пусть в схеме (12.46), (12.47) постоянные операторы $A^{(\alpha)} \geq 0$, $\alpha = 1, 2$. Тогда для разностного решения имеет место следующая оценка устойчивости по начальным данным и правой части:

$$\left\| (E + \frac{\tau}{2}A^{(2)})y^{n+1} \right\| \leq \left\| (E + \frac{\tau}{2}A^{(2)})y^0 \right\| + \sum_{k=0}^n \tau \left(\|\varphi_1^k\| + \|\varphi_2^k\| \right).$$

На основе этой оценки устанавливается, что схема переменных направлений сходится со скоростью $O(\tau^2 + |h_i|^2)$ в соответствующей, зависящей от операторов расщепления норме.

Необходимо выделить аддитивные схемы, которые относятся к классу безусловно устойчивых разностных схем при расщеплении на произвольное число операторов — схемы многокомпонентного расщепления. Аддитивные разностные схемы для задач с расщеплением на три и более попарно некоммутирующих операторов традиционно строятся на основе

понятия суммарной аппроксимации – схемы покомпонентного расщепления (локально-одномерные схемы).

Для двумерной задачи (12.7)–(12.9) имеем

$$\frac{y^{n+\alpha/2} - y^{n+(\alpha-1)/2}}{\tau} + A^{(\alpha)} (\sigma_{\alpha} y^{n+\alpha/2} + (1 - \sigma_{\alpha}) y^{n+(\alpha-1)/2}) = \varphi_{\alpha}^n, \quad (12.48)$$

$$\alpha = 1, 2, \quad n = 0, 1, \dots$$

При $\sigma_{\alpha} \geq 0,5$ схема покомпонентного расщепления (12.48) безусловно устойчива. Приведем соответствующую априорную оценку устойчивости по начальным данным и правой части.

Для правых частей φ_{α}^n , $\alpha = 1, 2$ будем использовать специальное представление

$$\varphi_{(\alpha)}^n = \overset{\circ}{\varphi}_{\alpha}^n + \overset{*}{\varphi}_{\alpha}^n, \quad \alpha = 1, 2, \quad \sum_{\alpha=1}^2 \overset{\circ}{\varphi}_{\alpha}^n = 0. \quad (12.49)$$

При $0,5 \leq \sigma_{\alpha} \leq 2$, $\alpha = 1, 2$ и $\tau > 0$ для решения задачи (12.48), (12.49) выполняется априорная оценка

$$\|y^{n+1}\| \leq \|y^0\| + \sum_{k=0}^n \tau \sum_{\alpha=1}^2 \left(\|\overset{*}{\varphi}_{\alpha}^k\| + \tau \|A^{(\alpha)} \sum_{\beta=\alpha}^2 \overset{\circ}{\varphi}_{\beta}^k\| \right).$$

При исследовании сходимости локально-одномерных схем существенно учитывается специальное представление для погрешности типа (12.49). Отметим также, что устойчивость локально-одномерных схем устанавливается не только в гильбертовых пространствах сеточных функций, но и при использовании принципа максимума — в равномерной норме.

12.3. Упражнения

В предложенных ниже примерах получены результаты по устойчивости двух- и трехслойных разностных схем, исследуются разностные схемы для параболических и гиперболических уравнений второго порядка.

Упражнение 12.1. Пусть в схеме (12.10), (12.11) $A = A^* > 0$. Покажите, что при

$$B \geq \frac{\tau}{2} A \quad (12.50)$$

разностная схема устойчива по начальным данным и правой части, и для разностного решения справедлива априорная оценка

$$\|y^{n+1}\|_A \leq \|u_0\|_A + \|\varphi^0\|_{A^{-1}} + \|\varphi^n\|_{A^{-1}} + \sum_{k=1}^n \tau \|\varphi_k^k\|_{A^{-1}}. \quad (12.51)$$

Решение. Представим решение задачи (12.10), (12.11) в виде

$$y^n = v^n + w^n, \quad (12.52)$$

где w^n есть решение стационарного уравнения

$$Aw^{n+1} = \varphi^n, \quad n = 0, 1, \dots,$$

и пусть $w(0) = w(\tau)$. Для v^n получим задачу

$$\begin{aligned} Bv_t + Av &= \tilde{\varphi}^n, \\ v^0 &= y^0 - w^0 \end{aligned}$$

с правой частью

$$\tilde{\varphi}^n = -(B - \tau A)w_t^n, \quad \tilde{\varphi}^0 = 0$$

при использовании стандартных обозначений

$$y_t = \frac{y^{n+1} - y^n}{\tau}.$$

Для решения этой задачи (см. оценку (12.23)) получим

$$\|v^{n+1}\|_A \leq \|v^0\|_A + \sum_{k=0}^n \tau \|B^{-1}\tilde{\varphi}^k\|_A.$$

Принимая во внимание, что $w^t = A^{-1}\varphi_t$, для последнего слагаемого получим

$$\|B^{-1}\tilde{\varphi}^k\|_A = \|A^{1/2}B^{-1}\tilde{\varphi}^k\| = \|(E - \tau C)A^{-1/2}\varphi_t^k\|,$$

где $C = A^{1/2}B^{-1}A^{1/2}$. В силу предположения (12.50), обеспечивающего устойчивость схемы, имеем

$$\|E - \tau C\| \leq 1$$

и поэтому

$$\|A^{1/2}B^{-1}\bar{\varphi}^k\| \leq \|A^{-1/2}\varphi_i^k\| = \|\varphi_i^k\|_{A^{-1}}.$$

Тем самым

$$\|v^{n+1}\|_A \leq \|v^0\|_A + \sum_{k=1}^n \tau \|\varphi_i^k\|_{A^{-1}}.$$

Принимая во внимание, что

$$\|v^0\|_A \leq \|y^0\|_A + \|w^0\|_A = \|y^0\|_A + \|\varphi^0\|_{A^{-1}},$$

получаем

$$\|v^{n+1}\|_A \leq \|y^0\|_A + \|\varphi^0\|_{A^{-1}} + \sum_{k=1}^n \tau \|\varphi_i^k\|_{A^{-1}}.$$

С учетом (12.52) приходим к доказываемой оценке (12.51).

Упражнение 12.2. Методом энергетических неравенств докажите априорную оценку устойчивости по начальным данным (12.30) трехслойной схемы (12.27), (12.28) при выполнении операторных неравенств (12.29).

Решение. Положим

$$v^n = \frac{1}{2}(y^n + y^{n-1}), \quad w^n = y^n - y^{n-1}$$

и с учетом тождества

$$y^n = \frac{1}{4}(y^{n+1} + 2y^n + y^{n-1}) - \frac{1}{4}(y^{n+1} - 2y^n + y^{n-1})$$

перепишем схему (12.28) в виде

$$B \frac{w^{n+1} + w^n}{2\tau} + R(w^{n+1} - w^n) - A(w^{n+1} - w^n) + A \frac{v^{n+1} + v^n}{2} = 0.$$

Домножим скалярно это уравнение на

$$2(v^{n+1} - v^n) = w^{n+1} + w^n,$$

что дает равенство

$$\frac{1}{2\tau} (B(w^{n+1} + w^n), w^{n+1} + w^n) + (R(w^{n+1} - w^n), w^{n+1} + w^n) - \\ - \frac{1}{4} (A(w^{n+1} - w^n), w^{n+1} + w^n) + (A(v^{n+1} + v^n), v^{n+1} - v^n) = 0.$$

Для самосопряженных операторов R и A и неотрицательного оператора B ($B \geq 0$) отсюда следует неравенство

$$\mathcal{E}_{n+1} \leq \mathcal{E}_n,$$

где с учетом введенных обозначений

$$\mathcal{E}_{n+1} = \frac{1}{4} (A(y^{n+1} + y^n), y^{n+1} + y^n) + \\ + (R(y^{n+1} - y^n), y^{n+1} - y^n) - \frac{1}{4} (A(y^{n+1} - y^n), y^{n+1} - y^n).$$

Это и есть доказываемая априорная оценка (12.30).

Упражнение 12.3. Постройте двухслойную разностную схему повышенного порядка аппроксимации для решения уравнения

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(x, t), \quad 0 < x < l, \quad 0 < t \leq T \quad (12.53)$$

с условиями (12.2), (12.3).

Решение. Будем рассматривать схему с весами и оптимизацию порядка аппроксимации проведем за счет выбора веса. Уравнению (12.53) поставим в соответствие разностную схему

$$\frac{y^{n+1} - y^n}{\tau} = \sigma y_{\bar{x}\bar{x}}^{n+1} + (1 - \sigma) y_{\bar{x}\bar{x}}^n + \varphi^n, \quad n = 0, 1, \dots$$

Погрешность аппроксимации на решениях уравнения (12.53) есть

$$\psi^n = \sigma u_{\bar{x}\bar{x}}^{n+1} + (1 - \sigma) u_{\bar{x}\bar{x}}^n - \frac{u^{n+1} - u^n}{\tau} + \varphi^n.$$

Перепишем ее в более удобном виде

$$\psi^n = \frac{1}{2} (u_{\bar{x}\bar{x}}^{n+1} + u_{\bar{x}\bar{x}}^n) + \left(\sigma - \frac{1}{2} \right) \tau \frac{u_{\bar{x}\bar{x}}^{n+1} - u_{\bar{x}\bar{x}}^n}{\tau} - \frac{u^{n+1} - u^n}{\tau} + \varphi^n.$$

С учетом

$$\begin{aligned}
 u_{xx} &= \frac{\partial^2 u}{\partial x^2} + \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4} + O(h^4), \\
 u^{n+1} &= u(x, t_n + 0,5\tau) + \frac{\tau}{2} \frac{\partial u}{\partial t}(x, t_n + 0,5\tau) + \\
 &\quad + \frac{\tau^2}{8} \frac{\partial^2 u}{\partial t^2}(x, t_n + 0,5\tau) + O(\tau^3), \\
 u^n &= u(x, t_n + 0,5\tau) - \frac{\tau}{2} \frac{\partial u}{\partial t}(x, t_n + 0,5\tau) + \\
 &\quad + \frac{\tau^2}{8} \frac{\partial^2 u}{\partial t^2}(x, t_n + 0,5\tau) + O(\tau^3), \\
 \frac{1}{2}(u^{n+1} + u^n) &= u(x, t_n + 0,5\tau) + \frac{\tau^2}{8} \frac{\partial^2 u}{\partial t^2}(x, t_n + 0,5\tau) + O(\tau^3), \\
 \frac{u^{n+1} - u^n}{\tau} &= \frac{\partial^2 u}{\partial t^2}(x, t_n + 0,5\tau) + O(\tau^2)
 \end{aligned}$$

на решениях уравнения (12.53) получим

$$\begin{aligned}
 \psi^n &= \left(\sigma - \frac{1}{2}\right) \tau \frac{\partial^3 u}{\partial x^2 \partial t}(x, t_n + 0,5\tau) + \\
 &\quad + \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4}(x, t_n + 0,5\tau) + O(h^4 + \tau^2).
 \end{aligned} \tag{12.54}$$

Принимая во внимание, что

$$\frac{\partial^3 u}{\partial x^2 \partial t} = + \frac{h^2}{12} \frac{\partial^4 u}{\partial x^4} + \frac{\partial^2 f}{\partial x^2},$$

из (12.54) получим

$$\begin{aligned}
 \psi^n &= \varphi^n - f(x, t_n + 0,5\tau) - \frac{h^2}{12} \frac{\partial^2 f}{\partial x^2}(x, t_n + 0,5\tau) + \\
 &\quad + \left(\left(\sigma - \frac{1}{2}\right) \tau + \frac{h^2}{12} \right) \frac{\partial^3 u}{\partial x^2 \partial t}(x, t_n + 0,5\tau) + O(h^4 + \tau^2).
 \end{aligned}$$

Выбирая

$$\sigma = \sigma^* = \frac{1}{2} - \frac{h^2}{12\tau},$$

$$\varphi^n = f(x, t_n + 0,5\tau) + \frac{h^2}{12} \frac{\partial^2 f}{\partial x^2}(x, t_n + 0,5\tau),$$

получим $\psi^n = O(h^4 + \tau^2)$.

Упражнение 12.4. *Покажите, что если в разностной схеме с весами (12.40), (12.41) оператор $A = A^* > 0$ и*

$$\sigma \geq \frac{1 + \varepsilon}{2} - \frac{1}{\tau \|A\|}, \quad \varepsilon = \text{const}, \quad (12.55)$$

то верна априорная оценка устойчивости по начальным данным и правой части

$$\|y^{n+1}\|_{\tilde{A}}^2 \leq \|u_0\|_{\tilde{A}}^2 + \frac{1 + \varepsilon}{2\varepsilon} \sum_{k=0}^n \tau \|\varphi^k\|^2, \quad (12.56)$$

где

$$\tilde{A} = (E + \sigma\tau A)A, \quad \|y\|_{\tilde{A}}^2 = \|y\|_A^2 + \sigma\tau \|Ay\|^2.$$

Решение. Умножая уравнение (12.40) на оператор $B = E + \sigma\tau A$, получим разностную схему

$$\tilde{B} + \tilde{A}y = \tilde{\varphi},$$

$$\tilde{B} = B^2, \quad \tilde{\varphi} = B\varphi.$$

При ограничениях (12.55) справедливо неравенство

$$\tilde{B} \geq \frac{1 + \varepsilon}{2} \tau \tilde{A}. \quad (12.57)$$

Это следует из

$$\begin{aligned} \tilde{B} - \frac{1 + \varepsilon}{2} \tau \tilde{A} &= (E + \sigma\tau A) \left(E + \left(\sigma - \frac{1 + \varepsilon}{2} \right) \tau A \right) \geq \\ &\geq (E + \sigma\tau A) \left(\frac{1}{\|A\|} + \left(\sigma - \frac{1 + \varepsilon}{2} \right) \tau \right) A \geq 0. \end{aligned}$$

При выполнении неравенства (12.57) имеет место (см. (12.24), (12.25)) априорная оценка в $H_{\tilde{A}}$:

$$\|y^{n+1}\|_{\bar{A}}^2 \leq \|u_0\|_{\bar{A}}^2 + \frac{1+\varepsilon}{2\varepsilon} \sum_{k=0}^n \tau \|\tilde{\varphi}^k\|_{\bar{B}^{-1}}^2.$$

Так как $\|\tilde{\varphi}_k\|_{\bar{B}^{-1}}^2 = \|\varphi_k\|^2$, то мы приходим к искомой оценке (12.56).

Упражнение 12.5. *Сформулируйте условия устойчивости явной трехслойной схемы второго порядка аппроксимации по времени и пространству для задачи (12.3), (12.5), (12.6).*

Решение. В рассматриваемом случае разностное уравнение имеет вид

$$\frac{y^{n+1} - 2y^n + y^{n-1}}{\tau^2} + Ay^n = \varphi^n, \quad n = 1, 2, \dots,$$

где разностный оператор A определен согласно (12.37). Схема записывается в каноническом виде (12.26) при

$$B = 0, \quad R = \frac{1}{\tau^2} E,$$

а условия устойчивости (12.29) приводят к следующим ограничениям на шаг по времени:

$$\tau \leq \frac{4}{\|A\|}.$$

Принимая во внимание (12.39), откуда получаем

$$\tau \leq \frac{4}{\Delta} = \left(\max_{0 \leq x \leq l} k(x) \right)^{-1/2} h.$$

Тем самым устойчивость явной схемы обеспечивается при $\tau = O(h)$.

Упражнение 12.6. *Покажите, что схема переменных направлений (12.46), (12.47) при $\varphi_n^{(\alpha)} = 0,5\varphi_n$, $\alpha = 1, 2$ сходится со скоростью $O(h^2 + \tau^2)$.*

Решение. Запишем соответствующую задачу для погрешности. Положим, как обычно, $z^n = y^n - u(x, t_n)$ и пусть $z^{n+1/2} = y^{n+1/2} - \bar{u}(x, t_n)$. Выбор точного решения, с которым связывается $y^{n+1/2}$, проведем позднее. Задача для погрешности имеет вид

$$\begin{aligned} \frac{z^{n+1/2} - z^n}{0,5\tau} + A^{(1)}z^{n+1/2} + A^{(2)}z^n &= \psi_1^n, \\ \frac{z^{n+1} - z^{n+1/2}}{0,5\tau} + A^{(1)}z^{n+1/2} + A^{(2)}z^{n+1} &= \psi_2^n. \end{aligned}$$

Для погрешности аппроксимации получим

$$\begin{aligned}\psi_1^n &= -\frac{\bar{u}^n - u^n}{0,5\tau} - A^{(1)}\bar{u}^n - A^{(2)}u^n + \varphi^n, \\ \psi_2^n &= -\frac{u^{n+1} - \bar{u}^n}{0,5\tau} - A^{(1)}\bar{u}^n - A^{(2)}u^{n+1} + \varphi^n.\end{aligned}$$

Положим

$$\bar{u}^n = \frac{1}{2}(u^{n+1} + u^n) + \frac{\tau^2}{4}A^{(2)}\frac{u^{n+1} - u^n}{\tau}.$$

В этом случае следует

$$\psi_2^n - \psi_1^n = -\frac{u^{n+1} - 2\bar{u}^n + u^n}{0,5\tau} - A^{(2)}(u^{n+1} - u^n) = 0.$$

Кроме того имеем

$$\begin{aligned}\psi_1^n &= -A^{(1)}\frac{u^{n+1} + u^n}{2} - A^{(2)}u^n + \varphi^n - \frac{u^{n+1} - u^n}{\tau} - \\ &\quad - \frac{\tau}{2}A^{(2)}\frac{u^{n+1} - u^n}{\tau} + O(\tau^2).\end{aligned}$$

На решениях уравнения (12.7)

$$\psi_1^n = \psi_2^n = O(\tau^2 + |h|^2).$$

Тем самым, при специальном определении промежуточного решения разностная схема переменных направлений (12.46), (12.47) с $\varphi_{(\alpha)}^n = 0,5\varphi_n$, $\alpha = 1, 2$ имеет второй порядок аппроксимации по времени и по пространству.

Для исследования точности рассмотрим сеточную задачу для погрешности. Используя априорную оценку для схемы переменных направлений при точном задании начальных условий получим

$$\left\| \left(E + \frac{\tau}{2}A^{(2)} \right) z^{n+1} \right\| \leq \sum_{k=0}^n \tau \left(\|\psi_1^k\| + \|\psi_2^k\| \right).$$

Тем самым схема переменных направлений сходится со скоростью $O(\tau^2 + |h|^2)$ в соответствующей норме.

12.4. Задачи

Задача 12.1. Докажите, что для оператора $C = C^* > 0$ при $\tau > 0$ неравенства

$$\|S\| = \|E - \tau C\| \leq \rho,$$

$$\frac{1 - \rho}{\tau} E \leq C \leq \frac{1 + \rho}{\tau} E$$

эквивалентны.

Задача 12.2. Покажите, что условие (12.20) необходимо и достаточно для устойчивости разностной схемы (12.13), (12.14) в H_B , если

$$A = A^* > 0, \quad B = B^* > 0.$$

Задача 12.3. Методом энергетических неравенств докажите оценку устойчивости по начальным данным и правой части (12.25) для разностной схемы (12.10), (12.11), когда $A = A^* > 0$, а для оператора B справедливо неравенство (12.24).

Задача 12.4. Пусть операторы A и B удовлетворяют условиям

$$B \geq \varepsilon E + 0,5\tau A, \quad A = A^* > 0,$$

где ε — любое положительное число. Тогда для разностной схемы (12.10), (12.11) верна априорная оценка

$$\|y^{n+1}\|_A^2 \leq \|u_0\|_A^2 + \frac{1}{2\varepsilon} \sum_{k=0}^n \tau \|\varphi^k\|^2.$$

Задача 12.5. Запишите трехслойную схему (12.26) с самосопряженными операторами A, B и R в виде двухслойной векторной схемы

$$B \frac{Y^{n+1} - Y^n}{\tau} + AY^n = \Phi^n, \quad n = 1, 2, \dots$$

с самосопряженным оператором A при определении вектора

$$Y^n = \left\{ \frac{1}{2}(y^n + y^{n-1}), y^n - y^{n-1} \right\}.$$

Задача 12.6. Покажите, что условия

$$B + \frac{\tau}{2} \frac{\varrho - 1}{\varrho + 1} A \geq 0, \quad R > \frac{1}{4} A > 0, \quad \varrho > 1$$

достаточны для ϱ -устойчивости трехслойной схемы (12.28) с самосопряженными операторами A, B и R .

Задача 12.7. Пусть $A = A^* > 0, R = R^* > 0$ в разностной схеме (12.26), (12.27). Докажите, что при

$$B \geq \varepsilon E, \quad R > \frac{1}{4} A, \quad \varepsilon = \text{const}$$

для разностного решения верна априорная оценка (12.35).

Задача 12.8. Аппроксимируйте краевые условия третьего рода (12.4) при численном решении задачи Коши для параболического уравнения (12.1).

Задача 12.9. Постройте монотонную разностную схему второго порядка точности для решения задачи (12.1)–(12.3).

Задача 12.10. Интегро-интерполяционным методом постройте разностную схему в случае, когда коэффициент $k(x)$ имеет разрыв первого рода в точке $x = x^* \in \omega_h$ и на разрыве имеют место условия сопряжения

$$\begin{aligned} u(x^* + 0, t) - u(x^* - 0, t) &= 0, \\ k \frac{\partial u}{\partial x}(x^* + 0, t) - k \frac{\partial u}{\partial x}(x^* - 0, t) &= 0. \end{aligned}$$

Задача 12.11. На основе принципа максимума сформулируйте условия устойчивости схемы с весами (12.40), (12.41) при численном решении задачи (12.1)–(12.3).

Задача 12.12. Докажите, что условие

$$A^* + \left(\sigma - \frac{1}{2} \right) \tau A^* A \geq 0$$

необходимо и достаточно для устойчивости схемы с весами (12.40), (12.41) с несамосопряженным оператором $A > 0$ по начальным данным ($\varphi^n = 0$, $n = 0, 1, \dots$) в пространстве H_{A^*A} , т. е. для выполнения оценки

$$\|Ay^{n+1}\| \leq \|Ay^n\|, \quad n = 0, 1, \dots$$

Задача 12.13. Исследуйте погрешность аппроксимации и условия устойчивости трехслойной разностной схемы с весами

$$\frac{y^{n+1} - y^{n-1}}{2\tau} + A(\sigma_1 y^{n+1} + (1 - \sigma_1 - \sigma_2)y^n + \sigma y^{n-1}) = \varphi^n, \\ n = 1, 2, \dots,$$

для задачи (12.1)–(12.3).

Задача 12.14. Рассмотрите двухпараметрический класс схем с весами

$$\frac{y^{n+1} - 2y^n + y^{n-1}}{\tau^2} + A(\sigma_1 y^{n+1} + (1 - \sigma_1 - \sigma_2)y^n + \sigma y^{n-1}) = \varphi^n, \\ n = 1, 2, \dots,$$

для уравнения (12.5). Сформулируйте условия устойчивости и получите оценки устойчивости по начальным данным и правой части.

Задача 12.15. Исследуйте точность разностной схемы Дугласа—Рэкфорда

$$\frac{y^{n+1/2} - y^n}{\tau} + A^{(1)}y^{n+1/2} + A^{(2)}y^n = \varphi_n, \\ \frac{y^{n+1} - y^{n+1/2}}{\tau} + A^{(2)}(y^{n+1} - y^n) = 0$$

при расщеплении (12.44), (12.45) для решения задачи (12.7)–(12.9).

Задача 12.16. Рассмотрите проблему аппроксимации граничных условий при использовании схемы переменных направлений при численном решении задачи Коши для уравнения (12.7) с неоднородными граничными условиями первого рода.

Задача 12.17. Рассмотрим двухслойную разностную схему, которая имеет канонический вид

$$B \frac{y^{n+1} - y^n}{\tau} + Ay^n = \varphi^n, \quad t_n \in \omega_\tau,$$

где оператор A имеет аддитивное представление (12.44) с постоянными операторами $A^{(\alpha)} \geq 0$, $\alpha = 1, 2$. Факторизованная схема соответствует выбору оператора B в виде

$$B = B_1 B_2,$$

где

$$B_\alpha = E + \sigma \tau A^{(\alpha)}, \quad \alpha = 1, 2.$$

Докажите, что при $\sigma \geq 0,5$ схема безусловно устойчива и для решения имеет место априорная оценка

$$\|B_2 y^{n+1}\| \leq \|B_2 y^0\| + \sum_{k=0}^n \tau \|\varphi^k\|.$$

Задача 12.18. Исследуйте точность в $L_2(\omega_h)$ локально-одномерной схемы (12.44), (12.45), (12.48) для задачи (12.7)–(12.9) на основе использования представления для погрешности (12.49).

Задача 12.19. Получите априорные оценки устойчивости для локально-одномерной схемы с представлением правой части в виде (12.49) в равномерной норме (в $L_\infty(\omega_h)$). Исследуйте сходимость в $L_\infty(\omega_h)$ локально-одномерной схемы (12.44), (12.45), (12.48) для задачи (12.7)–(12.9).

Задача 12.20. Для задачи (12.7)–(12.9) рассмотрите схему аддитивно-усредненную локально-одномерную схему

$$\frac{y_\alpha^{n+1} - y^n}{2\tau} + A^{(\alpha)}(\sigma_\alpha y_\alpha^{n+1} + (1 - \sigma_\alpha)y^n) = \varphi_\alpha^n,$$

$$\alpha = 1, 2, \quad n = 0, 1, \dots,$$

$$y^{n+1} = \frac{1}{2} \sum_{\alpha=1}^2 y_\alpha^{n+1}.$$

Сформулируйте условия устойчивости и получите априорную оценку с расщепленными правыми частями (12.49).

Литература

- [1] Бахвалов Н. С., Жидков Н. П., Кобельков Г. М. *Численные методы*. М.: Наука, 1987.
- [2] Бахвалов Н. С., Лапин А. В., Чижонков Е. В. *Численные методы в задачах и упражнениях*. М.: Высшая школа, 2000.
- [3] Березин И. С., Жидков Н. П. *Методы вычислений*. М.: Наука, 1966, т. 1; Физматгиз, 1962, т. 2.
- [4] Васильев Ф. П. *Численные методы решения экстремальных задач*. М.: Наука, 1980.
- [5] Волков Е. А. *Численные методы*. М.: Наука, 1987.
- [6] Гавурин М. К. *Лекции по методам вычислений*. М.: Наука, 1971.
- [7] Годунов С. К., Рябенский В. С. *Разностные схемы*. М.: Наука, 1977.
- [8] Дробышевич В. И., Дымников В. П., Ривин Г. С. *Задачи по вычислительной математике*. М.: Наука, 1980.
- [9] Завьялов Ю. С., Квасов Б. И., Мирошниченко В. Л. *Методы сплайн-функций*. М.: Наука, 1980.
- [10] Калиткин Н. Н. *Численные методы*. М.: Наука, 1978.
- [11] Коллатц Л., Альбрехт Ю. *Задачи по прикладной математике*. М.: Мир, 1978.
- [12] Коновалов А. Н. *Введение в вычислительные методы линейной алгебры*. Новосибирск: Наука, 1993.
- [13] Крылов В. И., Бобков В. В., Монастырный П. И. *Вычислительные методы*. М.: Наука, 1976, т. 1; 1977, т. 2.
- [14] Ляшко И. И., Макаров В. Л., Скоробогатько А. А. *Методы вычислений*. Киев: Высшая школа, 1977.
- [15] Марчук Г. И. *Методы вычислительной математики*. М.: Наука, 1989.
- [16] Марчук Г. И. *Методы расщепления*. М.: Наука, 1988.
- [17] Марчук Г. И., Агошков В. И. *Введение в проекционно-сеточные методы*. М.: Наука, 1981.
- [18] Ортега Дж., Пул У. *Введение в численные методы решения дифференциальных уравнений*. М.: Наука, 1986.

- [19] Ортега Дж., Рейнболдт В. *Итерационные методы решения нелинейных систем уравнений со многими неизвестными*. М.: Мир, 1975.
- [20] *Сборник от задачи по числени методи*. София: Наука и искусство, 1986.
- [21] Самарский А. А. *Введение в численные методы*. М.: Наука, 1997.
- [22] Самарский А. А. *Теория разностных схем*. М.: Наука, 1989.
- [23] Самарский А. А., Андреев В. Б. *Разностные методы для эллиптических уравнений*. М.: Наука, 1976.
- [24] Самарский А. А., Вабищевич П. Н. *Аддитивные схемы для задач математической физики*. М.: Наука, 1999.
- [25] Самарский А. А., Вабищевич П. Н. *Численные методы решения задач конвекции-диффузии*. М.: Эдиториал УРСС, 1999.
- [26] Самарский А. А., Гулин А. В. *Устойчивость разностных схем*. М.: Наука, 1973.
- [27] Самарский А. А., Гулин А. В. *Численные методы*. М.: Наука, 1989.
- [28] Самарский А. А., Николаев Е. С. *Методы решения сеточных уравнений*. М.: Наука, 1978.
- [29] *Сборник задач по методам вычислений*. М.: Физматлит, 1994.
- [30] *Современные численные методы решения обыкновенных дифференциальных уравнений*. М.: Мир, 1979.
- [31] Стренг Г., Фикс Дж. *Теория метода конечных элементов*. М.: Мир, 1980.
- [32] Сухарев А. Г., Тимохов А. В., Федоров В. В. *Курс методов оптимизации*. М.: Наука, 1986.
- [33] Тихонов А. Н., Арсенин В. Я. *Методы решения некорректных задач*. М.: Наука, 1979.
- [34] Фаддеев Д. К., Фаддеева В. П. *Вычислительные методы линейной алгебры*. М.: Физматгиз, 1963.
- [35] Яненко Н. Н. *Метод дробных шагов решения многомерных задач математической физики*. Новосибирск: Наука, 1967.