

БАКАЛАВР. АКАДЕМИЧЕСКИЙ КУРС

Под редакцией У. Г. Пирумова

ЧИСЛЕННЫЕ МЕТОДЫ

Учебник и практикум. 5-е издание



УМО ВО рекомендует
МО рекомендует

 Юрайт
издательство
biblio-online.ru



МОСКОВСКИЙ
АВИАЦИОННЫЙ
ИНСТИТУТ
(НАЦИОНАЛЬНЫЙ
ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ)



МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

ЧИСЛЕННЫЕ МЕТОДЫ

УЧЕБНИК И ПРАКТИКУМ
ДЛЯ АКАДЕМИЧЕСКОГО БАКАЛАВРИАТА

5-е издание, переработанное и дополненное

Под редакцией **У. Г. Пирумова**

*Рекомендовано Учебно-методическим отделом высшего
образования в качестве учебника
для студентов высших учебных заведений, обучающихся
по физико-математическим направлениям и специальностям*

*Допущено Министерством образования и науки
 Российской Федерации в качестве учебного пособия
 для студентов высших учебных заведений,
 обучающихся по направлению подготовки «Математика.
 Прикладная математика»*

**Книга доступна в электронной библиотечной системе
biblio-online.ru**

Москва • Юрайт • 2014

УДК 51
ББК 22.193я73
П33

Авторы:

Пирумов Ульян Гайкович — доктор технических наук, член-корреспондент РАН, заведующий кафедрой вычислительной математики и программирования МАИ (НИУ), дважды лауреат Государственной премии;

Гидаспов Владимир Юрьевич — кандидат физико-математических наук, старший научный сотрудник, заместитель заведующего кафедрой вычислительной математики и программирования МАИ (НИУ);

Иванов Игорь Эдуардович — кандидат физико-математических наук, старший научный сотрудник кафедры вычислительной математики и программирования МАИ (НИУ);

Ревизников Дмитрий Леонидович — доктор физико-математических наук, профессор кафедры вычислительной математики и программирования МАИ (НИУ);

Стрельцов Вячеслав Юльевич — кандидат физико-математических наук, доцент кафедры вычислительной математики и программирования МАИ (НИУ);

Формалев Владимир Федорович — доктор физико-математических наук, профессор кафедры вычислительной математики и программирования МАИ (НИУ).

Рецензенты:

Карташов Э. М. — доктор технических наук, профессор, заведующий кафедрой высшей и прикладной математики Московского государственного университета тонких химических технологий им. М. В. Ломоносова, заслуженный деятель науки РФ;

Миногаев В. Б. — доктор физико-математических наук, профессор, заведующий кафедрой общей и прикладной математики Московского государственного индустриального университета;

Пасконов В. М. — доктор физико-математических наук, профессор Московского государственного университета им. М. В. Ломоносова;

Яламов Ю. И. — доктор физико-математических наук, профессор, действительный член РАН.

Пирумов, У. Г.

Численные методы : учебник и практикум для академического бакалавриата / У. Г. Пирумов [и др.] ; под ред. У. Г. Пирумова, — 5-е изд., перераб. и доп. — М. : Издательство Юрайт, 2014. — 421 с. — Серия : Бакалавр. Академический курс.

ISBN 978-5-9916-4211-8

Учебное пособие содержит все традиционные разделы, предусмотренные программой по данной дисциплине. Материалдается по единой схеме, включающей в себя постановку задачи, описание алгоритма решения, детально разобранные типовые примеры и тщательно подобранный комплекс задач для самостоятельного решения.

Соответствует Федеральному государственному образовательному стандарту высшего образования четвертого поколения.

Для студентов технических вузов, аспирантов и преподавателей, а также инженеров, использующих в практической деятельности численные методы.

УДК 51
ББК 22.193я73

© Коллектив авторов, 1998

© Коллектив авторов, 2012, с изменениями

© ООО «Издательство Юрайт», 2014

ISBN 978-5-9916-4211-8

Предисловие

В учебном пособии изложены основные численные методы решения прикладных задач. Материал соответствует государственному образовательному стандарту по направлению «Прикладная математика». В некоторых разделах пособия, помимо традиционных, рассматриваются численные методы, развитые авторским коллективом.

Книга содержит четыре главы. Первая глава посвящена численным методам алгебры. В отличие от 1-го издания (М., 1998), она дополнена материалом по численным методам оптимизации для функций многих переменных и разделом по линейному программированию. Во второй главе описаны приближения функций, численное дифференцирование и интегрирование, а также добавлено изложение метода статистических испытаний. Третья глава посвящена численным методам решения обыкновенных дифференциальных уравнений. В нее дополнительно включен параграф по интегральным уравнениям. В четвертой главе приведены численные методы решения уравнений в частных производных. Каждая глава сопровождается введением, в котором дается краткое содержание главы, а также приводятся сведения из разделов высшей математики, необходимые для лучшего усвоения материала.

Авторы считают своим долгом выразить глубокую признательность рецензентам: профессорам В. Б. Миносцеву, Э. М. Карташову, В. М. Пасконову и Ю. И. Яламову. Авторы выражают благодарность за подготовку рукописи к изданию В. Н. Камзолову, Ал. А. Пирнпуу, В. В. Сеницкой.

Введение

В настоящее время численные методы являются мощным математическим средством решения многих научно-технических проблем. Это связано как с невозможностью в большинстве случаев получить точное аналитическое решение, так и со стремительным развитием компьютерной техники. Существуют многочисленные стандартные программы и объектно ориентированные пакеты прикладных программ. Однако научным и инженерно-техническим работникам важно понимать сущность основных численных методов и алгоритмов, поскольку зачастую интерпретация результатов расчетов нетривиальна и требует специальных знаний особенностей применяемых методов. Поэтому необходимо уделять большое внимание структуре погрешностей при решении конкретных задач и корректности вычислений.

СТРУКТУРА ПОГРЕШНОСТИ. Существует четыре источника погрешностей, полученных в результате численного решения: математическая и физическая модели, исходные данные, приближенность метода и ошибка округления.

Первые два источника погрешностей приводят к так называемой неустранимой погрешности. Эта погрешность может присутствовать, даже если решение сформулированной задачи найдено точно. Погрешность метода возникает из-за того, что точный оператор и исходные данные, в частности начальные и краевые условия, заменяются по определенным правилам приближенными. Так, производные заменяются их разностными аналогами, интегралы — суммами, функции — специальными многочленами; а при решении многих задач строятся бесконечные итерационные процессы, которые естественным образом прекращаются после конечного числа итераций. Как правило, погрешность метода может быть оценена и поддается контролю. Для некоторых методов такая оценка будет представлена.

Погрешность метода следует выбирать так, чтобы она была не более чем на порядок меньше неустранимой погрешности.

Погрешность округления возникает в связи с тем, что вычисления производятся с конечным числом значащих цифр. Округ-

ления производятся по следующему правилу: если в старшем из отбрасываемых разрядов стоит цифра меньше пяти, то содержимое сохраняемых разрядов не изменяется. В противном случае в младший сохраняемый разряд добавляется единица с тем же знаком, что и у самого числа. Очевидно, что погрешность, возникающая при округлении, не превышает младшего оставляемого разряда. Повторное округление проводить не следует, так как оно может привести к увеличению погрешности. Различают абсолютную и относительную погрешность. Пусть a — точное, вообще говоря, неизвестное числовое значение некоторой величины, а \bar{a} — известное приближенное числовое значение этой величины, тогда число

$$\Delta(\bar{a}) = |a - \bar{a}|$$

называют абсолютной погрешностью числа a , а величину

$$\delta(\bar{a}) = \frac{\Delta(\bar{a})}{|\bar{a}|}$$

— его относительной погрешностью. Нетрудно показать, что при сложении и вычитании складываются абсолютные погрешности, а при делении и умножении — относительные погрешности. Очевидно, что абсолютная погрешность характеризуется числом верных цифр после запятой, а относительная погрешность — числом верных значащих цифр.

Поскольку на современных компьютерах число записывается, как правило, с 10—12 десятичными знаками, то погрешность δ единичного округления порядка 10^{-10} — 10^{-12} обычно пренебрежимо мала по сравнению с неустранимой погрешностью и погрешностью метода. При решении больших задач производятся миллиарды операций и легко предположить, что ошибки могут заметно накапливаться, однако, поскольку они носят случайный характер, может происходить их взаимная компенсация. Зачастую строятся специальные алгоритмы, в частности итерационные, которые малочувствительны к ошибкам округления.

КОРРЕКТНОСТЬ. При численном решении основных задач необходимо знать какие-либо входные (исходные) данные — начальные, краевые (граничные) значения искомой функции, коэффициенты и правые части уравнения и т. д. Очевидно, что для исследователя важно установить, существует ли решение задачи, единственно ли оно и как оно зависит от входных данных.

Говорят, что задача поставлена корректно, если она разрешима при любых допустимых входных данных, когда имеется единственное решение и это решение непрерывно зависит от входных

данных, т. е. малому их изменению соответствует малое изменение решения. Тогда говорят, что задача устойчива.

Задача поставлена некорректно, если ее решение неустойчиво относительно входных данных, т. е. малому их изменению могут соответствовать большие изменения решения. Известно, что задача численного интегрирования корректна, а задача дифференцирования некорректна.

Классическим примером некорректной задачи является задача Коши для уравнения Лапласа. Эта некорректность исходной задачи проявляется и при ее численном решении.

В настоящее время развиты методы решения некорректных задач. К числу их относятся так называемые методы регуляризации, которые сводят решение исходной задачи к решению близкой к ней вспомогательной с некоторым малым параметром ϵ так, что при при $\epsilon \rightarrow 0$ решение вспомогательной задачи должно стремиться к решению исходной задачи. Далее для некоторых численных методов будут формулироваться условия корректности и устойчивости.

Глава 1

ЧИСЛЕННЫЕ МЕТОДЫ АЛГЕБРЫ

Методы, приведенные в главе, в значительной степени объединяет система понятий и идентичность подходов к решению задач.

Сначала приводятся численные методы решения систем линейных алгебраических уравнений. В линейной алгебре эту задачу называют первой основной задачей. К ней призывают задачи вычисления определителей и элементов обратной матрицы, которые иногда называют второй и третьей основными задачами линейной алгебры.

Далее излагаются численные методы уточнения корней нелинейных уравнений. В общем случае такие уравнения не имеют аналитических формул для своих корней или эти формулы слишком громоздки и неудобны для практического использования. В частности, в начале XIX века было доказано, что нелинейные уравнения выше четвертой степени неразрешимы в радикалах даже при использовании радикалов произвольной степени. Используемые численные методы уточнения корней нелинейных уравнений можно условно разделить на две группы. К первой группе относятся методы, которые назовем универсальными в том смысле, что в принципе они пригодны для отыскания корней уравнений любого вида. Эти методы носят итерационный характер. Подробно рассмотрены метод дихотомии (деления пополам), метод простой итерации, метод Ньютона и две его модификации: метод секущих и метод Ньютона с постоянным значением производной. Кратко дан метод парабол.

Для системы нелинейных уравнений приведено обобщение методов простой итерации, метода итераций Зейделя и метода Ньютона в приложении к одному нелинейному уравнению.

В последней теме главы освещается задача нахождения собственных значений и собственных векторов. Рассмотрены два метода отыскания всех собственных значений: итерационный метод вращения и метод интерполяции. Дополнительные сведения о методах решения полной проблемы собственных значений содержатся в разделе, посвященном методу Данилевского.

1.1. Основные понятия линейной алгебры

Приведем некоторые сведения из линейной алгебры, которые потребуются в дальнейшем. Рассмотрим прямоугольную матрицу

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

Две матрицы $A \equiv [a_{ij}]$ и $B \equiv [b_{ij}]$ размерности $m \times n$ равны друг другу, если $a_{ij} = b_{ij}$ для всех i и j .

Сумма двух матриц A и B размерности $m \times n$ есть матрица размерности $m \times n$:

$$A + B \equiv [a_{ij}] + [b_{ij}] \equiv [a_{ij} + b_{ij}].$$

Произведение матрицы A на скаляр α есть матрица размерности $m \times n$:

$$\alpha A \equiv \alpha [a_{ij}] = [\alpha a_{ij}].$$

Введем обозначение $c_{ik} = \sum_{j=1}^n a_{ij}b_{jk}$. Произведение матрицы A размерности $m \times n$ на матрицу B размерности $n \times r$ есть матрица C размерности $m \times r$:

$$C = AB \equiv [a_{ij}][b_{jk}] \equiv [c_{ik}].$$

Таким образом, элемент c_{ik} матрицы C есть сумма произведений i -й строки матрицы A на соответствующие элементы k -го столбца матрицы B , при этом число столбцов матрицы A должно равняться числу строк матрицы B . Из существования произведения AB вовсе не следует существование произведения BA . Для квадратных матриц ($m = n$) одного порядка существуют матрицы AB и BA , но, вообще говоря, $AB \neq BA$.

ОПРЕДЕЛИТЕЛЬ (ДЕТЕРМИНАНТ) квадратной матрицы будем обозначать $\det A$:

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Определитель квадратной матрицы с n^2 элементами (действительными или комплексными числами a_{ij}) есть сумма $n!$ членов вида $(-1)^r a_{1j_1} a_{2j_2} \dots a_{nj_n}$:

$$\det A = \sum_{j=1}^n a_{ij} A_{ij},$$

где A_{ij} — алгебраическое дополнение элемента a_{ij} . Отметим, что

$$(\det A)^2 \leq \prod_{i=1}^n \prod_{j=1}^n |a_{ij}|^2 \text{ (неравенство Адамара).}$$

Важным частным случаем квадратной матрицы является **ДИАГНОНАЛЬНАЯ МАТРИЦА**

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}.$$

При $a_{11} = a_{22} = \dots = a_{nn} = 1$ МАТРИЦА называется **ЕДИНЧНОЙ** и обозначается через E . Другим частным случаем квадратной матрицы является **ТРЕХДИАГОНАЛЬНАЯ МАТРИЦА**

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 & \dots & 0 & 0 \\ a_{21} & a_{22} & a_{23} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-1, n-1} & a_{n-1, n} \\ 0 & 0 & 0 & \dots & a_{n, n-1} & a_{nn} \end{bmatrix}.$$

С такими матрицами часто приходится иметь дело при решении дифференциальных уравнений. Матрица называется

АДАМАР ЖАК САЛОМОН (Hadamard Jacques Salomon; 1865—1963) — французский математик. Его основные работы относятся к теории дифференциальных уравнений с частными производными, теории функций комплексного переменного и теории чисел. Им написан известный учебник геометрии. В курсе математического анализа известны теорема Адамара — Коши о радиусе сходимости степенного ряда, а в теории определителей — неравенства Адамара для абсолютной величины определителя. Идеи Адамара оказали большое влияние на создание функционального анализа.

НИЖНЕЙ ТРЕУГОЛЬНОЙ, если все ее элементы, расположенные выше главной диагонали, равны нулю:

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 & 0 \\ a_{21} & a_{22} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & a_{nn} \end{bmatrix}.$$

Аналогично определяется **ВЕРХНЯЯ ТРЕУГОЛЬНАЯ** матрица. Квадратная матрица называется **СИММЕТРИЧНОЙ**, если ее элементы удовлетворяют соотношению $a_{ij} = a_{ji}$ ($i = 1, \dots, n$, $j = 1, \dots, n$). Если в матрице поменять строки со столбцами, то получим транспонированную матрицу

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix}.$$

Квадратная матрица A симметрична, если $A = A^T$.

Обозначим через x вектор-столбец и через x^T — вектор-строку. Матрица называется **ПОЛОЖИТЕЛЬНО ОПРЕДЕЛЕННОЙ**, если $x^T A x > 0$.

Матрица \bar{A} называется **КОМПЛЕКСНО СОПРЯЖЕННОЙ** с матрицей A , если ее элементы суть комплексно сопряженные с элементами матрицы A . Матрица $A^* = \bar{A}^T$ называется сопряженной с матрицей A . Если матрица A вещественная, то $A^* = A^T$.

Матрица A^{-1} называется **ОБРАТНОЙ** матрице A , если $AA^{-1} = E$. Необходимым и достаточным условием существования обратной матрицы A^{-1} является условие $\det A \neq 0$. Говорят, что в этом случае матрица A несобственная, или невырожденная.

МИНОРОМ порядка k матрицы A называется определитель k -го порядка, составленный из элементов, которые находятся на пересечении k строк и k столбцов матрицы A . РАНГОМ матрицы A называется число r такое, что все миноры порядка $r + 1$ и выше равны нулю.

ХАРАКТЕРИСТИЧЕСКИМ УРАВНЕНИЕМ матрицы A называется уравнение

$$\det \begin{bmatrix} a_{11} - \lambda & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} - \lambda & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{nn} - \lambda \end{bmatrix} = 0.$$

Корни λ_i этого уравнения называются СОБСТВЕННЫМИ ЧИСЛАМИ матрицы A . Левая часть уравнения $\det(A - \lambda E)$ называется характеристическим полиномом. СОБСТВЕННЫМ ВЕКТОРОМ матрицы A называется отличный от нуля вектор, удовлетворяющий условию $Ax = \lambda x$.

Две квадратные матрицы A и B называются ПОДОБНЫМИ, если $B = S^{-1}AS$, где S — невырожденная матрица и $A = SBS^{-1}$. Подобные матрицы имеют одинаковые характеристические многочлены и, следовательно, одинаковые собственные значения и определители, так как $\det A = \prod_{i=1}^n \lambda_i$. Собственные векторы матрицы A связаны с собственными векторами матрицы B соотношением $x = By$. Матрица A невырожденная, если все λ_i отличны от 0. Все собственные числа симметричной матрицы действительны.

Действительная матрица называется ОРТОГОНАЛЬНОЙ, если ее транспонированная матрица совпадает с обратной, т. е. $A^T = A^{-1}$ или $AA^T = A^TA = E$. Все собственные значения ортогональной матрицы по модулю равны единице. Строки (столбцы) ортогональной матрицы попарно ортогональны, суммы квадратов элементов каждой строки (столбца) ортогональной матрицы равны единице, определитель ортогональной матрицы равен ± 1 ; если матрица A ортогональна, то и матрица A^{-1} тоже ортогональна.

Всякая симметричная действительная матрица может быть приведена к диагональному виду подобным преобразованием

$$A = U\Lambda U^{-1},$$

где U — ортогональная матрица, Λ — диагональная матрица. Из свойств ортогональной матрицы следует, что

$$\Lambda = U^{-1}AU.$$

В дальнейшем зачастую рассматриваются множества, элементами которых являются числа, векторы, матрицы, функции. Сами множества обычно являются линейными нормированными пространствами, ибо в них определены операции сложения элементов и их умножения на число и введена норма каждого элемента $\|x\|$. НОРМОЙ называется вещественное неотрицательное число, удовлетворяющее следующим условиям:

- $\|x\| > 0$ при $x \neq 0$, $\|x\| = 0$ при $x = 0$.
- $\|ax\| = |a| \|x\|$.
- $\|x + y\| \leq \|x\| + \|y\|$.

Рассмотрим нормы в некоторых множествах. В множестве действительных чисел $\|x\| = |x|$. В множестве функций $x(t)$, опре-

деленных и непрерывных при $0 \leq t \leq 1$ (пространство С), определена чебышёвская норма $\|x\|_c = \max |x(t)|$. В конечномерном пространстве, элементами которого являются группы из n чисел $x = \{x_1, x_2, \dots, x_n\}$ (их можно считать координатами векторов), норма определена следующим образом:

$$\|x\|_p = \left[\frac{1}{n} \sum_{i=1}^n |x_i|^p \right]^{\frac{1}{p}}.$$

Между разными нормами существует соотношение

$$\|x\|_1 \leq \|x\|_2 \leq \|x\|_c \leq \sqrt{n} \|x\|_2 \leq n \|x\|_1,$$

где $\|x\|_c = \sup |x_i|$ по аналогии с пространством С. Поэтому из сходимости в одной из норм следует сходимость в остальных.

В пространстве квадратных матриц порядка n наиболее употребительны следующие нормы:

$$\begin{aligned} \|A\|_c &= \max_i \left[\sum_{j=1}^n |a_{ij}| \right], & \|A\|_1 &= \max_j \left[\sum_{i=1}^n |a_{ij}| \right], \\ \|A\|_M &= n \max_{i,j} |a_{ij}|, & \|A\|_E &= \left[\sum_{i,j=1}^n |a_{ij}|^2 \right]^{1/2}. \end{aligned}$$

Норма $\|A\|_c$ — максимальное число среди сумм модулей элементов строк матрицы, норма $\|A\|_1$ — максимальное число среди сумм модулей элементов столбцов матрицы. Эти нормы не имеют специального названия, норма $\|A\|_M$ называется максимальной, а $\|A\|_E$ — сферической или евклидовой. Необходимо отметить, что $\|A + B\| \leq \|A\| + \|B\|$; $\|AB\| \leq \|A\| \cdot \|B\|$.

Интересна связь между нормами матриц и векторов, на которые матрицы действуют. Норма матрицы называется согласованной с нормой вектора, если $\|Ax\| \leq \|A\| \cdot \|x\|$. Сферическая норма согласована с $\|x\|_2$, а максимальная согласована со всеми рассмотренными выше нормами.

Введем понятие предела векторов и матриц. Рассмотрим последовательность векторов $x^{(1)}, x^{(2)}, \dots$ с компонентами $x_1^{(1)}, \dots, x_n^{(1)}, x_1^{(2)}, \dots, x_n^{(2)}, \dots$. Если существуют пределы $\lim_{k \rightarrow \infty} x_i^{(k)}$ ($i = 1, 2, \dots, n$), то говорят, что вектор с компонентами x_1, x_2, \dots, x_n является пределом последовательности $x^{(1)}, x^{(2)}, \dots$. Аналогично определяется предел последовательности матриц. Необходимым и достаточным условием сходимости векторов $x^{(k)}$ к x является условие $\|x^{(k)} - x\| \rightarrow 0$, при этом $\|x^{(k)}\| \rightarrow \|x\|$. Аналогичное утверждение справедливо для матриц.

1.2. Основные трудности решения систем линейных уравнений. Классификация методов решения

Требуется найти решение системы линейных уравнений

$$Ax = b, \quad (1.1)$$

где $A = (a_{ij})$ — квадратная матрица коэффициентов при неизвестных; $x = (x_j)$ — вектор-столбец неизвестных; $b = (b_j)$ — вектор-столбец правых частей системы. С точки зрения классической теории линейных алгебраических систем их решение не вызывает затруднений. По правилу Крамера система n линейных уравнений с n неизвестными имеет единственное решение, если определитель системы отличен от нуля ($\det A \neq 0$) и значение каждого из неизвестных вычисляется как отношение двух определителей порядка n , т. е.

$$x_j = \det A_j / \det A, \quad j = 1, \dots, n. \quad (1.2)$$

Здесь $\det A_j$ — определитель матрицы, получаемой заменой j -го столбца матрицы A столбцом правых частей. При непосредственном вычислении определителей как алгебраической суммы $n!$ произведений элементов для отыскания решения системы линейных уравнений по правилу Крамера требуется приблизительно $n \cdot n!$ арифметических операций типа умножения*. Будет показано, что использование метода исключения Гаусса позволяет уменьшить время, необходимое для решения задачи (1.1), до величины менее одной секунды.

Другое важное обстоятельство, связанное с решением систем линейных алгебраических уравнений, состоит в следующем. С точки зрения теории линейных систем различаются два случая: определитель матрицы системы не равен нулю ($\det A \neq 0$), т. е. система уравнений является невырожденной, либо определитель матрицы системы равен нулю ($\det A = 0$), в этом случае система называется вырожденной. Во втором случае система ли-

*Под операциями типа умножения подразумеваются операции умножения, деления и возведения в степень.

КРАМЕР ГАБРИЕЛЬ (Cramer Gabriel; 1704—1752) — швейцарский математик. Крамер установил правило решения систем линейных уравнений с буквенными коэффициентами (правило Крамера), а также заложил основы теории определителей.

бо не имеет решения, либо имеет неединственное решение. С точки зрения практических вычислений существуют «почти невырожденные системы» — системы, определитель которых близок к нулю, но отличен от нуля ($\det A \approx 0$). Небольшие изменения коэффициентов матрицы системы или правых частей системы в «почти невырожденных системах» могут привести к большим погрешностям решения.

Все эти случаи хорошо иллюстрируются на примере решения системы двух линейных уравнений:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, \\ a_{21}x_1 + a_{22}x_2 = b_2. \end{cases}$$

На рис. 1.1 каждому уравнению соответствует прямая на плоскости x_1Ox_2 , а точка пересечения этих прямых есть решение системы (1.1).

Если $\det A = 0$, то наклоны прямых равны и они либо параллельны, либо совпадают. При $\det A \approx 0$ небольшие погрешности в коэффициентах и правых частях могут привести к большим погрешностям в решении, т. е. к неточному определению положения точки пересечения.

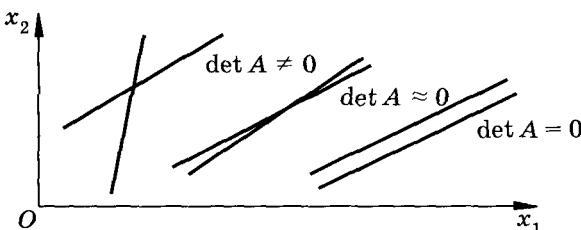


Рис. 1.1

Системы такого типа, в которых есть малые погрешности в коэффициентах системы или в правых частях (эти погрешности могут быть, в частности, результатом округлений при вычислениях или записи чисел в память компьютера), называются ПЛОХО ОБУСЛОВЛЕННЫМИ.

Плохо обусловленная система геометрически соответствует почти параллельным прямым. В теоретических исследованиях обусловленность характеризуется числом обусловленности $\chi = \|A\| \cdot \|A^{-1}\|$, при любой норме $\chi \geq 1$. Чем больше это число, тем хуже обусловленность системы; так, при $\chi \approx 10^3 - 10^4$ система уже плохо обусловлена. На практике, однако, ограничиваются проверкой условия $\det A \approx 0$.

Системы линейных уравнений можно было бы рассматривать как частный случай нелинейных уравнений. Однако относитель-

ная простота линейных систем обусловила появление специальных высокоеффективных методов их решения. По этой же причине введение и пояснение основных понятий могут быть выполнены относительно просто. Вместе с тем линейные системы представляют собой большой самостоятельный интерес по двум причинам. Во-первых, к системам линейных уравнений приводят многие практические задачи. Во-вторых, важным и неиссякаемым источником систем линейных алгебраических уравнений являются практически все разделы вычислительной математики: нахождение решения системы линейных алгебраических уравнений необходимо в задачах уточнения корней систем нелинейных уравнений, аппроксимации функций, отыскания собственных значений и собственных векторов матрицы, решения обыкновенных дифференциальных уравнений, уравнений в частных производных и т. д. Методы решения линейных уравнений и метод прогонки используются и для некоторых классов интегральных уравнений и интегродифференциальных уравнений.

Будут рассмотрены наиболее употребительные прямые и итерационные методы. ПРЯМЫЕ МЕТОДЫ дают решение задачи за конечное (точно определяемое для каждого метода) число операций. Здесь намеренно не употребляется термин «точные методы», так как из-за ошибок округления при вычислениях с конечным числом знаков решение всегда получается с погрешностями. Причины этого и способы уменьшения влияния ошибок округления будут обсуждены ниже. Из прямых методов будут рассмотрены метод Гаусса, метод Гаусса с выбором главного элемента для системы линейных алгебраических уравнений общего вида и метод прогонки для систем линейных уравнений с трехдиагональной матрицей.

ГÁУСС КАРЛ ФРИДРИХ (Gauß Carl Friedrich; 1777—1855) — великий немецкий математик, которому также принадлежат работы по астрономии, физике, геодезии. Г. доказал основную теорему алгебры о существовании хотя бы одного корня у всякого алгебраического уравнения. Ему принадлежат важные работы по теории чисел, дифференциальной геометрии, теории вероятностей, теории бесконечных рядов. Вычисляя погрешности при измерениях, Г. предложил метод наименьших квадратов. Г. разработал теорию геодезических линий, построил первую квадратичную форму, доказал «славную» теорему (*théoreme érgerium*) об одинаковости меры кривизны (K) соответствующих точек поверхностей, совмещаемых путем изгибаания. Отличительными чертами творчества Г. являются глубокая органическая связь в его исследованиях между теоретической и прикладной математикой и широта проблематики.

ИТЕРАЦИОННЫЕ МЕТОДЫ дают решение как предел бесконечной последовательности приближенных решений, в которых каждое последующее более точное приближение находится по уже найденному предыдущему решению (или предыдущим решениям). Из итерационных методов решения систем линейных алгебраических уравнений рассмотрены метод простой итерации и метод Зейделя.

1.3. Метод исключения Гаусса

Систему уравнений (1.1) представим в виде

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots \dots \dots \dots \dots \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \end{cases} \quad (1.3)$$

или

$$\sum_{j=1}^n a_{ij}x_j = b_i, \quad i = 1, \dots, n.$$

Известно большое число схем метода исключения, приспособленных для ручного или машинного счета матриц общего или специального вида.

Метод Гаусса можно интерпретировать как метод, в котором первоначально матрица приводится к верхней треугольной форме (прямой ход), а далее — к единичной (обратный ход). Очевидно, что если матрица единичная, то $x_i = b_i$.

Пусть матрица системы (1.3) — верхняя треугольная, поэтому $a_{ij} = 0$ при $i > j$, т. е. все элементы ниже главной диагонали равны нулю. Тогда из последнего уравнения сразу определяем x_n . Подставляя x_n в предпоследнее уравнение, находим x_{n-1} и т. д.

Общие формулы имеют вид

$$x_k = \frac{b_k}{a_{kk}} \quad \text{при } k = n, \\ x_k = \frac{1}{a_{kk}} \left[b_k - \sum_{l=k+1}^n a_{kl}x_l \right] \quad \text{при } k = n-1, n-2, \dots, 1. \quad (1.4)$$

При $k > l$ коэффициенты $a_{kl} = 0$.

Приведем матрицу системы (1.3) к верхней треугольной. Вычтем из второго уравнения системы (1.3) первое, умноженное на такое число, при котором коэффициент при x_1 обратится в нуль. То же проделаем со всеми остальными уравнениями. В результате все коэффициенты первого столбца, лежащие ниже главной диагонали, обратятся в нуль. Затем, используя второе уравнение, обратим в нуль соответствующие коэффициенты второго столбца. Последовательно продолжая этот процесс, приведем матрицу системы к верхней треугольной форме.

Запишем общие формулы метода Гаусса. Пусть проведено исключение коэффициентов из $(k - 1)$ -го столбца. Тогда останутся уравнения с ненулевыми элементами ниже главной диагонали:

$$\sum_{j=k}^n a_{ij}^{(k)} x_j = b_i^{(k)}, \quad k \leq i \leq n.$$

Умножим k -ю строку на число $c_{mk} = a_{mk}^{(k)} / a_{kk}^{(k)}$, $m > k$ и вычтем из m -й строки. Первый ненулевой элемент этой строки обратится в нуль, а остальные изменятся по формулам

$$a_{ml}^{(k+1)} = a_{mk}^{(k)} - c_{mk} a_{kl}^{(k)},$$

$$b_m^{(k+1)} = b_m^{(k)} - c_{mk} b_k^{(k)}, \quad k < m.$$

Проведя вычисления по этим формулам при всех указанных индексах, обратим в нуль элементы k -го столбца, лежащие ниже главной диагонали. Аналогичная процедура приводит матрицу системы к верхней треугольной форме, при этом весь процесс приведения называется ПРЯМЫМ ХОДОМ МЕТОДА ГАУССА. Вычисление неизвестных по формулам (1.4) называют ОБРАТНЫМ ХОДОМ метода.

Обратный ход можно совершить иначе, если обратить в нуль и все коэффициенты, лежащие выше главной диагонали. Например, элементы n -го столбца обращаются в нуль, если $a_{nn}^{(n)}$ умножить на $(-a_{kn}^{(n)} / a_{nn}^{(n)})$ и сложить с соответствующей строкой. Аналогично обращаются в нуль и все остальные столбцы. Если, кроме того, разделить затем каждое уравнение на соответствующий элемент, стоящий на главной диагонали, то матрица системы становится единичной, а неизвестные $x_i = b_i^{(2n)}$, где $b_i^{(2n)}$ — коэффициенты правой части i -го уравнения после указанных преобразований.

На некотором шаге прямого хода может оказаться, что коэффициент $a_{ii}^{(k)} \neq 0$, но мал по сравнению с остальными элементами

матрицы системы и, в частности, мал по сравнению с элементами первого столбца. Деление коэффициентов системы на малую величину может привести к значительным ошибкам округления.

Для уменьшения ошибок округления поступают следующим образом. Среди элементов первого столбца $a_{ki}^{(k)}$ каждой промежуточной матрицы выбирают наибольший по модулю (главный) элемент и путем перестановки i -й строки со строкой, содержащей главный элемент, добиваются того, что главный элемент становится ведущим. Такая модификация метода исключения Гаусса называется методом Гаусса с выбором главного элемента. Случай появления нулевых элементов обходится при этом сам собой.

Для реализации метода требуется примерно $n^3/3$ операций типа умножения и $n^3/3$ операций типа сложения*. Полезно помнить, что оценка числа операций определяется в основном операциями, затрачиваемыми при выполнении прямого хода метода Гаусса. Обратный ход метода Гаусса требует примерно n^2 операций. Следовательно, если требуется решить несколько систем линейных алгебраических уравнений вида $Ax = b$ с одной и той же матрицей и различными правыми частями, то общее число операций при решении S систем будет оцениваться величиной $(2/3)n^3 + Sn^2$. В этом случае целесообразно реализовать алгоритм метода Гаусса в виде двух подпрограмм: первая подпрограмма должна реализовывать прямой ход алгоритма и получать на выходе верхнюю треугольную матрицу, а вторая подпрограмма должна, используя полученную матрицу, вычислять решение системы для произвольной правой части.

1.4. Вычисление определителя и обратной матрицы методом исключения

Определитель треугольной матрицы равен произведению диагональных элементов. В результате выполнения прямого хода метода исключения система линейных уравнений приводится к верхней треугольной матрице. Следовательно, определитель мат-

*Под операциями типа сложения подразумеваются операции сложения, вычитания, взятия модуля, а также отделения целой части.

рицы системы может быть вычислен как произведение диагональных элементов:

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

где k — количество перестановок строк при использовании метода исключения с выбором главного элемента. Вычисление определителя требует примерно $(2/3)n^3$ операций.

Для вычисления обратной матрицы вспомним, что $AA^{-1} = E$, где A^{-1} — обратная матрица, а E — единичная матрица. Обозначим элементы обратной матрицы через α . Тогда соотношение $AA^{-1} = E$ можно записать в виде

$$\sum_{k=1}^n a_{ik} a_{kl} = \delta_{il}, \quad 1 \leq i \leq n, \quad (1.5)$$

где δ_{il} — символ Кронекера,

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

Видно, что если рассматривать l -й столбец обратной матрицы как вектор, то он является решением системы (1.5) с матрицей A со специальной правой частью и вектором правых частей, в котором на l -м месте стоит единица, а на остальных — нули. Решение этой системы для каждого значения l дает элементы l -го столбца обратной матрицы. Таким образом, для обращения матрицы необходимо решить n систем линейных уравнений n -го порядка с одинаковой матрицей A и различными правыми частями. Так как приведение матрицы A к треугольной выполняется только один раз, то для обращения матрицы с использованием метода исключения требуется примерно $2n^3$ операций.

1.5. Метод прогонки для решения систем линейных уравнений с трехдиагональной матрицей

Системы линейных уравнений с трехдиагональной матрицей коэффициентов при неизвестных являются наиболее важным и распространенным случаем систем специального вида. В таких системах отличны от нуля только элементы, лежащие на главной

диагонали и на нижней и верхней диагоналях, прилегающих к ней. К системам с трехдиагональными матрицами приводят, например, задачи о сплайн-интерполяции, о решении разностными методами обыкновенных дифференциальных уравнений и уравнений в частных производных.

Метод прогонки принадлежит к числу прямых методов решения систем линейных уравнений и используется в тех случаях, в которых многие коэффициенты матрицы равны нулю. Это обстоятельство учтено при реализации метода прогонки, в котором исключаются преобразования с нулевыми элементами. В методе прогонки применительно к системе линейных уравнений, имеющих трехдиагональную матрицу, можно выделить следующие этапы.

- Приведение трехдиагональной матрицы к верхней треугольной (прямой ход). В случае трехдиагональной матрицы это означает приведение к двухдиагональной, т. е. приведение исходной системы к системе, содержащей по два неизвестных в каждом уравнении, кроме последнего, в котором содержится только одно неизвестное.
- Запись обратного хода в виде $x_i = P_{i+1}x_{i+1} + Q_{i+1}$, так как преобразованная матрица — двухдиагональная.
- Вывод рекуррентного соотношения для P_{i+1} и Q_{i+1} через P_i и Q_i и получение соотношения для P_2 и Q_2 ($P_1 = Q_1 = 0$).
- Осуществление обратного хода метода прогонки и определение всех неизвестных.

Рассматриваемый метод прогонки представляет собой модификацию метода исключения Гаусса, использующую специальный регулярный вид матрицы системы. Запишем систему линейных алгебраических уравнений с трехдиагональной матрицей в виде

$$\begin{aligned} a_i x_{i-1} - b_i x_i + c_i x_{i+1} &= d_i, \quad i = 1, 2, \dots, n, \\ a_1 = c_n &= 0. \end{aligned} \tag{1.6}$$

Запись (1.6) представляет собой так называемый КАНОНИЧЕСКИЙ ВИД СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ МЕТОДА ПРОГОНКИ. При этом матрица системы (1.6) имеет вид

$$A = \begin{bmatrix} -b_1 & c_1 & 0 & 0 & \dots & 0 & 0 \\ a_2 & -b_2 & c_2 & 0 & \dots & 0 & 0 \\ \dots & a_3 & -b_3 & c_3 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \dots & -b_{n-1} & c_{n-1} \\ 0 & 0 & 0 & \dots & \dots & a_n & -b_n \end{bmatrix}.$$

Прямой ход метода прогонки сводится к исключению неизвестного x_{i-1} в каждом уравнении системы. Получаемая в результате прямого хода система содержит в каждом уравнении только два неизвестных x_i и x_{i+1} , и матрица ее — верхняя треугольная с двумя диагоналями. Запишем i -ю строку преобразованной двухдиагональной матрицы в виде

$$x_i = P_{i+1}x_{i+1} + Q_{i+1}. \quad (1.7)$$

Если система (1.6) приведена к виду (1.7), то обратный ход метода Гаусса очевиден. Однако использование общих алгоритмов прямого и обратного хода нецелесообразно. Построим эффективную вычислительную схему, которая и составляет суть метода прогонки. Для этого, уменьшив в (1.7) индекс на единицу, запишем

$$x_{i-1} = P_i x_i + Q_i.$$

Подставляя x_{i-1} в систему (1.6), получим соотношение

$$a_i(P_i x_i + Q_i) - b_i x_i + c_i x_{i+1} = d_i,$$

из которого нетрудно получить

$$x_i = [b_i - a_i P_i]^{-1} c_i x_{i+1} + [b_i - a_i P_i]^{-1} [a_i Q_i - d_i].$$

Сравнивая это соотношение с (1.7), можем записать рекуррентные соотношения

$$P_{i+1} = c_i [b_i - a_i P_i]^{-1}, \quad Q_{i+1} = [a_i Q_i - d_i] [b_i - a_i P_i]^{-1} \quad (1.8)$$

для вычисления так называемых ПРОГОНОЧНЫХ КОЭФФИЦИЕНТОВ.

Подчеркнем, что последующие значения прогоночных коэффициентов P_{i+1} , Q_{i+1} вычисляются только по известным коэффициентам системы (1.6) и известным предыдущим значениям прогоночных коэффициентов P_i , Q_i .

Для начала прямого хода метода прогонки необходимо задать начальные (стартовые) значения прогоночных коэффициентов, например P_1 , Q_1 . Отметим, что, вообще говоря, начальные значения коэффициентов P_1 , Q_1 в рассмотренной схеме вычислений не требуются, так как значения коэффициентов P_2 , Q_2 вычисляются только через коэффициенты первого уравнения системы (1.6): при $i = 1$ из (1.6) получаем соотношение $-b_1 x_1 + c_1 x_2 = d_1$. Сравнивая это выражение с (1.7) при $i = 1$, получаем $P_2 = c_1/b_1$; $Q_2 = -d_1/b_1$, а значение x_1 в обратном ходе вычисляем по соотношению $x_1 = P_2 x_2 + Q_2$. Использование P_1 , Q_1 в качестве начальных значений целесообразно по двум причинам: сохраняется однородность вычислительного алгоритма для всех $i = 2, 3, \dots, n$; упрощается обсуждение и доказательство условия корректности и устойчивости метода прогонки. Из того обстоятельства, что

коэффициенты P_2, Q_2 не зависят от P_1, Q_1 (в соотношениях (1.8) при $i = 1$ коэффициенты P_1, Q_1 умножаются на $a_1 = 0$), следует, что можно задать любые значения для прогоночных коэффициентов P_1, Q_1 . Далее будет ясно, почему удобно положить $P_1 = Q_1 = 0$. Для начала обратного хода метода прогонки необходимо для вычисления $x_i = P_{i+1}x_{i+1} + Q_{i+1}$ задать значение x_{n+1} . Так как $c_n = 0$, то из первого соотношения (1.8) вытекает, что $P_{n+1} = 0$ и, следовательно, можно задать любое значение для x_{n+1} . Обычно полагают $x_{n+1} = 0$, и тогда $x_n = Q_{n+1}$.

Метод прогонки устойчив, если $|P_i| \leq 1$. Метод прогонки корректен, если $b_i - a_i P_i \neq 0$.

Отметим, что при устойчивости метода прогонки ошибки округления не возрастают, а подавляются. Пусть $\tilde{x}_i, \tilde{x}_{i+1}$ — вычисленные с погрешностями значения решения. Пусть при этом коэффициенты P_{i+1}, Q_{i+1} вычисляются точно. Тогда по (1.7)

$$\tilde{x}_i - x_i = P_{i+1}(\tilde{x}_{i+1} - x_{i+1})$$

или

$$\varepsilon_i = \tilde{x}_i - x_i = P_{i+1}\varepsilon_{i+1}.$$

Из этой формулы при $|P_{i+1}| \leq 1$ следует, что в данном случае погрешность не возрастает.

Достаточным условием корректности метода прогонки и устойчивости его к погрешностям является условие преобладания диагональных коэффициентов:

$$|b_i| \geq |a_i| + |c_i|, \quad i = 1, 2, \dots, n.$$

В самом деле, если хотя бы для одного значения i выполняется строгое неравенство $|P_i| < 1$, то можно записать цепочку неравенств:

$$|P_{i+1}| = \frac{|c_i|}{|b_i - a_i P_i|} \leq \frac{|c_i|}{|b_i| - |a_i||P_i|} \leq \frac{|c_i|}{|a_i| + |c_i| - |a_i||P_i|} < 1.$$

Было показано, что можно положить $|P_1| = 0$ и тогда, во-первых, для всех i выполняется $|P_i| < 1$, что обеспечивает затухание погрешности, и, во-вторых, для всех i выполняется условие $|b_i - a_i P_i| > |c_i| \geq 0$ и, таким образом, не возникает ситуаций деления на нуль. Так как условие $|b_i| \geq |a_i| + |c_i|$ является только достаточным, то невыполнение его не означает нарушения корректности и устойчивости.

Для реализации метода требуется примерно $8n$ операций, из которых $3n$ составляют операции типа умножения и $5n$ — операции типа сложения. При численном решении дифференциаль-

ных уравнений используются различные варианты метода прогонки: метод встречных прогонок, потоковая прогонка, матричная прогонка для систем векторных уравнений. Отметим, что

$$\det A = \prod_{i=1}^n [a_i P_i - b_i].$$

Метод прогонки обобщается на случай, при котором в системе (1.6) a_i , b_i , c_i — квадратные матрицы размерности $m \times m$, а x_i , d_i — векторы размерности m . Тогда соотношения (1.7) и (1.8) метода прогонки, в которых все действия совершаются уже над матрицами, а не над скалярами, сохраняются, а $[b_i - a_i P_i]^{-1}$ в (1.8) является соответствующей обратной матрицей.

Условие устойчивости матричной прогонки выглядит как $|\det P_i| < 1$, а условие корректности и устойчивости имеет вид $\det(B_i^{-1}C_i) + \det(B_i^{-1}A_i) < 1$.

1.6. Итерационные методы решения линейных уравнений

В итерационных методах предполагается осуществление трех следующих этапов: построение для вычисления последовательных приближений итерационного процесса, сходящегося к точному решению (т. е. построение последовательности векторов $x^{(0)}, x^{(1)}, x^{(2)}, \dots, x^{(k)}$, сходящейся к точному решению x^*); определение критерия сходимости этого процесса, позволяющего определить момент достижения требуемой точности; исследование скорости сходимости и оптимизация итерационного процесса с целью уменьшения числа операций, необходимых для достижения требуемой точности.

Итерационные методы позволяют получить решение с наперед заданной точностью, если доказана сходимость метода. Стого точного решения итерационные методы не дают, поскольку оно достигается как предел последовательности векторов. Прямой метод, вообще говоря, дает точное решение, но из-за ошибок округления, имеющих место на любых компьютерах, оно не может быть достигнуто, и *a priori* даже трудно оценить, насколько это решение отличается от точного. В связи с отмеченным итерационные методы иногда позволяют получить решение с большей точностью, чем прямые.

Рассмотрим несколько итерационных методов решения линейных уравнений.

Метод простой итерации

В методе простой итерации система (1.1) линейных алгебраических уравнений $Ax = b$ приводится к эквивалентной системе вида

$$x = \alpha x + \beta. \quad (1.9)$$

Решение системы (1.9) и, следовательно, решение исходной системы (1.1) ищется как предел последовательности векторов при $k \rightarrow \infty$:

$$x^{(k+1)} = \alpha x^{(k)} + \beta, \quad k = 0, 1, 2, \dots, \quad (1.10)$$

где $x^{(0)}$ — начальное приближение для вектора решения.

Достаточное условие сходимости метода простой итерации определяется следующей теоремой.

ТЕОРЕМА 1. Если какая-либо норма матрицы α , согласованная с рассматриваемой нормой вектора x , меньше единицы ($\|\alpha\| < 1$), то последовательность $x^{(k)}$ в методе простой итерации сходится к точному решению x^* системы (1.9) со скоростью, не меньшей скорости геометрической прогрессии со знаменателем $q \leq \|\alpha\|$ при любом начальном приближении $x^{(0)}$.

ДОКАЗАТЕЛЬСТВО. Для доказательства теоремы введем погрешность $\varepsilon^{(k)} = x^* - x^{(k)}$. Вычитая из соотношения $x^* = \alpha x^* + \beta$ равенство (1.10), получаем $\varepsilon^{(k+1)} = \alpha \varepsilon^{(k)}$. Переходя к нормам, имеем

$$\|\varepsilon^{(k)}\| = \|\alpha^k \varepsilon^{(0)}\| \leq \|\alpha^k\| \cdot \|\varepsilon^{(0)}\| \leq \|\alpha\|^k \cdot \|\varepsilon^{(0)}\|.$$

Отметим, что неравенство $\|\alpha^k \varepsilon^{(0)}\| \leq \|\alpha^k\| \cdot \|\varepsilon^{(0)}\|$ из предыдущего выражения является условием согласованности нормы матрицы и вектора. Если $\|\alpha\| < 1$, то при любом векторе начальной погрешности (или иначе — при любом начальном векторе $x^{(0)}$) норма погрешности $\|\varepsilon^{(k)}\|$ *) стремится к нулю не медленнее геометрической прогрессии со знаменателем $q \leq \|\alpha\|$.

Если в качестве нормы матрицы выбрать норму $\|\alpha\|_c = \max_i \sum_{j=1}^n |\alpha_{ij}|$ или $\|\alpha\|_1 = \max_j \sum_{i=1}^n |\alpha_{ij}|$, то для решения вопроса о сходимости метода простой итерации можно воспользоваться

*) Верхний индекс k в скобках означает номер итерации, а без скобок, как обычно, степень.

следствием из теоремы 1: метод простой итерации сходится, если для матрицы α выполняется одно из следующих условий:

$$\begin{aligned} \sum_{j=1}^n |\alpha_{ij}| &< 1, \quad i = 1, 2, \dots, n, \\ \sum_{i=1}^n |\alpha_{ij}| &< 1, \quad j = 1, 2, \dots, n. \end{aligned} \tag{1.11}$$

Простейшим и распространенным способом приведения системы $Ax = b$ к виду (1.9), удобному для итераций, является выделение диагональных элементов, при этом каждое i -е уравнение разрешается относительно i -го неизвестного:

$$a_{ii}x_i = \left[b_i - \sum_{j=1, j \neq i}^n a_{ij}x_j \right], \quad i = 1, 2, \dots, n, \tag{1.12}$$

и метод простой итерации запишется в виде

$$a_{ii}x_i^{(k+1)} = \left[b_i - \sum_{j=1, j \neq i}^n a_{ij}x_j^{(k)} \right].$$

Матрица α при этом имеет вид

$$\left[\begin{array}{ccccc} 0 & -\frac{a_{12}}{a_{11}} & -\frac{a_{13}}{a_{11}} & \dots & -\frac{a_{1n}}{a_{11}} \\ -\frac{a_{21}}{a_{22}} & 0 & -\frac{a_{23}}{a_{22}} & \dots & -\frac{a_{2n}}{a_{22}} \\ \dots & \dots & \dots & \dots & \dots \\ -\frac{a_{n1}}{a_{nn}} & -\frac{a_{n2}}{a_{nn}} & \dots & -\frac{a_{n,n-1}}{a_{nn}} & 0 \end{array} \right].$$

Элемент этой матрицы можно записать в виде $\alpha_{ij} = -\frac{a_{ij}}{a_{ii}}[1 - \delta_{ij}]$, где δ_{ij} — символ Кронекера. В этом случае достаточное условие сходимости метода простой итерации может быть сформулировано как условие преобладания диагональных элементов матрицы A , что следует из (1.11) и записи матрицы α , т. е.

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

КРОНЕКЕР ЛЕОПОЛЬД (Kronecker Leopold; 1823—1891) — немецкий математик, основные работы которого относятся к алгебре и теории чисел, теории квадратичных форм и теории групп, а также теории эллиптических функций. Большое значение имеют его исследования по арифметической теории алгебраических величин.

Еще раз подчеркнем, что рассмотренные формы условия сходимости метода итерации являются лишь достаточными. Их выполнение гарантирует сходимость метода, но их невыполнение в общем случае не означает, что метод простой итерации расходится. Необходимым и достаточным условием сходимости метода простой итерации является условие того, что целая часть $|\lambda_{\max}| < 1$ (где λ_{\max} — максимальное по модулю собственное значение матрицы A); это условие редко используется в практике вычислений.

Перейдем к вопросу об оценке погрешности решения. Представляют интерес два соотношения оценки погрешности решения $\varepsilon^{(k)} = x^* - x^{(k)}$: первое связывает норму погрешности с нормой разности двух последовательных приближений $\|x^{(k)} - x^{(k-1)}\|$ и может быть использовано для оценки погрешности только в процессе вычислений; второе связывает норму погрешности с нормами вектора начального приближения $\|x^{(0)}\|$ и вектора свободного члена $\|\beta\|$ в системе (1.9). Необходимые соотношения даются следующими двумя теоремами.

ТЕОРЕМА 2. Если какая-либо норма матрицы α , согласованная с рассматриваемой нормой вектора x , меньше единицы ($\|\alpha\| < 1$), то имеет место следующая оценка погрешности:

$$\|\varepsilon^{(k)}\| \leq \frac{\|\alpha\|}{1 - \|\alpha\|} \|x^{(k)} - x^{(k-1)}\|. \quad (1.13)$$

ДОКАЗАТЕЛЬСТВО. Вычтем из равенства $x^* = \alpha x^* + \beta$ равенство (1.10):

$$x^* - x^{(k)} = \alpha [x^* - x^{(k-1)}].$$

Вычитая из обеих частей значение приближения $x^{(k-1)}$, преобразуем это соотношение к виду

$$x^* - x^{(k-1)} = [x^{(k)} - x^{(k-1)}] + \alpha [x^* - x^{(k-1)}].$$

Перейдя к нормам, получим

$$\|\varepsilon^{(k-1)}\| \leq \|x^{(k)} - x^{(k-1)}\| + \|\alpha\| \cdot \|\varepsilon^{(k-1)}\|$$

или

$$[1 - \|\alpha\|] \cdot \|\varepsilon^{(k-1)}\| \leq \|x^{(k)} - x^{(k-1)}\|.$$

Так как по условию теоремы $[1 - \|\alpha\|] > 0$, то

$$\|\varepsilon^{(k-1)}\| \leq \frac{1}{1 - \|\alpha\|} \|x^{(k)} - x^{(k-1)}\|.$$

Используя соотношение $\varepsilon^{(k)} = \alpha\varepsilon^{(k-1)}$, из которого следует, что $\|\varepsilon^{(k)}\| \leq \|\alpha\| \cdot \|\varepsilon^{(k-1)}\|$, окончательно получим:

$$\|\varepsilon^{(k)}\| \leq \frac{\|\alpha\|}{1 - \|\alpha\|} \|x^{(k)} - x^{(k-1)}\|.$$

ТЕОРЕМА 3. Если какая-либо норма матрицы α , согласованная с рассматриваемой нормой вектора x , меньше единицы ($\|\alpha\| < 1$), то имеет место следующая оценка погрешности:

$$\|\varepsilon^{(k)}\| \leq \|\alpha\|^k \cdot \|x^{(0)}\| + \frac{\|\alpha\|^k}{1 - \|\alpha\|} \|\beta\|.$$

Сделаем два замечания. Во-первых, соотношение (1.13) может быть записано в виде

$$\|\varepsilon^{(k)}\| \leq \frac{\|\alpha\|^k}{1 - \|\alpha\|} \|x^{(1)} - x^{(0)}\|,$$

позволяющим получить оценку погрешности по результатам двух первых итераций. Во-вторых, при использовании метода итераций в качестве оценки погрешности вычислений иногда рекомендуется использовать норму разности двух последовательных приближений. Из соотношений для погрешности следует, что в общем случае это неверно. Если норма $\|\alpha\|$ близка к единице, то коэффициент при $\|x^{(k)} - x^{(k-1)}\|$ может быть достаточно большим.

Погрешности последовательных итераций связаны соотношением

$$\|\varepsilon^{(k+1)}\| = \|\alpha\| \cdot \|\varepsilon^{(k)}\|,$$

т. е. погрешность изменяется на шаге линейно. Говорят, что метод имеет *линейную сходимость* или первый порядок сходимости. Вместе с тем количество итераций, необходимое для достижения требуемой точности, зависит от значения $\|\alpha\|$ и начального приближения $x^{(0)}$.

Итак, на примере метода простой итерации продемонстрированы три этапа итерационных методов: построение последовательности векторов, порождаемых формулой (1.10); определение условия сходимости по теореме 1 и оценка скорости сходимости с помощью теорем 2 и 3.

Метод Зейделя

В методе простой итерации не используется кажущаяся очевидной возможность улучшения сходимости итерационного процесса — немедленное введение в расчет вновь вычисленных компонент

вектора $x^{(k)}$. Эта возможность используется в итерационном методе Зейделя. Итерационный процесс для системы (1.9) выполняется при этом по соотношению

$$x_i^{(k+1)} = \sum_{j=1}^{i-1} \alpha_{ij} x_j^{(k+1)} + \sum_{j=i+1}^n \alpha_{ij} x_j^{(k)} + \beta_i, \quad i = 1, 2, \dots, n \quad (1.14)$$

или для системы (1.1)

$$a_{ii} x_i^{(k+1)} = b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(k)}.$$

Не вдаваясь в подробности, отметим, что метод итераций Зейделя часто действительно приводит к более быстрой сходимости, чем метод простой итерации. Однако возможны случаи, когда метод итераций Зейделя сходится медленнее метода простой итерации, и даже случаи, когда метод простой итерации сходится, а метод итераций Зейделя расходится.

Отметим, что *метод Зейделя сходится*, если матрица A положительно определенная и симметрична.

Покажем, что метод итераций Зейделя эквивалентен некоторому методу простой итерации со специальным образом построенной матрицей α и вектором β в соотношении (1.10). Для этого запишем систему (1.14) в виде $x^{(k+1)} = Hx^{(k+1)} + Fx^{(k)} + \beta$, где F — верхняя треугольная матрица из коэффициентов матрицы α , а $H = \alpha - F$. Перепишем систему в виде $(E - H)x^{(k+1)} = Fx^{(k)} + \beta$, где E — единичная матрица. Матрица $(E - H)$ — нижняя треугольная матрица с диагональными элементами, равными единице. Следовательно, определитель этой матрицы отличен от нуля (равен единице) и она имеет обратную матрицу $(E - H)^{-1}$. Тогда

$$x^{(k+1)} = (E - H)^{-1} Fx^{(k)} + (E - H)^{-1} \beta.$$

Сопоставляя это соотношение с решением (1.10), можем заключить, что действительно метод итераций Зейделя эквивалент методу простой итерации в том смысле, что для установления условия и критерия сходимости метода итераций Зейделя можно воспользоваться теоремами, приведенными для метода простой

ЗЕЙДЕЛЬ ФИЛИПП ЛЮДВИГ (Seidel Philipp Ludwig; 1821—1896) — немецкий астроном и математик, основные работы которого относятся к математическому анализу. Одновременно и независимо от Дж. Г. Стокса З. ввел понятие равномерной сходимости последовательности и ряда. Им предложен итерационный метод решения системы линейных алгебраических уравнений (метод Зейделя).

итерации, если положить $\alpha = (E - H)^{-1}F$. Итерационный процесс для системы (1.12) записывают в более общей форме, а именно:

$$a_{ii}x_i^{(k+1)} = \omega b_i + (1 - \omega) a_{ii}x_i^{(k)} - \omega \left[\sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} + \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right],$$

вводя в итерационный процесс для $x_i^{(k+1)}$ значение $x_i^{(k)}$, которое отсутствует в методе простой итерации и методе Зейделя.

Итерационный процесс при $\omega > 1$ называют МЕТОДОМ ВЕРХНЕЙ РЕЛАКСАЦИИ, при $\omega = 1$ (метод Зейделя) — методом ПОЛНОЙ РЕЛАКСАЦИИ и при $\omega < 1$ — методом НИЖНЕЙ РЕЛАКСАЦИИ.

В простом случае при специально выбранном ω можно дать оценку числа итераций N , необходимых для достижения заданной точности ϵ . В методе простой итерации $N_{(\epsilon)} = 0,2n^2\ln(1/\epsilon)$, методе Зейделя $N = 0,1n^2\ln(1/\epsilon)$, методе верхней релаксации $N = 0,64n\ln(1/\epsilon)$. Очевидно, что число итераций тем больше, чем больше порядок системы; при этом имеет место квадратичная зависимость в первых двух методах и линейная зависимость в методе верхней релаксации. Очевидно также, что число итераций тем больше, чем меньше ϵ .

Представленные оценки не являются универсальными, и хотя в обсуждаемом случае самым медленным является метод простой итерации, а самым быстрым — метод верхней релаксации, это соотношение в других случаях может изменяться.

Можно дать наиболее общую запись итерационного процесса для системы (1.1). Имеем

$$B_k \frac{x^{(k+1)} - x^{(k)}}{\tau_k} + Ax^{(k)} = b, \quad (1.15)$$

где B_k — некоторая матрица, выбираемая для обеспечения быстрой сходимости метода. Во многом она определяется конкретной системой (1.1) и искусством вычислителя. Говорят, что итерационный процесс стационарный, если B_k и τ_k не зависят от k ; в дальнейшем индекс k при B и τ опускается. Из (1.15) очевидно, что если итерационный процесс сходится, т. е. $|x^{(k+1)} - x^{(k)}| \rightarrow 0$, то он сходится к решению системы (1.1).

Пусть матрицы D и L определены следующим образом:

$$D = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}, \quad L = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ a_{21} & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & 0 \end{bmatrix}.$$

Тогда метод простой итерации есть частный случай процесса (1.15) при $B = D$ и $\tau = 1$, метод Зейделя — частный случай при $B = D + L$, $\tau = 1$, а метод релаксации — частный случай при $B = D + \omega L$, $\tau = \omega$.

Матрицу B_k следует выбирать как можно ближе к матрице A , но так, чтобы обращение этой матрицы было более простой задачей (матрица B — треугольная, трехдиагональная и т. д.).

1.7. Решение нелинейных уравнений.

Два этапа отыскания корня

Пусть имеется нелинейное уравнение

$$f(x) = 0. \quad (1.16)$$

Требуется найти корни этого уравнения, т. е. те значения x , которые обращают уравнение (1.16) в тождество. В процессе приближенного отыскания корней уравнения (1.16) обычно выделяют два этапа: отделение корня и уточнение корня.

Под отделением корня понимается определение промежутка, содержащего один и только один корень уравнения. Одна из точек этого промежутка принимается за начальное приближение корня. В зависимости от метода, который предполагается использовать для уточнения корня, требуется определение тех или иных свойств отделенного корня и поведения функции на отрезке отделения. Например, при использовании простейшего метода уточнения корня — метода дихотомии, необходимо и достаточно установить лишь непрерывность функции на отрезке отделения. При использовании других методов может потребоваться выяснить, является ли корень действительным, какова кратность корня, установить непрерывность и монотонность функции и ее некоторых низших производных.

В общем случае этап отделения корня уравнения (1.16) не может быть алгоритмизирован. Для некоторых классов уравнений (наиболее известным из которых является класс алгебраических уравнений) разработаны специальные приемы отделения корней, существенно облегчающие такое отделение и позволяющие автоматизировать этот процесс. Некоторые из этих приемов будут приведены при рассмотрении методов решений алгебраических уравнений. Нередко отделение корней нелинейных уравнений выполняется «вручную» с использованием всей возможной

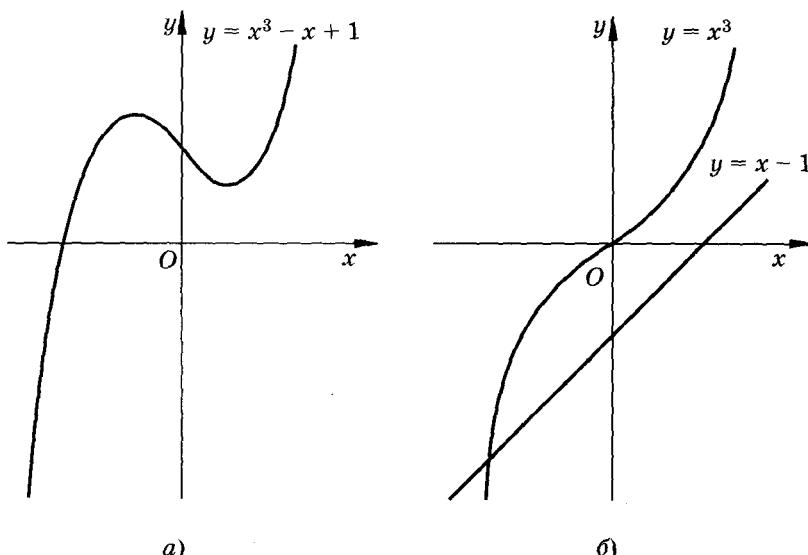


Рис. 1.2 (а, б)

информации о функции $f(x)$. В ряде случаев приближенное значение корня может быть определено из физических соображений, если речь идет о решении нелинейного уравнения, связанного с конкретной прикладной задачей. Успешно применяется графический метод определения действительных корней, обладающий большой наглядностью и позволяющий относительно просто устанавливать возможность существования кратных корней. При графическом отделении корней бывает полезным представить уравнение (1.16) в эквивалентном виде $f_1(x) = f_2(x)$ и искать точки пересечения функций $f_1(x)$ и $f_2(x)$. Например, для уравнения $f(x) \equiv x^3 - x + 1$ вместо построения графика $y = f(x)$ (рис. 1.2, а) проще построить графики функций $y = f_1(x) = x^3$ и $y = f_2(x) = x - 1$ (рис. 1.2, б).

В ряде случаев может быть полезной теорема, известная из курса математического анализа.

ТЕОРЕМА. Если непрерывная на $[a, b]$ функция $f(x)$, определяющая уравнение $f(x) = 0$, на концах отрезка $[a, b]$ принимает значения разных знаков, т. е. $f(a) \cdot f(b) < 0$, то на этом отрезке содержится по крайней мере один корень уравнения. Если же функция $f(x)$ непрерывна и дифференцируема и ее производная сохраняет знак внутри отрезка $[a, b]$, то на этом отрезке находится только один корень x^* уравнения.

В случае, когда на концах интервала функция имеет одинаковые знаки, на этом интервале корни либо отсутствуют, либо их четное число.

Известно, что интервал, на котором расположены корни x_p многочлена n -й степени $P_n(x) = \sum_{i=0}^n a_i x^i$, в том числе и комплексные, выражается соотношением

$$|x_p| \leq 1 + \frac{1}{|a_n|} \max [|a_0|, |a_1|, \dots, |a_{n-1}|].$$

Кроме того, по правилу знаков Декарта разность между числом перемен знаков последовательности a_0, a_1, \dots, a_n и числом положительных корней является либо положительным числом, либо нулем (в случае действительных корней). Это правило распространяется и на отрицательные корни при замене x на $-x$. Правило Декарта позволяет также оценить число действительных корней на интервале $[a, b]$. Для этого обозначим $x = (a + by)/(1 + y)$ и применим правило знаков к уравнению

$$P(y) = P_n\left(\frac{a + by}{1 + y}\right) = 0.$$

Для отделения корня полезно также использовать теорему Гюа.

ТЕОРЕМА ГЮА. Если все корни алгебраического уравнения являются действительными числами, то для последовательности коэффициентов a_0, a_1, \dots, a_n квадрат каждого некрайнего коэффициента больше произведения соседних с ним коэффициентов, т. е.

$$a_k^2 > a_{k-1} a_{k+1}, \quad k = 1, 2, \dots, n-1.$$

ТЕОРЕМА. Если для каких-либо k выполнено неравенство $a_k^2 \leq a_{k-1} a_{k+1}$, то многочлен имеет по крайней мере пару комплексных корней.

ДЕКАРТ РЕНЕ, латинизированное имя — КАРТЕЗИЙ (Descartes René, Cartesius; 1596—1650) — французский философ, математик, физик и физиолог. В труде «Геометрия» Д. сформулировал метод прямолинейных координат, заложив основы аналитической геометрии; одним из первых ввел понятия переменной величины и функции; дал классификацию кривых с подразделением их на алгебраические и трансцендентные. В своих трудах Д. много внимания уделил исследованию свойств уравнений и сформулировал правило знаков для определения числа положительных и отрицательных корней (правило Декарта).

На втором этапе уточнения при нахождении корня используют два типа методов: ПРЯМЫЕ и ИТЕРАЦИОННЫЕ. В прямых методах корень уравнения может быть найден за конечное, заранее известное число операций. Прямыми методами удается решить некоторые простейшие алгебраические и тригонометрические уравнения.

В итерационных методах корень x^* определяется как предел некоторой последовательности $x^{(0)}, x^{(1)}, \dots, x^{(k)}$ и решение не может быть достигнуто за конечное, заранее известное число операций.

Основные методы решения нелинейных уравнений и систем являются итерационными, и к их числу принадлежат метод дихотомии (половинного деления), метод простой итерации, метод Ньютона (метод касательных), метод секущих, метод парабол (метод Мюллера), метод Зейделя. Далее эти методы будут рассмотрены.

Важной характеристикой итерационных методов является скорость сходимости процесса. Говорят, что метод имеет n -й порядок сходимости, если $|x^{(k+1)} - x^*| = C |x^{(k)} - x^*|^n$, где C — постоянная, не зависящая от n . При $n = 1$ имеет место сходимость первого порядка, или *линейная сходимость*, а при $n = 2$ — второго порядка, или *квадратичная*. Говорят, что метод является *одношаговым*, если для построения итерационной последовательности нужно вычислить функцию в одной точке, *двушаговым* — в двух и т. п.

Сравнение различных методов следует проводить по числу операций при реализации одной итерации и по скорости сходимости.

Изложенные методы решения нелинейных уравнений и систем широко используются в численных методах оптимизации.

1.8. Метод половинного деления

Пусть действительный корень уравнения $f(x) = 0$ отделен и функция $f(x)$ непрерывна на интервале $[a, b]$ отделения корня. Построим процесс сужения интервала $[a, b]$ так, чтобы искомый

ДЕ ГЮА ДЕ МАЛЬВ ЖАН ПОЛЬ (De Gua de Malves Jean Paul; 1712—1786) — французский математик, аббат, основные труды которого относятся к алгебре и аналитической геометрии. В алгебре Г. занимался общей теорией уравнений, в геометрии — общей теорией кривых высшего порядка.

корень всегда находился внутри суженного интервала. Очевидно, что в этом случае погрешность приближенного значения корня не превышает $|b^{(k)} - a^{(k)}|$, где $b^{(k)}$, $a^{(k)}$ — граничные точки интервала на k -й итерации. Найдем середину отрезка $x^{(1)} = (a + b)/2$ и вычислим $f((a + b)/2)$. Составим произведения $f(a) \cdot f((a + b)/2)$ и $f(b) \cdot f((a + b)/2)$. Из двух половин отрезков выберем тот, в котором произведение является отрицательной величиной, и обозначим новые границы отрезка через $a^{(1)}$, $b^{(1)}$. Затем новый отрезок разделим пополам, вновь составим аналогичные произведения и выберем тот из отрезков, в котором произведение — величина отрицательная.

Погрешность метода половинного деления, который также называется МЕТОДОМ ДИХОТОМИИ*, определяется достаточно очевидным соотношением (которое, впрочем, может быть строго доказано)

$$\varepsilon^{(k)} = |x^* - x^{(k)}| \leq |b^{(k)} - a^{(k)}| \leq |b - a|/2^k,$$

которое указывает на скорость сходимости метода: с увеличением k погрешность стремится к нулю не медленнее геометрической прогрессии со знаменателем $q = 1/2$. Метод дихотомии прост и надежен, всегда сходится, хотя и медленно, устойчив к ошибкам округления. Метод дихотомии, однако, не обобщается на системы уравнений.

1.9. Метод простой итерации нахождения корней нелинейных уравнений

При использовании метода простой итерации для уточнения корня уравнение $f(x) = 0$ заменяется эквивалентным уравнением

$$x = \varphi(x). \quad (1.17)$$

Это означает, что из $f(x^*) = 0$ следует $x^* = \varphi(x^*)$ и наоборот. Привести уравнение (1.16) к уравнению (1.17) можно многими способами, например, положив $\varphi(x) = x + \psi(x)f(x)$, где $\psi(x)$ — непрерывная произвольная знакопостоянная функция.

* Дихотомия — от греческого слова διχοτομέω — разделяю на две части.

Геометрически на интервале отделения корня уравнение (1.17) представляется в виде двух пересекающихся линий $y = \phi(x)$ и $y = x$ (рис. 1.3). Полагая, что известно начальное приближение $x^{(0)}$ для значения корня x^* , построим итерационный процесс

$$\begin{aligned} x^{(k+1)} &= \phi(x^{(k)}), \\ k &= 0, 1, 2, \dots, \end{aligned} \quad (1.18)$$

изображенный на рис. 1.3 ломаной линией со стрелочками, указывающими направление движения. Для представленного на рис. 1.3 случая взаимного расположения линий $y = x$ и $y = \phi(x)$ неограниченное повторение вычислений по соотношению (1.18) позволяет сколь угодно близко подойти к точному значению корня x^* .

Исследуем сходимость метода. Если $\phi(x)$ имеет непрерывную производную, то из теоремы Лагранжа о конечном приращении

$$x^{(k+1)} - x^* = \phi(x^{(k)}) - \phi(x^*) \leq (x^{(k)} - x^*) |\phi'(x)| \quad (1.19)$$

следует, что точка ξ лежит между точками $x^{(k)}$ и x^* . Поэтому если всюду $|\phi'(x)| \leq q < 1$, то отрезки $|x^{(k)} - x^*|$ убывают не медленнее геометрической прогрессии со знаменателем $q < 1$. Действительно, из (1.19), которое можно рассматривать как рекуррентное соотношение, следует, что $|x^{(k)} - x^*| = q^k |x^{(0)} - x^*|$ и последовательность $x^{(k)}$ сходится при любом нулевом приближении.

Итак, условие

$$|\phi'(x)| < 1 \quad (1.20)$$

является достаточным условием сходимости итераций. Если $|\phi'(x)| > 1$, то итерации могут не сходиться. Если $|\phi'(x)| < 1$, но вдали от корня $|\phi'(x)| > 1$, то итерации сходятся, если начальное приближение выбрано достаточно близко к корню. При произвольном начальном приближении сходимости может не быть. Таким образом, в методе простой итерации важен выбор начального приближения. Из соотношения (1.19) следует, что если на интервале отделения корня выполняется условие

$$|\phi'(x)| \leq q < 1,$$

то погрешность на каждой итерации уменьшается для любого начального приближения не медленнее членов геометрической прогрессии со знаменателем q .

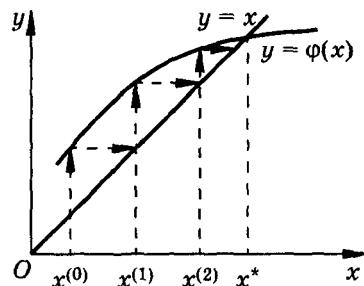


Рис. 1.3

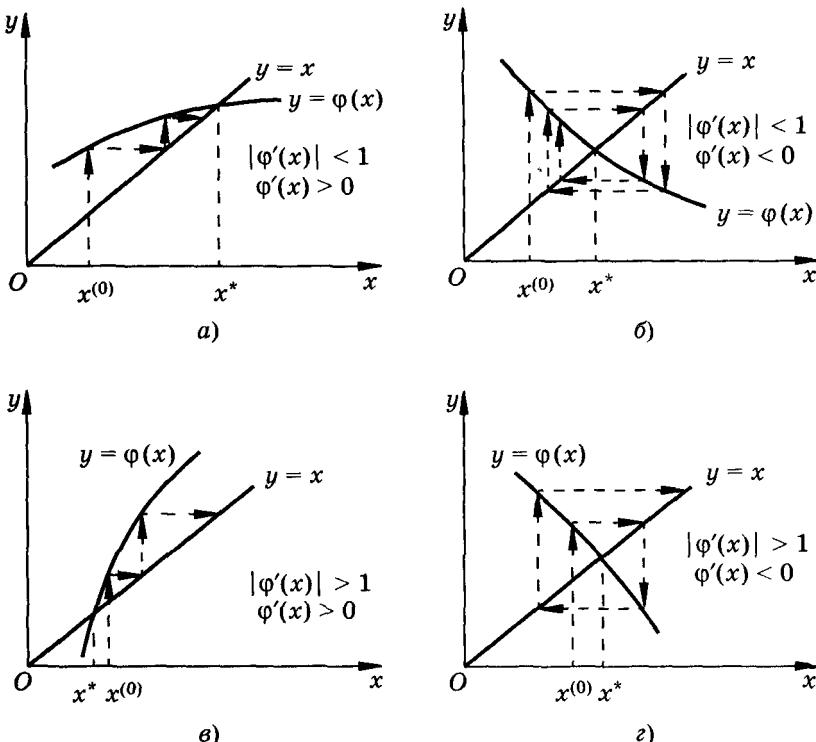


Рис. 1.4

Четыре случая взаимного расположения линий $y = x$ и $y = \varphi(x)$ вблизи корня и соответствующие им итерационные процессы показаны на рис. 1.4. Рис. 1.4, а и 1.4, б соответствуют случаю $|\varphi'(x)| < 1$ — процесс итераций сходится. При этом в первом случае $\varphi'(x) > 0$ и сходимость носит односторонний характер (рис. 1.4, а), а во втором $\varphi'(x) < 0$ и сходимость носит двусторонний характер (рис. 1.4, б). Рис. 1.4, в и 1.4, г соответствуют случаю $|\varphi'(x)| > 1$ — процесс итерации расходится, при этом имеет место односторонняя и двусторонняя расходимость.

Подчеркнем, что условие (1.20) сходимости метода итераций является лишь достаточным. При этом все приближения должны попадать в отрезок отделения корня. Выполнение условия (1.20) гарантирует сходимость процесса (1.18), но невыполнение условия (1.20) в общем случае не означает, что итерационный процесс окажется расходящимся. Например, для случая, проиллюстрированного рис. 1.5, условие (1.20) на интервале $[x, x^{(0)}]$ не выполняется, но метод итераций сходится.

Используя соотношения (1.19) и (1.20), можно записать

$$\begin{aligned} |x^* - x^{(k+1)}| &= \\ &= \varepsilon^{(k+1)} \leq q \varepsilon^{(k)} = q |x^* - x^{(k)}|, \end{aligned}$$

где $q = \max |\varphi'(x)|$, $x \in [a, b]$. Из этого соотношения следует, что скорость сходимости метода итерации зависит от величины q : чем меньше q , тем быстрее сходится метод.

Исходное уравнение $f(x) = 0$ может быть преобразовано к виду

$x = \varphi(x)$ многими способами, и, очевидно, для метода итерации целесообразно брать то уравнение $x = \varphi(x)$, для которого q имеет наименьшее значение.

Для пояснения рассмотрим классический пример вычисления квадратного корня. Исходное уравнение $f(x) = x^2 - a = 0$ ($a \geq 0$) преобразуем к виду $x = \varphi(x)$ тремя способами, приведенными в табл. 1.1 в первом столбце.

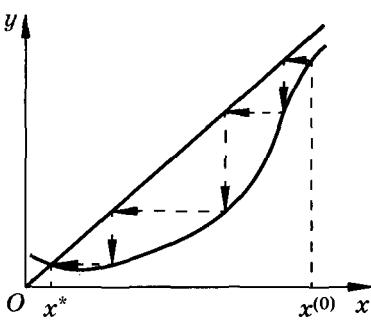


Рис. 1.5

Таблица 1.1

$\varphi(x)$	$\varphi'(x)$	Поведение $ \varphi'(x) $	Сходимость метода
$\frac{a}{x}$	$-\frac{a}{x^2}$	$ \varphi'(x) \rightarrow 1$ при $x \rightarrow \pm\sqrt{a}$	не сходится
$x^2 + x - a$	$2x + 1$	$ \varphi'(x) < 1$ при $x \in (-1, 0)$, $ \varphi'(x) \geq 1$ при $x \notin (-1, 0)$	сходится в ограниченном интервале к отрицательному значению корня
$\frac{(x + a/x)}{2}$	$\frac{(x - a/x^2)}{2}$	$ \varphi'(x) \rightarrow 0$ при $x \rightarrow \pm\sqrt{a}$	сходится, и очень быстро

Анализ поведения $|\varphi'(x)|$ вблизи корня (третий столбец таблицы) показывает, что при удачном выборе представления $x = \varphi(x)$ можно обеспечить высокую скорость сходимости итерационного процесса без ограничения диапазона параметра a . Третье уравнение $x = (x + a/x)/2$ используется для вычисления квадратного корня на компьютерах. Таким образом, в методе простой итерации важен выбор вида функций $\varphi(x)$. Отметим, что метод простой итерации обобщается на случай систем нелинейных уравнений.

1.10. Метод Ньютона и некоторые его модификации

Вновь рассмотрим уравнение (1.16). Полагая, что погрешность $\varepsilon^{(k)} = x^* - x^{(k)}$ мала, а функция $f(x)$ имеет непрерывную вторую производную, разложим $f(x^*)$ в ряд Тейлора:

$$f(x^*) = f(x^{(k)} + \varepsilon^{(k)}) = f(x^{(k)}) + \varepsilon^{(k)} f'(x^{(k)}) + \frac{(\varepsilon^{(k)})^2}{2} f''(\xi) + \dots,$$

где $\xi \in [x^{(k)}, x^*]$. Учитывая, что $f(x^*) = 0$ и оставляя только линейную часть разложения в ряд (отсюда и другое название метода — МЕТОД ЛИНЕАРИЗАЦИИ), можем записать приближенное, линейное относительно погрешности, уравнение

$$f(x^{(k)}) + \bar{\varepsilon}^{(k)} f'(x^{(k)}) = 0,$$

из которого для погрешности имеем

$$\bar{\varepsilon}^{(k)} = -f(x^{(k)})/f'(x^{(k)}). \quad (1.21)$$

Так как использована лишь линейная часть разложения в ряд, то при подстановке (1.21) в соотношение $x^* = x^{(k)} + \varepsilon^{(k)}$, следующее из соотношения для погрешности, получим вместо x^* лишь приближенное уточненное значение корня, которое обозначим $x^{(k+1)}$. Тогда можем записать основное соотношение метода Ньютона в виде

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})}. \quad (1.22)$$

НЬЮТОН ИСААК (Newton Isaac; 1643—1727) — великий английский физик, математик, механик и астроном. Н. сформулировал основные законы классической механики, открыл закон всемирного тяготения и разработал наряду с Г. Лейбницем дифференциальное и интегральное исчисления. Исследования Н. в области механики потребовали создания новых математических методов, основанных на введении переменных величин («флюэнт») и скоростей их изменения («флюксий»). Основную идею нового метода: по данной флюэнте найти флюксию и наоборот, т. е. по заданной функции найти производную и обратно — по производной найти функцию, Н. выразил в письме к Г. Лейбнику. Н. предложен метод численного решения алгебраических уравнений (метод Ньютона) и дано решение задачи о проведении через $n + 1$ данные точки параболической кривой n -го порядка (интерполяционная формула Ньютона).

Это соотношение позволяет построить последовательность приближений $x^{(k)}$, $k = 0, 1, 2, \dots$ к точному значению корня по заданному начальному приближению $x^{(0)}$.

Геометрически процесс (1.22) означает замену на каждой итерации кривой $y = f(x)$ на касательную к ней в точке $[x^{(k)}, f(x^{(k)})]$ и определение значения $x^{(k+1)}$ как координаты точки пересечения касательной и оси абсцисс (рис. 1.6). С рассмотренной интерпретацией соотношения (1.22) связано еще одно название метода — МЕТОД КАСАТЕЛЬНЫХ.

Достаточное условие сходимости метода Ньютона получим из соответствующего условия для метода простой итерации. Сопоставляя соотношения (1.17) и (1.22), можно заключить, что метод Ньютона представляет собой специальный случай метода простой итерации, в котором $\varphi(x) = x - f(x)/f'(x)$.

Используя условие сходимости метода итераций $|\varphi'(x)| < 1$ и выражение

$$\varphi'(x) = \frac{f(x)f''(x)}{[f'(x)]^2},$$

нетрудно получить достаточное условие сходимости метода Ньютона в форме

$$|f(x)f''(x)| < [f'(x)]^2. \quad (1.23)$$

Поскольку $f(x^*) = 0$, то $\varphi'(x^*) = 0$, и итерации по соотношению (1.22) сходятся к точному значению корня при произвольном начальном приближении, но вдали от корня сходимость может быть немонотонной.

Для оценки скорости сходимости метода Ньютона запишем соотношение

$$x^{(k+1)} - x^* = \varphi(x^{(k)}) - \varphi(x^*).$$

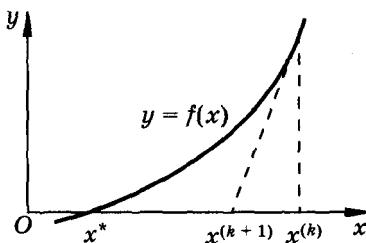


Рис. 1.6

ТЭЙЛОР БРУК (Taylor Brook; 1685—1731) — английский математик. В 1712 г. Т. вывел общую формулу для разложения функций в степенные ряды. Т. положил начало математическому изучению задачи о колебаниях струны. Ему принадлежат также заслуги в разработке теории конечных разностей.

Далее разложим $\varphi(x^{(k)})$ в ряд Тейлора:

$$\varphi(x^{(k)} - x^* + x^*) = \varphi(x^*) + \varphi'(x^*)(x^{(k)} - x^*) + \frac{1}{2} \varphi''(x^*)(x^{(k)} - x^*)^2 + \dots$$

Подставляя это разложение в предыдущую формулу и учитывая, что $\varphi'(x) = 0$, получаем

$$\varepsilon^{(k+1)} \approx (1/2) \varphi''(x) (\varepsilon^{(k)})^2.$$

Из этого соотношения следует, что метод Ньютона имеет вблизи корня второй порядок сходимости: на каждой итерации ошибка меняется пропорционально квадрату ошибки на предыдущей итерации. Достоинства метода Ньютона состоят в его квадратичной сходимости, возможности обобщения на случай систем уравнений, а также в том, что он является одношаговым. Однако метод Ньютона расходится в тех областях, где $f'(x) \equiv 0$. Кроме того, если функция $f(x)$ задана таблично, то вычисление $f'(x)$ затруднено.

Указанная трудность устраняется в МЕТОДЕ СЕКУЩИХ (методе хорд).

Существует вариант метода Ньютона с постоянным значением производной; значение производной вычисляется только в начальной точке $x = x^{(0)}$ (рис. 1.7), и далее для всех итераций значения производных $f'(x)$ полагаются постоянными, равными $f'(x^{(0)})$.

Последовательные приближения вычисляются по формуле

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(0)})}.$$

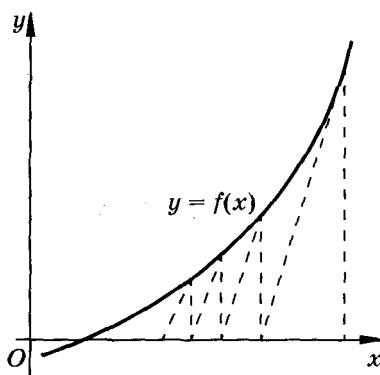


Рис. 1.7

Геометрически замена производной на каждой итерации постоянным значением означает, что наклон прямой, точка пересечения которой с осью абсцисс принимается за новое приближение для корня уравнения, считается постоянным. Часто эта модификация метода Ньютона рекомендуется как позволяющая уменьшить объем вычислений за счет отказа от вычисления производной на каждой итерации. Однако следует учитывать, что метод Ньютона с постоянным значением

производной имеет лишь первый порядок сходимости — вблизи корня соотношение для погрешности имеет вид

$$\varepsilon^{(k+1)} \approx \varepsilon^{(k)} [1 - f'(x^{(k)})/f'(x^{(0)})]$$

и, следовательно, сходится медленнее, чем метод Ньютона. Это, в конечном счете, нередко приводит к увеличению общего объема вычислений.

1.11. Метод секущих

В методе секущих, иначе называемом МЕТОДОМ ХОРД, приближенное значение производной $f'(x)$ в формуле (1.22) определяется по двум последовательным приближениям $[x^{(k-1)}, f(x^{(k-1)})]$ и $[x^{(k)}, f(x^{(k)})]$ по соотношению

$$f'(x^{(k)}) \approx \frac{f(x^{(k)}) - f(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}, \quad (1.24)$$

что приводит к замене касательной в точке $[x^{(k)}, f(x^{(k)})]$ секущей, проведенной через две точки кривой $y = f(x)$ (рис. 1.8) или, что то же самое, — к аппроксимации функции $f(x)$ на этом интервале линейной функцией.

Условия сходимости метода секущих аналогичны условиям сходимости метода Ньютона. Порядок сходимости метода секущих определяется соотношениями

$$\varepsilon^{(k+1)} \approx \alpha \varepsilon^{(k)} \varepsilon^{(k-1)} = \alpha^{1/p} [\varepsilon^{(k)}]^p,$$

где $\alpha = (1/2) f''(\xi)/f'(x^{(k)})$,

$$p = \sqrt{5}/2 \approx 1,618.$$

К особенностям метода следует отнести следующее: в методе не требуется непосредственного вычисления производной $f'(x)$ на каждой итерации, которое может привести к существенному уменьшению объема вычислений; метод является двухшаговым, и, в частности, на первой итерации вычислений необходимо знать два начальных значения $x^{(0)}$ и $x^{(1)}$;

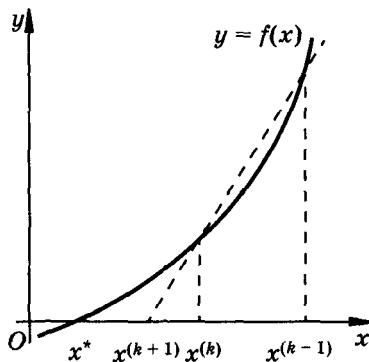


Рис. 1.8

сходимость метода может быть немонотонной даже в малой окрестности корня; в знаменателе формулы для вычисления $x^{(k+1)}$ стоит разность двух величин $[f(x^{(k)}) - f(x^{(k-1)})]$, которые имеют вблизи корня малые и близкие значения, что может привести к заметным погрешностям вычислений, особенно для кратных корней.

1.12. Метод парабол

Рассмотренный метод секущих можно интерпретировать как метод, в котором на каждой итерации исходная функция аппроксимируется линейной функцией (секущей), построенной по двум точкам, принадлежащим $f(x)$. Развивая далее идеи аппроксимации, можно для построения итерационных формул использовать информацию о функции в нескольких точках, предшествующих точке $x^{(k)}$. В методе парабол по трем последовательным приближениям

$$[x^{(k-2)}, f(x^{(k-2)})], [x^{(k-1)}, f(x^{(k-1)})], [x^{(k)}, f(x^{(k)})]$$

строится многочлен второй степени (парабола), приближающий исходную функцию. Иначе этот метод называют МЕТОДОМ МЮЛЛЕРА или методом КВАДРАТИЧНОЙ ИНТЕРПОЛЯЦИИ. За новое приближение берется обычно ближайший к $x^{(k)}$ корень соответствующего квадратного уравнения. Геометрическая интерпретация метода парабол дана на рис. 1.9.

В качестве $x^{(k+1)}$ выбирается тот из корней квадратного уравнения, для которого величина $|x^{(k+1)} - x^{(k)}|$ наименьшая. Доказывается, что погрешность метода определяется соотношением

$$\varepsilon^{(k+1)} \approx \varepsilon^{(k)} \varepsilon^{(k-1)} \varepsilon^{(k-2)} \approx (\varepsilon^{(k)})^p,$$

где $p = 1,839$.

РЕГИОМОНТАНУС МЮЛЛЕР ИОГАНН (Regiomontanus Müller Johann; 1436—1476) — немецкий астроном и математик. Р. Родился в г. Кёнигсберге и использовал латинизированное написание своего имени Кёнигсбергский (König — король, Berg — гора). Р. внес большой вклад в тригонометрию и астрономию: составил астрономические таблицы, предложил метод определения географических долгот на море (метод лунных расстояний), занимался плоской и сферической тригонометрией. В своих работах Р. использовал алгебраические символы. Ему принадлежат исследования в области решения уравнений, в частности, кубических уравнений.

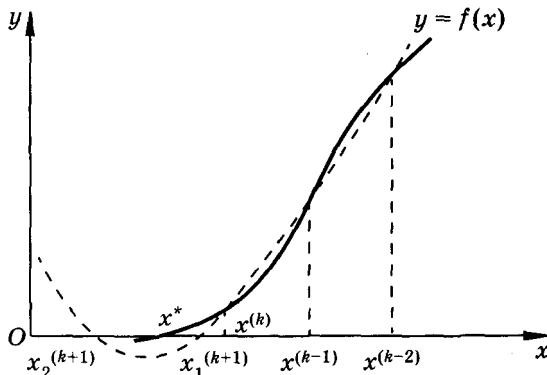


Рис. 1.9

Это означает, что, несмотря на привлечение дополнительной информации о функции, метод парабол имеет порядок сходимости, лишь немного превышающий порядок сходимости метода секущих. Вместе с тем возникают задачи решения квадратного уравнения, выбора одного из двух корней многочлена и, самое важное, определения области гарантированной сходимости метода. Если три приближения для построения многочлена выбраны далеко от корня и содержат погрешности, то возможно самое неожиданное поведение решения.

Отметим, что метод парабол успешно применяется для отыскания корней многочленов, в том числе комплексных; при этом метод обладает тем замечательным свойством, что начальное приближение может быть действительным. Метод парабол является трехшаговым методом.

1.13. Методы нахождения корней систем

нелинейных уравнений.

Ускорение сходимости по Эйткену

Для уточнения корней систем нелинейных уравнений наиболее часто используют методы итерации (метод простой итерации и метод Зейделя) и метод Ньютона. Как и в случае уточнения корней одного нелинейного уравнения, для систем нелинейных уравнений требуется определение хорошего начального приближения (отделение корня), гарантирующего сходимость метода и высокую скорость сходимости. Для системы двух уравнений это

может быть сделано графически, но для систем высоких порядков доведетворительных методов отделения корней не существует.

Метод простой итерации

Систему нелинейных уравнений запишем в векторной форме

$$f(x) = 0, \quad (1.25)$$

где $x = (x_1, x_2, \dots, x_n)^T$ — вектор-столбец неизвестных*, $f(x) = [f_1(x), f_2(x), \dots, f_n(x)]^T$ — вектор-столбец функций. В методе простой итерации система (1.25) приводится к эквивалентной системе вида $x = \varphi(x)$, где $\varphi(x) = [\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)]^T$, или

$$\begin{cases} x_1 = \varphi_1(x_1, x_2, \dots, x_n), \\ x_2 = \varphi_2(x_1, x_2, \dots, x_n), \\ \dots \dots \dots \dots \\ x_n = \varphi_n(x_1, x_2, \dots, x_n). \end{cases} \quad (1.26)$$

Полагая известным начальное приближение для корня $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})^T$, построим итерационный процесс $x^{(k+1)} = \varphi(x^{(k)})$, $k = 0, 1, 2, \dots$, или

$$\begin{cases} x_1^{(k+1)} = \varphi_1[x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}], \\ x_2^{(k+1)} = \varphi_2[x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}], \\ \dots \dots \dots \dots \\ x_n^{(k+1)} = \varphi_n[x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}]. \end{cases} \quad (1.27)$$

Рассмотрим поведение вектора погрешности

$$\varepsilon^{(k)} = x^{(k)} - x^* = [x_1^{(k)} - x_1^*, \dots, x_n^{(k)} - x_n^*],$$

полагая, что погрешность — величина малая. Для компонент вектора x можем записать

$$x_i^{(k+1)} = x_i^* - \varepsilon_i^{(k+1)} = \varphi_i[x_1^* - \varepsilon_1^{(k)}, \dots, x_n^* - \varepsilon_n^{(k)}], \quad i = 1, \dots, n.$$

Полагая наличие у функций $\varphi_i(x)$ непрерывных частных производных и используя соотношение $x_i^* = \varphi_i(x^*)$, можем (см. (1.19)), используя разложение в ряд, получить

$$\varepsilon_i^{(k+1)} \approx \sum_{l=1}^n \varepsilon_l^{(k)} \frac{\partial \varphi_i(x)}{\partial x_l} \Bigg|_{x=x^*}, \quad i = 1, \dots, n. \quad (1.28)$$

*Верхний индекс T означает операцию транспонирования.

Из (1.28) следует, что в методе простой итерации вектор погрешности испытывает линейное преобразование, или, иначе, метод имеет первый порядок сходимости.

Если обозначить матрицу производных системы функций $\varphi_i(x)$ для $i = 1, \dots, n$ (МАТРИЦУ ЯКОБИ) через

$$A_\varphi = \begin{bmatrix} \frac{\partial \varphi_1(x)}{\partial x_1} & \frac{\partial \varphi_1(x)}{\partial x_2} & \dots & \frac{\partial \varphi_1(x)}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial \varphi_n(x)}{\partial x_1} & \frac{\partial \varphi_n(x)}{\partial x_2} & \dots & \frac{\partial \varphi_n(x)}{\partial x_n} \end{bmatrix}_{x=x^*},$$

то систему (1.28) можно переписать в виде

$$\varepsilon^{k+1} = A_\varphi \varepsilon^k. \quad (1.29)$$

Достаточное условие сходимости итерационного процесса (1.27) формулируется следующим образом: если какая-либо норма матрицы A_φ , согласованная с рассматриваемой нормой вектора x , меньше единицы, то метод итераций сходится. Условие сходимости $\|A_\varphi\| < 1$ есть обобщение на случай нелинейной системы условия (1.20) для одного уравнения.

| Метод итераций Зейделя

Нередко сходимость метода простой итерации можно улучшить, если вновь вычисленные значения компонент вектора неизвестных немедленно включить в расчет. В этом случае итерационный процесс имеет вид

$$x_i^{(k+1)} = \varphi_i[x_1^{(k+1)}, \dots, x_{i-1}^{(k+1)}, x_i^{(k)}, \dots, x_n^{(k)}], \quad i = 1, \dots, n.$$

Сходимость этого процесса также линейная. Как и при решении систем линейных уравнений, может быть поставлена задача об отыскании оптимальной на каждой итерации последовательности уточнения компонент вектора решения. Удовлетворитель-

ЯКОБИ КАРЛ ГУСТАВ ЯКОБ (Jacobi Carl Gustav Jacob; 1804—1851) — немецкий математик, один из создателей теории эллиптических функций. Я. принадлежат открытия в области теории чисел, алгебры, вариационного исчисления, интегрального исчисления и теории дифференциальных уравнений. Он ввел в употребление функциональные определители (якобианы) и указал на их роль при замене переменных в кратных интегралах и при решении уравнений с частными производными. Я. исследовал класс ортогональных многочленов (многочлены Якоби).

ных методов построения оптимальной последовательности нет. На практике иногда используется упорядочение неизвестных по убыванию разности их значений на двух последовательных итерациях.

Метод Ньютона

Основная идея метода Ньютона — решение системы нелинейных уравнений $f(x) = 0$ сводится к решению последовательности линейных задач, дающих в пределе решение исходной задачи. Линейная задача получается путем выделения из нелинейных уравнений главной линейной части.

Рассмотрим погрешность вычисления корня на k -й итерации $\varepsilon^{(k)} = x^* - x^{(k)}$, где $\varepsilon^{(k)} = [\varepsilon_1^{(k)}, \varepsilon_2^{(k)}, \dots, \varepsilon_n^{(k)}]'$. Полагая, что функции f_i непрерывны и дифференцируемы в окрестности корня и $\varepsilon_i^{(k)} (i = 1, \dots, n)$ — малые величины, разложим $f(x^*) = f(x^{(k)} + \varepsilon^{(k)}) = 0$ в ряд Тейлора, сохранив лишь линейную часть разложения. Получим систему уравнений

$$\sum_{j=1}^n \varepsilon_j^{(k)} \frac{\partial f_i(x)}{\partial x_j} \Bigg|_{x=x^*} = -f_i(x^{(k)}), \quad i = 1, \dots, n, \quad (1.30)$$

линейную относительно компонент вектора погрешностей. Если использовать эту систему для отыскания компонент вектора погрешностей, то в силу приближенности системы (1.30) — оставлена лишь линейная часть — найденное значение вектора погрешности будет лишь приближенным. Тогда при подстановке полученного решения в соотношение $x^* = \varepsilon^{(k)} - x^{(k)}$ будем иметь вместо x^* приближенное уточненное значение корня, которое обозначим через $x^{(k+1)}$. Используя запись системы (1.30) в виде

$$A_f \varepsilon^{(k)} = -f^{(k)},$$

где A_f — матрица производных системы функций f_i (матрица Якоби), можем записать итерационный процесс для нахождения вектора x :

$$x^{(k+1)} = x^{(k)} - [A_f^{(k)}]^{-1} f^{(k)}, \quad k = 0, 1, 2, \dots,$$

где $[A^{(k)}]^{-1}$ — матрица, обратная матрице Якоби. Представленная формула является обобщением формулы (1.22) на случай систем нелинейных уравнений. Уточненное значение вектора может быть вновь использовано для получения следующего приближения к корню, что приводит к итерационному процессу. Заметим, что в большинстве случаев предпочтительным является не вычисление обратной матрицы $[A_f^{(k)}]^{-1}$, а получение каким-либо методом

решения $\varepsilon^{(k)}$ линейной системы (1.30) и вычисление нового приближенного значения по соотношению $x^{(k+1)} = x^{(k)} + \varepsilon^{(k)}$.

Итерационный процесс (1.30) сходится, если определитель матрицы $A_f^{(k)}$ отличен от нуля, т. е. $\det(A_f^{(k)}) \neq 0$. Требуется, однако, хорошее отделение корня, но достаточное условие сходимости метода слишком громоздко, чтобы им можно было воспользоваться на практике.

На каждой итерации метода Ньютона требуется вычислять матрицу производных $A_f^{(k)}$ и решать систему линейных уравнений (1.30). Можно попытаться уменьшить объем вычислений за счет отказа от вычисления матрицы $A_f^{(k)}$ на каждой итерации и использования на всех итерациях постоянного значения $A_f^{(0)}$, вычисленного по начальному приближению. Напомним, что при этом может быть дополнительно существенно уменьшен объем вычислений, если для решения последовательности линейных систем использовать алгоритм, позволяющий выполнить преобразование матрицы $A_f^{(0)}$ к верхней треугольной только один раз. Следует иметь в виду, однако, что, во-первых, указанная модификация метода Ньютона гарантирует лишь линейную сходимость итераций (против квадратичной в окрестности корня в методе Ньютона) и, во-вторых, константа в линейной зависимости погрешности при неудачном выборе начального приближения может оказаться весьма большой и сходимость будет медленной. Таким образом увеличивается число итераций, необходимое для достижения заданной точности, и уменьшение общего объема вычислений не гарантировано.

Ускорение сходимости по Эйткену

Предположим, что отношение $(x^{(k-1)} - x^*)/(x^{(k)} - x^*) = q$ есть величина постоянная и неизменная в процессе итераций. Тогда

$$\frac{x^{(k-1)} - x^*}{x^{(k)} - x^*} = q = \frac{x^{(k-2)} - x^*}{x^{(k-1)} - x^*}.$$

Из этого соотношения следует, что

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

ЭЙТКЕН АЛЕКСАНДР КРЭГ (Aitken Alexander Craig; 1895—1967) — английский математик. Основные его труды относятся к численному анализу, статистике и алгебре. В алгебре Э. проводил исследования по теории определителей. В области численного анализа ему принадлежит идея об ускорении сходимости численного метода. Э. также разработал метод последовательной линейной интерполяции.

Полученный таким образом корень x^* можно принять за следующее приближенное значение $x^{(k+1)}$. Предложенный способ пригоден как для одного нелинейного уравнения, так и для систем нелинейных уравнений. Это предположение означает, что метод применим к процессам с линейной сходимостью (простые итерации), но неприменим к методам Ньютона, секущих, парабол и т. п.

1.14. Введение в проблему собственных значений

Во многих задачах механики, физики, химии и связанных с ними задачах алгебры и вычислительной математики вместе с квадратной матрицей A порядка n приходится рассматривать уравнение

$$\det(A - \lambda E) = 0, \quad (1.31)$$

которое называют ХАРАКТЕРИСТИЧЕСКИМ или ВЕКОВЫМ УРАВНЕНИЕМ матрицы A , и однородную систему

$$(A - \lambda E)x = 0, \quad (1.32)$$

где E — единичная матрица порядка n , x — вектор-столбец.

Уравнение (1.31) есть алгебраическое уравнение (многочлен) степени n вида

$$(-1)^n (\lambda^n - p_1 \lambda^{n-1} - p_2 \lambda^{n-2} - \dots - p_n) = (-1)^n P_n(\lambda). \quad (1.33)$$

Корни этого многочлена λ_i , $i = 1, \dots, n$ называются собственными (характеристическими) значениями (числами) матрицы A . Однородная система (1.32) имеет ненулевое решение лишь в случае, при котором корни λ_i являются собственными значениями матрицы A . Соответствующие каждому λ_i ненулевые решения x системы (1.32) называются собственными векторами матрицы A .

С точки зрения линейной алгебры ЗАДАЧА О СОБСТВЕННЫХ ЗНАЧЕНИЯХ и векторах включает в себя составление характеристического многочлена (1.33), отыскание корней этого многочлена (собственных значений), решение для каждого λ_i однородной системы (1.32), т. е. определение собственных векторов матрицы A .

Говорят о полной проблеме собственных значений, если требуется определить все собственные значения (и собственные векторы) матрицы A , и о частичных проблемах собственных значений, если требуется найти лишь отдельные собственные значения, например, максимальное или минимальное. Обобщенная проблема собственных значений возникает при решении системы

$Ax = \lambda Bx$, а общая — тогда, когда все коэффициенты матрицы A зависят от λ .

Существуют три группы методов решения второй проблемы линейной алгебры.

- В первой группе методов (методы отражения, прямой вращения, итерационный вращения и QL -алгоритм) матрица приводится к какому-либо специальному виду, например диагональному или трехдиагональному, когда вычисление собственных чисел — простая задача.

- Во второй группе (методы Данилевского, Леверье, Крылова, Ланцша) матрицы приводятся к такой форме, для которой вычисление коэффициентов характеристического многочлена было бы тривиальным, т. е. производилось бы восстановление характеристического многочлена.

- В третьей группе — методов интерполяции — собственные числа определяются с использованием интерполяционных многочленов.

ЛÁППО-ДАНИЛÉВСКИЙ ИВАН АЛЕКСАНДРОВИЧ (1896—1931) — математик, основные труды которого относятся к теории функций от матриц и аналитической теории линейных дифференциальных уравнений.

ЛЕВЕРЬÉ УРБЕН ЖАН ЖОЗЕФ (Le Verrier Urbain Jean Joseph; 1811—1877) — французский астроном, работы которого посвящены проблемам небесной механики. Изучая неправильности в движении Урана, Л. показал, что причиной, их вызывающей, является находящаяся за пределами его орбиты неизвестная планета. Независимо от Дж. Адамса Л. вычислил (1846) положение этой планеты (Нептуна). Составленные им планетные таблицы потребовали колоссального вычислительного труда.

КРЫЛОВ АЛЕКСЕЙ НИКОЛАЕВИЧ (1863—1945) — кораблестроитель, механик и математик. Работы К. по кораблестроению имеют большую ценность для математики и механики. К. занимался вопросами наиболее рациональной организации численных расчетов, разработал способ улучшения сходимости тригонометрических рядов, предложил метод решения векового уравнения. К. построил первую в России машину для интегрирования дифференциальных уравнений.

ЛÁНЦШОУ КОРНЕЛИУС (Lanczos Cornelius; 1893—1974) — американский математик венгерского происхождения. Л. занимался в основном математическим анализом и приближенными вычислениями. Известны ленточные алгоритмы Ланцша, а также блочный метод Ланцша в алгебре.

1.15. Итерационный метод вращения для нахождения собственных значений

Метод предложен в 1846 г. Якоби и позволяет решить полную проблему собственных значений для эрмитовых матриц. Для простоты рассмотрим метод вращений для частного случая эрмитовых матриц — действительных симметрических матриц. В основе метода лежит тот факт, что действительная симметрическая матрица может быть приведена подобным преобразованием к диагональному виду. Метод Якоби основан на использовании нескольких оригинальных идей.

- Используется преобразование подобия $S^{-1}AS$, сохраняющее собственные числа.
- При этом преобразовании используются ортогональные матрицы, приводящие матрицу A к диагональной матрице Λ . Отметим, что *a priori* соответствующая ортогональная матрица неизвестна и цель метода — в конечном итоге ее получить.
- Согласно соотношению $\Lambda = U^{-1}AU$, в качестве ортогональной матрицы в каждой итерации используется матрица вращения.
- Итерационный процесс строится таким образом, чтобы после каждой итерации происходило уменьшение суммы квадратов внедиагональных элементов.
- Реализация такого итерационного процесса на каждом этапе обращает в нуль максимальный элемент матрицы и определяет угол ϕ в матрице вращения.

Матрицей элементарного вращения называется матрица вида

$$U_{ij} = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 & 0 \\ 0 & 1 & \dots & \dots & \dots & \dots & \dots & 0 & 0 \\ \dots & \dots \\ 0 & 0 & \dots & \cos \phi & \dots & -\sin \phi & \dots & 0 & 0 \\ \dots & \dots \\ 0 & 0 & \dots & \dots & 1 & \dots & \dots & 0 & 0 \\ \dots & \dots \\ 0 & 0 & \dots & \sin \phi & \dots & -\cos \phi & \dots & 0 & 0 \\ \dots & \dots \\ 0 & 0 & \dots & \dots & \dots & \dots & \dots & 0 & 1 \end{bmatrix} \leftarrow \begin{array}{l} i \\ j \end{array} .$$

Здесь отмечены i -е и j -е строки и столбцы матрицы; она является ортогональной.

Название МАТРИЦА ВРАЩЕНИЯ связано с тем, что в пространстве (плоскости) \mathbb{R}^2 матрица является матрицей преобразования декартовых координат при повороте осей на угол φ . Непосредственным умножением можно убедиться, что $U_{ij}^{-1} = U_{ij}^T$, ибо, если матрица $U_{ij}(\varphi)$ является ортогональной, обратная матрица равна транспонированной. Столь простая возможность получения обратной матрицы, необходимой для выполнения подобного преобразования, весьма ценна с точки зрения эффективности метода. В итерационном методе вращения строится бесконечная последовательность подобных преобразований с помощью матриц вращения — такая, что в пределе исходная матрица преобразуется к диагональному виду. Рассмотрим алгоритм метода и покажем, что он действительно обеспечивает требуемое преобразование.

Предположим, что на некоторой k -й итерации построена матрица $A^{(k)}$. Выбираем в ней максимальный по модулю внедиагональный элемент $a_{ij}^{(k)}$. Так как матрица симметрична, то без ущерба для общности можно считать, что $i < j$. По индексам i и j строим матрицу вращения $U_{ij}^{(k)}$ и выполняем подобное преобразование над матрицей $A^{(k)}$. В результате получаем матрицу $A^{(k+1)} = U_{ij}^{(k)T} A^{(k)} U_{ij}^{(k)}$, отличающуюся от матрицы $A^{(k)}$ только элементами i -х и j -х строк и столбцов.

Выпишем соотношения для вычисления элементов i -х и j -х строк и столбцов, введя для упрощения промежуточную матрицу $B^{(k)} = A^{(k)} U_{ij}^{(k)}$. Учитывая вид матрицы $U_{ij}(\varphi)$, отметим, что все столбцы матрицы $B^{(k)}$, кроме i -х и j -х, будут такими же, как и у матрицы $A^{(k)}$. Для элементов i -го и j -го столбцов этой матрицы

$$\begin{aligned} b_{mi}^{(k)} &= a_{mi}^{(k)} \cos \varphi^{(k)} + a_{mj}^{(k)} \sin \varphi^{(k)}, \\ b_{mj}^{(k)} &= -a_{mi}^{(k)} \sin \varphi^{(k)} + a_{mj}^{(k)} \cos \varphi^{(k)}, \quad m = 1, 2, \dots, n. \end{aligned} \tag{1.34}$$

Заметим, что для всех $m \neq i$ и $m \neq j$, т. е. для внедиагональных элементов, можем записать $b_{mi}^{(k)} = a_{mi}^{(k+1)}$, так как последующее умножение матрицы B слева на U_{ij}^T не изменяет внедиагональных элементов i -го и j -го столбцов.

Для элементов i -й и j -й строк матрицы $A^{(k+1)}$

$$\begin{aligned} a_{im}^{(k+1)} &= b_{im}^{(k)} \cos \varphi^{(k)} + b_{jm}^{(k)} \sin \varphi^{(k)}, \\ a_{jm}^{(k+1)} &= -b_{im}^{(k)} \sin \varphi^{(k)} + b_{jm}^{(k)} \cos \varphi^{(k)}, \quad m = 1, 2, \dots, n. \end{aligned} \tag{1.35}$$

Матрицу элементарного вращения $U_{ij}^{(k)}$ выбираем таким образом, чтобы обратить максимальный по модулю внедиагональный элемент в нуль. Для этого, используя соотношения (1.34) и (1.35), записываем выражение для $a_{ij}^{(k+1)}$:

$$a_{ij}^{(k+1)} = \frac{1}{2} (a_{jj}^{(k)} - a_{ii}^{(k)}) \sin 2\varphi^{(k)} + \cos 2\varphi^{(k)}.$$

Приравнивая $a_{ij}^{(k+1)}$ к нулю, получаем выражение для определения угла $\varphi^{(k)}$, обеспечивающего при подобном преобразовании с помощью матрицы вращения приравнивание к нулю элемента матрицы $A^{(k+1)}$ с индексами i и j и симметричного ему элемента. Имеем

$$\operatorname{tg} 2\varphi^{(k)} = \frac{2a_{ij}^{(k)}}{(a_{ii}^{(k)} - a_{jj}^{(k)})}. \quad (1.36)$$

Таким образом, подобное преобразование при выборе матрицы вращения с соотношением (1.36) обращает в нуль элемент с индексами i и j и симметричный ему элемент матрицы $A^{(k+1)}$. При этом, однако, в соответствии с соотношениями (1.34) и (1.35) происходит изменение всех остальных элементов i -х и j -х строк и столбцов (и, в частности, тех, которые, может быть, на предыдущих итерациях были обращены в нуль). Можно показать, что в построенном процессе все внедиагональные элементы матрицы при достаточном числе подобных преобразований (итераций) становятся сколь угодно малыми. Имеет место соотношение

$$t[A^{(k+1)}] = t[A^{(k)}] - 2[a_{ij}^{(k)}]^2$$

или

$$t[A^{(k+1)}] < t[A^{(k)}].$$

Можно также показать, что $t[A^{(k)}] \leq q^k t(A)$, где значение q подчиняется неравенству $q < 1 - 2/(n(n+1))$, т. е. сходимость итерационного метода вращения ухудшается с увеличением n .

Собственные векторы являются столбцами матрицы $U = \prod_{k=0}^K U_{ij}^{(k)}$. Здесь через K обозначено число вращений. В методе вращения для вычисления собственных значений требуется $50n^3$ операций. В качестве хорошего нулевого приближения к методу Гаусса служит метод вращения Гивенса, применимый к неэрмитовым матрицам.

1.16. Метод Данилевского для построения характеристического многочлена матрицы

В методе Данилевского исходная матрица A с помощью подобных преобразований $\Phi = S^{-1}AS$ приводится к так называемой канонической форме Фробениуса

$$\Phi = \begin{bmatrix} p_1 & p_2 & p_3 & \dots & p_{n-1} & p_n \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}. \quad (1.37)$$

Разложив определитель матрицы $(\Phi - \lambda E)$ по элементам первой строки, получим

$$|\Phi - \lambda E| = (-1)^n (\lambda^n - p_1\lambda^{n-1} - p_2\lambda^{n-2} - \dots - p_n) = (-1)^n P_n(\lambda),$$

т. е. элементы первой строки матрицы (1.37) являются коэффициентами характеристического многочлена матрицы Φ , и так как матрица Φ получена подобными преобразованиями матрицы A , то они являются также и коэффициентами характеристического многочлена исходной матрицы A .

Переход от матрицы A к матрице Φ осуществляется за $(n-1)$ преобразование подобия, последовательно преобразующее строки матрицы A , начиная с последней, в строки матрицы Φ . Рассмотрим схему преобразований. Пусть после некоторого шага матрица A приведена к виду

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1(k-1)} & b_{1k} & \dots & b_{1(n-1)} & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2(k-1)} & b_{2k} & \dots & b_{2(n-1)} & b_{2n} \\ \dots & \dots \\ b_{k1} & b_{k2} & \dots & b_{k(k-1)} & b_{kk} & \dots & b_{k(n-1)} & b_{kn} \\ 0 & 0 & \dots & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 & 0 \end{bmatrix}$$

(здесь и далее индекс шага для сокращения опущен).

Приведение k -й строки к строке матрицы Фробениуса осуществляется следующим образом. Весь $(k - 1)$ -й столбец делится на $b_{k(k-1)}$, так что вместо элемента $b_{k(k-1)}$ получаем 1. Далее этот столбец умножается на элемент $-b_{ki}$ и складывается с i -м столбцом, при этом i изменяется от 1 до n (кроме, очевидно, $i = (k - 1)$). В результате k -я строка принимает форму Фробениуса. Нетрудно показать, что описанная процедура эквивалентна умножению матрицы B справа на матрицу S , имеющую вид

$$S = \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ -\frac{b_{k1}}{b_{k(k-1)}} & -\frac{b_{k2}}{b_{k(k-1)}} & \dots & \frac{1}{b_{k(k-1)}} & \dots & -\frac{b_{kn}}{b_{k(k-1)}} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 1 \end{bmatrix}.$$

Непосредственным умножением легко убедиться, что матрица S^{-1} , обратная матрице S , имеет вид

$$S^{-1} = \begin{bmatrix} 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ b_{k1} & b_{k2} & \dots & b_{kk} & \dots & b_{kn} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 1 \end{bmatrix}.$$

При умножении матрицы B на матрицу S справа k -я строка матрицы B принимает требуемый вид, а строки матрицы Фробениуса с $(k - 1)$ -й по n -ю сохраняют свой вид.

Однако матрица BS не подобна матрице A , и, чтобы сделать ее подобной, нужно матрицу BS умножить слева на матрицу S^{-1} .

Умножение матрицы BS слева на матрицу S^{-1} не меняет строк с k -й по n -ю этой матрицы. Применив последовательно описанный процесс к строкам матрицы A , начиная с последней, получим матрицу Фробениуса (1.37).

ФРОБЕНИУС ФЕРДИНАНД ГЕОРГ (Frobenius Ferdinand Georg; 1849—1917) — немецкий математик. Его основные работы относятся к алгебре, в частности, теории матриц, теории конечных групп и их представлений матрицами.

1.17. Метод интерполяции для построения характеристического многочлена

В основе метода интерполяции для построения характеристического многочлена лежит известная теорема алгебры о единственности интерполяционного многочлена. Задача об определении коэффициентов p_i ($i = 1, \dots, n$) характеристического многочлена степени n сводится к задаче построения интерполяционного многочлена степени n по известным значениям его в узловых точках. Коэффициент при неизвестном в степени n в уравнении (1.33) равен $(-1)^n$. Для этого задается n различных значений $\lambda = \beta_j$, $j = 1, 2, \dots, n$, и для каждого из этих значений вычисляется значение определителя $\det(A - \beta_j E) = \alpha_j$. Значения β_j могут выбираться произвольно, но для уменьшения влияния ошибок округления желательно узлы интерполяции располагать равномерно внутри интервала $(\lambda_{\min}, \lambda_{\max})$. Для оценки границ интервала можно использовать соотношение $\max |\lambda| < \|A\|$, где $\|A\|$ — любая норма матрицы A .

Для определения p_i ($i = 1, \dots, n$) записывается система n линейных уравнений

$$\beta_j^n - \sum_{i=1}^n p_i \beta_j^{n-i} = \alpha_j, \quad j = 1, \dots, n.$$

Если выбранные значения β_j различны, то определитель этой системы отличен от нуля и система уравнений имеет единственное решение. В силу теоремы о единственности интерполяционного многочлена построенный таким образом многочлен является характеристическим многочленом матрицы A . Отметим, что обычно для построения многочлена используются интерполяционные формулы, которые будут рассмотрены далее.

Одним из существенных достоинств метода интерполяции является его общность: метод может быть использован при построении характеристического многочлена для матриц произвольной структуры. Вместе с тем общность метода является и его недостатком, так как не все особенности матриц, образующихся в практических задачах, могут быть учтены.

Для применения метода интерполяции требуется вычислить n определителей порядка n . Вместе с построением интерполяционного многочлена по интерполяционной формуле метод требует примерно n^4 операций, что приводит для больших значений n к довольно большому объему вычислений. Однако более серьезным

обстоятельством, ограничивающим применение метода интерполяции, является то, что нули многочлена (т. е. искомые собственные значения) могут быть весьма чувствительны к погрешностям коэффициентов многочлена. Неизбежные и, возможно, достаточно большие погрешности вычисления коэффициентов P_n в методе интерполяции приводят к существенному снижению точности определения собственных значений при больших n . По этой причине метод интерполяции редко используется при $n > 10$.

Последний недостаток метода интерполяции можно устранить, если использовать метод ЛОКАЛЬНОЙ ИНТЕРПОЛЯЦИИ. В нем восстанавливается не весь характеристический многочлен, а некоторая его часть вблизи корня. Вычислив $\det(A - \lambda E)$ в трех точках, можно построить в окрестности корня многочлен второй степени и далее, решая квадратное уравнение, определить корни. Для нахождения корня используется метод парабол. Метод локальной интерполяции обладает тем преимуществом, что избавляет от необходимости строить многочлен высших степеней. Он пригоден для матриц любых видов и может быть использован для решения частичной, полной, обобщенной и общей проблем.

В методах интерполяции, таким образом, решение второй проблемы достигается путем многократного решения первой проблемы линейной алгебры, имея в виду вычисление определителей.

1.18. Численные методы оптимизации

ОПТИМИЗАЦИЯ — процесс НАХОЖДЕНИЯ ЭКСТРЕМУМА (глобального максимума или минимума) некоторой ФУНКЦИИ или выбор наилучшего (оптимального) варианта из множества возможных. Приведем математическую формулировку задачи.

Необходимо найти минимум скалярной функции $f(\bar{x})$ векторного аргумента x на множестве

$$\mathbf{X} = \{x: g_i(\bar{x}) \geq 0; h_i(\bar{x}) = 0; i = 1, \dots, N\},$$

где g_i и h_i также скалярные функции*.

*Здесь и далее будем говорить о поиске минимума функции $f(\bar{x})$, так как поиск максимума $f(\bar{x})$ сводится к поиску минимума функции $F(x) = -f(x)$.

Функция $f(\bar{x})$ называется ЦЕЛЕВОЙ ФУНКЦИЕЙ; множество X — ДОПУСТИМЫМ МНОЖЕСТВОМ, а решение \bar{x}^* задачи оптимизации — ОПТИМАЛЬНОЙ ТОЧКОЙ.

Задачи оптимизации можно разбить на три группы.

- Задачи безусловной оптимизации. В этом случае имеется максимум (минимум) целевой функции. Дополнительные ограничения не накладываются.

- Задачи условной оптимизации линейного программирования. В этом случае целевая функция $f(\bar{x})$ и дополнительные условия $g_i(\bar{x})$ и $h_i(\bar{x})$ линейны.

- Задачи условной оптимизации нелинейного программирования. В этом случае либо целевая функция, либо одно из ограничений нелинейны.

1.19. Численные методы отыскания безусловного экстремума функции одной переменной

Задача отыскания минимума функции одной переменной нередко возникает в практических приложениях. Кроме того, многие численные методы отыскания минимума функции многих переменных сводятся к задачам одномерного поиска.

Один из подходов к отысканию экстремума функции $f(x)$ вытекает из известной теоремы анализа и заключается в замене задачи на экстремум задачей поиска решений нелинейного уравнения или системы нелинейных уравнений

$$f'(x) = 0. \quad (1.38)$$

Различные методы нахождения корней нелинейных уравнений описаны в § 1.8—1.13. В этих разделах приведены многие современные методы решения нелинейных уравнений и систем. Такой путь тем не менее имеет тот недостаток, что решение уравнения (1.38) может быть достаточно сложной задачей и итерационный процесс решения уравнения (1.38) может сходиться очень медленно.

Однако для поиска экстремума одномерной функции можно не использовать необходимые условия экстремума.

Предположим, что функция является **УНИМОДАЛЬНОЙ**, т. е. имеет на рассматриваемом отрезке $[a, b]$ только один локальный минимум*.

В этом случае, вычисляя последовательно целевую функцию при возрастающих значениях x , будем получать ее изменяющиеся значения, пока не достигнем точки минимума. При этом не требуется дифференцируемости и даже непрерывности функции $f(x)$.

Правило выбора последовательности точек x определяет быстроту отыскания интервала, в котором находится точка минимума. Этот интервал значений, в котором находится точка x^* , называется **ИНТЕРВАЛОМ НЕОПРЕДЕЛЕННОСТИ**.

| Метод золотого сечения

Метод дихотомии имеет тот недостаток, что из каждого трех значений целевой функции, вычисленных в интервале неопределенности, в дальнейшем используются только два. Третье, вычисленное в точке $X = (a + b)/2 \pm \epsilon/2$, дополнительной информации не несет и в дальнейших расчетах не используется. В методе золотого сечения целевая функция вычисляется в точках интервала неопределенности, расположенных таким образом, чтобы одно из значений целевой функции давало новую полезную информацию на следующем шаге. Сущность метода состоит в том, что интервал неопределенности X делится на две неравные части z_1 и z_2 так, что отношение длины большего отрезка z_1 к длине всего интервала неопределенности равно отношению длины меньшего отрезка z_2 к длине большего (рис. 1.10). Подобное деление осуществлял еще Евклид. Таким образом,

$$z_1 + z_2 = X, \quad \frac{z_1}{X} = \frac{z_2}{z_1}.$$

Точка $x^ \in [a, b]$ является точкой локального минимума функции $f(x)$ на $[a, b]$, если при некотором достаточно малом $\epsilon > 0$ для всех $x \neq x^*$, $x \in [a, b]$, удовлетворяющих условию $|x - x^*| < \epsilon$, выполнено неравенство $f(x^*) \leq f(x)$.

ЕВКЛИД, ЭВКЛИД (Εὐκλείδης) — древнегреческий математик. Его научная деятельность протекала в Александрополисе в начале III в. до н. э. Евклид является автором работ по математике, астрономии, оптике, музыке. Главный его труд («Начала») содержит изложение планиметрии, стереометрии и ряда вопросов теории чисел. В частности, Е. была выполнена задача деления отрезка в данном отношении с помощью циркуля и линейки.

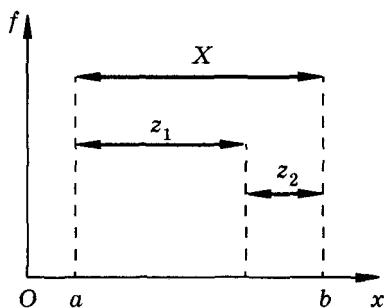


Рис. 1.10

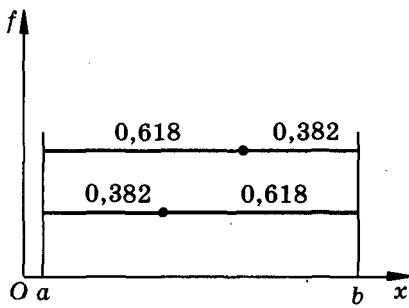


Рис. 1.11

Исключая из этих уравнений X , получаем квадратное уравнение относительно z_2/z_1 :

$$\left(\frac{z_2}{z_1}\right)^2 + \frac{z_2}{z_1} - 1 = 0. \quad (1.39)$$

Решая уравнение (1.39), получаем $z_2/z_1 = 0,618$. На рис. 1.11 показано деление интервала неопределенности в этом отношении.

Первые две точки располагаются симметрично на расстоянии 0,618 от концов интервала. В дальнейшем сохраняется один из этих интервалов, в котором располагается одна из точек, и симметрично ей располагается следующая. Таким образом, одно из значений целевой функции, которое требуется вычислить на следующем шаге, уже известно из предыдущего.

1.20. Численные методы отыскания безусловного экстремума функции многих переменных

Численные методы поиска экстремума функции многих переменных не только представляют самостоятельный интерес, но и используются как этап при поиске экстремума функции с ограничениями.

Один из подходов к задаче поиска экстремума функции $f(\bar{x})$ векторного аргумента размерности N (как и в случае одномерного

поиска) заключается в решении системы нелинейных алгебраических уравнений

$$\frac{\partial f}{\partial x_i} = 0, \quad i = 1, \dots, N, \quad (1.40)$$

получающейся при использовании необходимого условия экстремума. Методы для решения систем нелинейных уравнений, описанные в п. 1.14, могут плохо сходиться.

Другой подход заключается в построении такой последовательности $\bar{X}^{(0)}, \bar{X}^{(1)}, \dots$, для которой значения функции f образуют убывающую сходящуюся последовательность. Такие методы называются ПРЯМЫМИ МЕТОДАМИ решения задач оптимизации. ПОСЛЕДОВАТЕЛЬНОСТЬ ВЕКТОРОВ $\bar{X}^{(0)}, \bar{X}^{(1)}, \dots, \bar{X}^{(n)}$, удовлетворяющих условию $f(\bar{X}^{(0)}) > f(\bar{X}^{(1)}) > f(\bar{X}^{(n)})$, называется РЕЛАКСАЦИОННОЙ, а методы построения этих последовательностей — МЕТОДАМИ СПУСКА.

В методах спуска точки релаксационной последовательности вычисляются следующим образом:

$$\bar{X}^{(k+1)} = X^{(k)} + \alpha_k \bar{P}_k, \quad (1.41)$$

где \bar{P}_k — направление спуска от точки $\bar{X}^{(k)}$ к точке $\bar{X}^{(k+1)}$, а α_k — шаг вдоль этого направления. Различные варианты метода спуска отличаются друг от друга способами выбора направления спуска и шага вдоль него.

Эти методы принято делить на классы в зависимости от порядка используемых в методе производных целевой функции. Если в методе используется значение самой целевой функции, то это методы нулевого порядка, если значение первой производной — первого порядка и т. д. Методы одномерного поиска, рассмотренные в предыдущем разделе, являются методами нулевого порядка. Методы второго порядка для получения результата с заданной степенью точности требуют наименьшего числа итераций. Однако вычисление вторых производных может потребовать большого количества машинных операций. Поэтому быстрее (в смысле затрат машинного времени) могут сходиться методы первого и нулевого порядков. Хотя они и требуют большего количества итераций, но каждая итерация может потребовать меньшего количества арифметических операций, так что суммарное количество операций будет меньше. Поэтому с точки зрения затрат машинного времени эти методы могут оказаться предпочтительнее.

| Градиентные методы

Естественно в формуле (1.41) при поиске минимума $f(\bar{x})$ выбирать в качестве направления спуска то направление, вдоль кото-

рого функция убывает всего скорее. Этим направлением является направление, противоположное градиенту, — АНТИГРАДИЕНТ*. Таким образом, формула (1.41) примет вид

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} - \alpha_k \bar{f}'(\bar{x}^{(k)}), \quad (1.42)$$

где $\bar{f}'(\bar{x}^{(k)})$ — градиент функции f в точке $\bar{x}^{(k)}$. Методы спуска, в которых в качестве направления спуска используется направление, совпадающее с антиградиентом, называются ГРАДИЕНТНЫМИ и отличаются друг от друга выбором шага вдоль направления спуска. Эти методы являются МЕТОДАМИ ПЕРВОГО ПОРЯДКА.

| Градиентные методы с дроблением шага

Шаг спуска при использовании формулы (1.42) выбирается из того условия, что значение функции на каждом шаге меньше, чем на предыдущем, т. е.

$$f(\bar{x}^{(k)}) > f(\bar{x}^{(k)} - \alpha_k \bar{f}'(\bar{x}^{(k)})). \quad (1.43)$$

При этом возникает проблема с выбором шага α_k . Если шаг задать очень малым, таким, что условие (1.43) будет заведомо выполняться, это может потребовать очень большого числа итераций. Но больший шаг может привести к нарушению условия (1.43) либо привести к колебаниям около точки минимума. Поэтому итерационный процесс, удовлетворяющий условию (1.43), строится следующим образом. Выбирается шаг α_0 , и с этим шагом, постоянным для нескольких итераций, проводится расчет по формуле (1.42) и для k -й итерации проверяется условие (1.42). Если оно выполняется, то с этим шагом проводятся и следующие итерации. Если условие (1.42) не выполняется, то шаг α_0 уменьшается до тех пор, пока условие (1.43) не будет выполнено. Графическая иллюстрация метода приведена на рис. 1.12.

К точке минимума $\bar{x}^{(k)}$ приближается некоторая ломаная

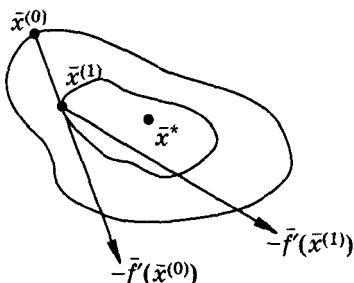


Рис. 1.12

*Из курса математического анализа известно, что градиент функции $f(\bar{x})$ в точке $\bar{x}^{(k)}$ направлен в сторону наискорейшего возрастания функции $f(\bar{x})$ и ортогонален линии уровня (поверхности постоянного значения функции $f(\bar{x})$), проходящей через точку $\bar{x}^{(k)}$.

$\bar{x}^{(0)}, \bar{x}^{(1)}, \dots, \bar{x}^{(k)}$, ортогональная в каждой точке соответствующей линии уровня. Эта ломаная называется градиентной кривой и удовлетворяет уравнению

$$\frac{d\bar{x}}{d\alpha} = -\bar{f}'(\bar{x}).$$

| Метод наискорейшего спуска

В методе наискорейшего спуска шаг на каждой итерации выбирается из условия минимума функции $f(\bar{x})$ вдоль направления спуска, т. е.

$$f(\bar{x}^{(k)} - \alpha_k \bar{f}'(\bar{x}^{(k)})) = \min f(\bar{x}^{(k)} - \alpha \bar{f}'(\bar{x}^{(k)})).$$

В этом методе на каждой итерации приходится решать задачу одномерной оптимизации, что, безусловно, усложняет алгоритм и требует дополнительных вычислений. Тем не менее метод дает выигрыш машинного времени, так как движение вдоль направления спуска осуществляется с оптимальным шагом.

Геометрическая иллюстрация метода приведена на рис. 1.13. Здесь каждая последующая точка находится как точка касания антиградиента и линии минимального уровня в данном направлении. Звенья градиентной кривой в этом случае ортогональны друг другу.

Действительно, шаг α на каждой итерации выбирается из условия минимума по α функции

$$\varphi(\alpha) = f(\bar{x}^{(k)} - \alpha \bar{f}'(\bar{x}^{(k)})).$$

Рис. 1.13

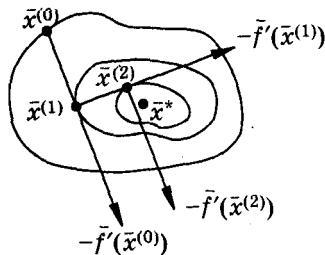
Дифференцируя φ по α , получаем

$$\frac{d\varphi}{d\alpha} = \frac{d}{d\alpha} [f(\bar{x}^{(k)} - \alpha \bar{f}'(\bar{x}^{(k)}))] = \frac{\partial f}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial \alpha} = -\bar{f}'(\bar{x}^{(k+1)}) \bar{f}'(\bar{x}^{(k)}) = 0,$$

т. е. направления спуска на двух последовательных итерациях ортогональны.

| Метод покоординатного спуска

Идея метода покоординатного спуска заключается в том, что спуск осуществляется не по всем координатам одновременно, а последовательно по каждой из них следующим образом.



Пусть $\bar{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$ — точка начального приближения (здесь нижний индекс — номер координаты). Осуществим спуск по первой координате x_1 . Для этого фиксируем все координаты, кроме первой, вычисляем частную производную функции f по первой координате $\frac{\partial f}{\partial x_1}\Big|_{\bar{x} = \bar{x}^{(0)}}$ и находим следующее приближение:

$$\bar{x}^{(1)} = \bar{x}^{(0)} - \alpha_1 \frac{\partial f}{\partial x_1}\Big|_{\bar{x} \approx \bar{x}^{(0)}} \cdot \bar{e}_1,$$

где $\bar{e}_1 = (1, 0, \dots, 0)$ — единичный вектор оси x_1 .

Следующая итерация происходит аналогично: фиксируются все координаты, кроме x_2 , и $\bar{x}^{(2)}$ вычисляется по формуле

$$\bar{x}^{(2)} = \bar{x}^{(1)} - \alpha_2 \frac{\partial f}{\partial x_2}\Big|_{\bar{x} = \bar{x}^{(0)}} \cdot \bar{e}_2,$$

где $\bar{e}_2 = (0, 1, \dots, 0)$ — единичный вектор оси x_2 .

Таким образом, в методе покоординатного спуска на каждой итерации меняется только одна координата, т. е. спуск осуществляется по ломаной, состоящей из отрезков прямых, параллельных осям координат. Геометрическая интерпретация метода для случая функции двух независимых переменных изображена на рис. 1.14. Здесь через x^* обозначено решение задачи оптимизации.

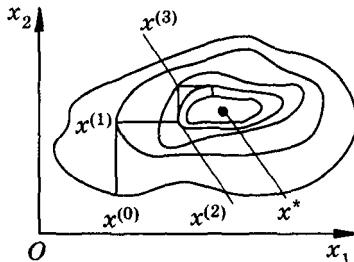


Рис. 1.14

Сходимость градиентных методов

В рассмотренных выше методах последовательность $f^{(k)}(\bar{x})$ сходится к точке минимума функции $f(\bar{x})$ из любой точки начального приближения $\bar{x}^{(0)}$. Если функция $f(x)$ ограничена снизу, ее градиент удовлетворяет УСЛОВИЮ ЛИПШИЦА (т. е. существует

ЛІПШІЦ РУДОЛЬФ ОТТО СІГІЗМУНД (Lipschitz Rudolf Otto Sigismund; 1832—1903) — немецкий математик. Ему принадлежат труды в различных областях анализа, теории чисел и многомерной геометрии. Л. установил одно из условий единственности решения задачи Коши для обыкновенного дифференциального уравнения (условие Липшица).

постоянная R такая, что неравенство $\|f'(x) - f'(\bar{y})\| \leq R\|\bar{x} - \bar{y}\|$ справедливо для любых \bar{x} и \bar{y} , принадлежащих допустимому множеству) и выбор значения шага на каждой итерации производится одним из описанных выше способов.

Теорема гарантирует сходимость либо к нижней грани, либо к стационарной точке. Градиентный метод с выбором шага по (1.42) гарантирует сходимость в точку минимума, если $f(\bar{x})$ является сильно выпуклой функцией. Скорость сходимости оценивается неравенством

$$\|\bar{x}^{(k+1)} - \bar{x}^*\| \leq \frac{M-m}{M+m} \|\bar{x}^{(k)} - \bar{x}^*\|, \quad (1.44)$$

т. е. итерации сходятся как геометрическая прогрессия со знаменателем прогрессии $q = (M-m)/(M+m)$. В качестве M и m в (1.44) могут, например, быть взяты максимальное и минимальное значения собственных чисел матрицы $f''(x)$.

В качестве критерия окончания итерационного процесса обычно берется одно из следующих условий:

$$\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\| \leq \varepsilon, \quad \left\| \frac{\partial f}{\partial \bar{x}} \right\|_{\bar{x} = \bar{x}^{(k)}} \leq \varepsilon, \quad \|f(\bar{x}^{(k+1)}) - f(\bar{x}^{(k)})\| \leq \varepsilon.$$

Если величины M и m незначительно отличаются друг от друга, то значение q мало и скорость сходимости метода достаточно хорошая. Если же $M \gg m$, значение q близко к единице и градиентные методы сходятся плохо. В этом случае рельеф функции носит так называемый *овражный* характер, т. е. линии уровня отличаются большой кривизной.

ПРИМЕР. Рассмотрим функцию $f(x, y) = x^2 + 2Sy^2$. Ее линии уровня сильно вытянуты вдоль оси x . Траектория градиентного спуска для этой функции характеризуется спуском на дно *оврага* и затем медленным движением вдоль его дна к точке минимума. Изменить рельеф функции можно введением новых переменных, что может изменить характер сходимости итерационного процесса. Так, заменой $x_1 = x$; $y_1^2 = 2Sy^2$ сведем приведенную функцию к функции $f(x_1, y_1) = x_1^2 + y_1^2$, линиями уровня которой являются окружности и градиент направлен в точку минимума функции.

| Метод сопряженных градиентов

Метод сопряженных градиентов, как и градиентные методы, относится к методам первого порядка. Однако по скорости сходимости он приближается к методам второго порядка. При этом квадратичная функция n переменных минимизируется таким

методом за n или менее шагов. Если учесть, что в окрестности точки минимума любая гладкая функция с хорошей степенью точности аппроксимируется квадратичной, то метод сопряженных градиентов имеет хорошую скорость сходимости и для неквадратичных функций. При этом он становится не конечным, а итерационным.

Как и для градиентных методов, алгоритм нахождения последовательности $\bar{x}^{(k)}$ строится по формуле

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} - \alpha_k \bar{P}_k,$$

но направление спуска \bar{P}_k выбирается из условия H -сопряженности* векторов \bar{P}_k и \bar{P}_{k-1} . Таким образом, для минимизации квадратичной функции n переменных необходимо определить n взаимно H -сопряженных направлений \bar{P}_k .

Рассмотрим алгоритм одного из методов сопряженных градиентов — метода Флетчера — Ривса, в котором выбор H -сопряженных направлений осуществляется совместно с одномерной оптимизацией $f(x)$ по α .

В точке $x^{(0)}$ вычисляется антиградиент $f(\bar{x})$. На k -м шаге решается задача оптимизации функции $f(\bar{x}^{(k)} + \alpha_k \bar{P}_k)$ по α , в результате этого определяется шаг α_k и точка $\bar{x}^{(k+1)}$. Если $f'(\bar{x}^{(k+1)}) = 0$, то задача определения минимума функции $f(\bar{x})$ решена. В ином случае направление спуска \bar{P}_{n+1} из точки $\bar{x}^{(k+1)}$ ищется следующим образом:

$$\bar{P}_{k+1} = -\bar{f}'(\bar{x}^{(k+1)}) + \frac{(\bar{f}'(\bar{x}^{(k+1)}), \bar{f}'(\bar{x}^{(k+1)}))}{(\bar{f}'(\bar{x}^{(k)}), \bar{f}'(\bar{x}^{(k+1)}))} \bar{P}_k. \quad (1.45)$$

Исследование алгоритма Флетчера — Ривса (1.45), как уже отмечалось, позволяет найти минимум квадратичной функции и аргументов не более чем за n шагов.

1.21. Численные методы отыскания условного экстремума

Рассмотренные методы относятся к случаю отыскания безусловного экстремума целевой функции $f(\bar{x})$, где $\bar{x} = (x_1, \dots, x_n)$. Однако в инженерной практике часто встречаются задачи, в которых

*Два вектора \bar{x} и \bar{y} называют H -сопряженными по отношению к матрице H , если $(x, Hy) = 0$.

необходимо найти экстремум целевой функции $f(\bar{x})$, причем должны удовлетворяться еще и условия:

$$g_i(\bar{x}) \geq 0, \quad i = 1, \dots, m, \quad (1.46)$$

называемые уравнениями связи.

Точка \bar{x} является ТОЧКОЙ ОТНОСИТЕЛЬНОГО МИНИМУМА функции $f(\bar{x})$, если существует $\varepsilon > 0$ такое, что для всех x , удовлетворяющих условию $\|x - x^*\| < \varepsilon$ и уравнениям связи, выполняется неравенство

$$f(x) \geq f(\bar{x}^*).$$

Исследуем наиболее распространенные методы поиска условного экстремума.

Метод исключения

Рассмотрим случай, когда все m ограничений (1.46) представляют собой равенства, т. е.

$$g_i(\bar{x}) = 0, \quad i = 1, \dots, m. \quad (1.47)$$

Предположим, что якобиан, составленный из частных производных $g_i(\bar{x})$ по первым m переменным, не равен нулю, т. е.

$$\begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_m} \end{vmatrix} \neq 0.$$

Следовательно, система (1.47) разрешима относительно x_1, x_2, \dots, x_m , и ее можно представить в виде

$$x_i = \varphi_i(x_{m+1}, x_{m+2}, \dots, x_n), \quad i = 1, \dots, m. \quad (1.48)$$

Переменные $x_{m+1}, x_{m+2}, \dots, x_n$ принято называть независимыми, а x_1, x_2, \dots, x_m — зависимыми. Под зависимыми и независимыми переменными в данном разделе понимаются функции, входящие в формулы задачи на условный экстремум.

Подставляя выражения для зависимых переменных (1.48) в функцию $f(\bar{x})$, получим задачу отыскания безусловного экстремума функции $(n-m)$ переменных

$$\begin{aligned} f(\varphi_1(x_{m+1}, \dots, x_n), \dots, \varphi_m(x_{m+1}, \dots, x_n), x_{m+1}, \dots, x_n) = \\ = \tilde{f}(x_{m+1}, \dots, x_n). \end{aligned}$$

Хотя метод исключения прост по своей идее, он имеет недостаток, заключающийся в том, что решение системы (1.47) относительно части компонент найти бывает трудно, а порой и невозможно. Поэтому чаще используется другой путь отыскания точки относительного экстремума, не требующий явного выражения зависимых переменных через независимые, — метод множителей Лагранжа.

Метод множителей Лагранжа

Из курса математического анализа известно, что в точке x^* безусловного экстремума функции $f(\bar{x})$ ее полный дифференциал равен нулю, т. е.

$$df(x^*) = \sum_{j=1}^m \frac{\partial f}{\partial x_j} dx_j + \sum_{k=m+1}^n \frac{\partial f}{\partial x_k} dx_k = 0, \quad (1.49)$$

где dx_j — дифференциалы зависимых, а dx_k — дифференциалы независимых переменных, удовлетворяющие соотношениям

$$\sum_{j=1}^m \frac{\partial g_i}{\partial x_j} dx_j + \sum_{k=m+1}^n \frac{\partial g_i}{\partial x_k} dx_k = 0, \quad i = 1, \dots, m, \quad (1.50)$$

полученным при дифференцировании уравнений связи.

Исключим из уравнений (1.49) и (1.50) дифференциалы зависимых переменных. Для этого поступим следующим образом: умножим каждое из уравнений (1.50) на произвольные множители $\lambda_1, \lambda_2, \dots, \lambda_m$ соответственно и сложим с уравнением (1.49). В результате получим выражение

$$\begin{aligned} & \sum_{j=1}^m \left(\frac{\partial f}{\partial x_j} + \lambda_1 \frac{\partial g_1}{\partial x_j} + \lambda_2 \frac{\partial g_2}{\partial x_j} + \dots + \lambda_m \frac{\partial g_m}{\partial x_j} \right) dx_j + \\ & + \sum_{k=m+1}^n \left(\frac{\partial f}{\partial x_k} + \lambda_1 \frac{\partial g_1}{\partial x_k} + \lambda_2 \frac{\partial g_2}{\partial x_k} + \dots + \lambda_m \frac{\partial g_m}{\partial x_k} \right) dx_k = 0. \end{aligned} \quad (1.51)$$

Выберем множители $\lambda_1, \lambda_2, \dots, \lambda_m$ таким образом, чтобы коэффициенты при дифференциалах зависимых переменных обратились в нуль, т. е. чтобы выполнялось равенство

$$\left. \left(\frac{\partial f}{\partial x_j} + \lambda_1 \frac{\partial g_1}{\partial x_j} + \lambda_2 \frac{\partial g_2}{\partial x_j} + \dots + \lambda_m \frac{\partial g_m}{\partial x_j} \right) \right|_{\bar{x} = \bar{x}^*} = 0, \quad j = 1, \dots, m. \quad (1.52)$$

Предположим, что якобиан функции $g_i(x)$, состоящий из частных производных по первым m элементам, отличен от нуля:

$$\begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_m} \end{vmatrix} \neq 0.$$

Если это условие выполняется, то уравнения (1.52), являющиеся системой линейных алгебраических уравнений относительно коэффициентов $\lambda_1, \lambda_2, \dots, \lambda_m$, имеют единственное решение. В этом случае в уравнении (1.51) останутся члены, содержащие дифференциалы только независимых переменных. Следовательно, и коэффициенты при дифференциалах зависимых переменных должны быть равны нулю, т. е. должны удовлетворяться соотношения

$$\frac{\partial f}{\partial x_k} + \lambda_1 \frac{\partial g_1}{\partial x_k} + \lambda_2 \frac{\partial g_2}{\partial x_k} + \dots + \lambda_m \frac{\partial g_m}{\partial x_k} = 0, \quad k = m+1, \dots, n. \quad (1.53)$$

Таким образом, получена система, содержащая $(n+m)$ уравнений (1.47), (1.52), (1.53) относительно $(n+m)$ неизвестных $\lambda_1, \lambda_2, \dots, \lambda_m, x_1, x_2, \dots, x_n$. Найденная в результате решения этой системы точка \bar{x} является точкой экстремума функции $f(x)$ и удовлетворяет уравнениям связи (1.47). Этую точку принято называть УСЛОВНО-СТАЦИОНАРНОЙ ТОЧКОЙ. Для того чтобы определить, является ли условно-стационарная точка точкой минимума или точкой максимума, необходимо использовать производные более высокого порядка, чем первый.

Перечислим основные этапы метода множителей Лагранжа.

- Составляется функция Лагранжа

$$L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x).$$

- Приравниваются к нулю ее частные производные по x_m и по λ :

$$\frac{\partial L}{\partial x_j} = \frac{\partial f}{\partial x_j} + \sum_{i=1}^m \lambda_i \frac{\partial g_i}{\partial x_j} = 0, \quad j = 1, \dots, n, \quad (1.54)$$

$$\frac{\partial L}{\partial \lambda_i} = g_i(x) = 0, \quad i = 1, \dots, m. \quad (1.55)$$

- Ищется решение системы (1.54), (1.55)
- $x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m$.
- Определяется характер условно-стационарной точки.

1.22. Линейное программирование

При оптимизации экономических планов возникают задачи на минимум линейной функции n переменных

$$L(x) \equiv \sum_{i=1}^n c_i x_i = \min \quad (1.56)$$

при наличии линейных дополнительных условий трех типов

$$x_i \geq 0, \quad 1 \leq i \leq n, \quad (1.57)$$

$$\sum_{i=1}^n a_{ji} x_i = b_j, \quad 1 \leq j \leq m, \quad (1.58)$$

$$\sum_{i=1}^n a_{ji} x_i \leq b_j, \quad m \leq j \leq M. \quad (1.59)$$

Каждое из условий типа неравенств (1.57) или (1.59) определяет полупространство, ограниченное гиперплоскостью; все эти условия вместе определяют выпуклый n -мерный многогранник J , являющийся пересечением соответствующих полупространств. С математической точки зрения условия (1.57) и (1.59) однотипны, но по традиции их записывают указанным образом. Условия типа равенств (1.58) выделяют из n -мерного пространства $(n-m)$ -мерную плоскость. Ее пересечение с областью J дает выпуклый $(n-m)$ -мерный многогранник G ; задача состоит в том, чтобы найти минимум линейной функции (1.56) в многограннике G .

ПРИМЕР. Задача распределения производства однотипной продукции по разным заводам.

Пусть x_i — выпускаемое i -м заводом количество продукции (оно должно быть неотрицательным), c_i — себестоимость одного изделия на этом заводе, a_{ji} при $j > m$ — расход сырья j -го вида и a_{ji} при $2 \leq j \leq m$ — расход заработной платы и других аналогичных показателей j -го вида при выпуске единицы продукции на

данном заводе. Положим $a_{1j} = 1$; тогда b_1 является суммарным выпуском продукции по всем заводам, а b_j при $2 \leq j \leq m$ — полной заработной платой и аналогичными данными по всей отрасли, суммы (1.59) — расходом сырья по всем заводам, а L — себестоимостью общей продукции. Требуется, чтобы себестоимость продукции была минимальной, выпуск продукции, расход заработной платы и т. д. — заданными, а фонды сырья b_j при $m < j$ не перерасходовались. Задача заключается в том, чтобы распределить неотрицательные плановые задания x_i по всем заводам так, чтобы удовлетворить всем этим требованиям.

Отметим терминологию, установившуюся в экономике. Вектор x , удовлетворяющий всем дополнительным условиям, называют *планом*; если он к тому же соответствует вершине многогранника G , то *опорным планом*. Решение экстремальной задачи (1.56) называют *оптимальным планом*, столбцы прямоугольной матрицы A — *векторами условий*, а столбец b — *вектором ограничений*. В задачах экономики обычно все коэффициенты a , b , c неотрицательны, хотя для последующего изложения это несущественно.

Многогранник условий G — выпуклый (он может быть и неограниченным). Поэтому внутри него линейная функция $L(x)$ не может достигать минимума. Ее минимум (если он существует) достигается обязательно в какой-нибудь вершине многогранника. При вырождении он может достигаться во всех точках ребра или даже p -мерной ограничивающей плоскости ($p < n - m$). Поэтому теоретически задача линейного программирования проста. Достаточно вычислить значения функции в конечном числе точек — в вершине многогранника — и найти среди этих значений наименьшее.

Сложность заключается в том, что типичное в экономике число переменных — это сотни и даже тысячи. Предложен ряд способов уменьшения числа искомых вершин, в частности **СИМПЛЕКС-МЕТОД**.

Глава 2

ПРИБЛИЖЕНИЕ ФУНКЦИЙ. ЧИСЛЕННОЕ ИНТЕГРИРОВАНИЕ И ДИФФЕРЕНЦИРОВАНИЕ

Материал этой главы объединяет общий подход к решению задач. Сначала излагается метод интерполяции и строятся интерполяционные многочлены в форме Лагранжа и Ньютона. Затем рассмотрена сплайн-интерполяция. В качестве примера метода среднеквадратичной аппроксимации изучается метод наименьших квадратов. Формулы численного дифференцирования получены дифференцированием интерполяционного многочлена и разложением в ряд Тейлора. Обсуждается точность формул и пути повышения точности численного дифференцирования. Излагаемые затем формулы численного дифференцирования получены на основе многочленной интерполяции. Исследованы точность и характер сходимости этих формул, рассмотрена возможность уточнения результата численного интегрирования (метод Рунге—Ромберга). В конце главы даны теоретические основы метода статистических испытаний.

2.1. Постановка задачи об аппроксимации функций

Для практических приложений важными являются следующие задачи. *Первая* состоит в замене некоторой функции, заданной аналитически или таблично, другой функцией, близкой к исходной, но более простой и удобной для вычислений. Например, замена функции многочленом позволяет получать простые формулы численного интегрирования и дифференцирования; замена таблицы приближающей функцией позволяет получать значения

в ее промежуточных точках. Эти примеры можно продолжить и дальше.

Возникает также и *вторая задача — восстановление функции на некотором отрезке по заданным на этом отрезке значениям функции в дискретном множестве точек*. Отметим также, что теория приближения функций является важным вспомогательным аппаратом при численном решении дифференциальных уравнений.

В общем случае при постановке задачи приближения необходимо действовать по следующему алгоритму.

- Определить, какой класс приближенных функций необходимо выбрать. Ответ на такой вопрос зависит от вида приближаемой функции и целей, для которых в дальнейшем будет использоваться приближающая функция. Широко используются следующие классы функций: многочлены, тригонометрические функции, показательные функции и др.
- Выбрать критерий близости исходной и приближающей функций. В качестве критерия можно выбрать, например, точное совпадение приближаемой и приближающей функций в узловых точках (лагранжева интерполяция); минимум суммы квадратов отклонения в узловых точках (метод наименьших квадратов) и др. Как и при выборе класса приближающих функций, выбор критерия близости исходной и приближающей функций определяется целью построения приближающей функции и может существенно повлиять на результаты. При аппроксимации экспериментальных результатов целесообразно использовать среднеквадратичное приближение. На рис. 2.1 показаны два варианта приближения функций: кривая I соответствует лагранжевой интерполяции, а кривая II — среднеквадратичному приближению. Здесь точками обозначены точные значения функции $f(x)$.

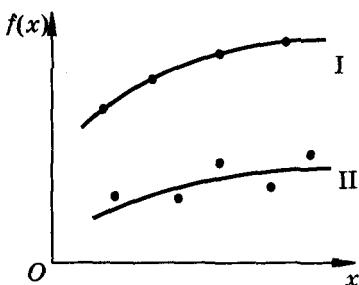


Рис. 2.1

На рис. 2.1 видно, что кривая II не проходит через узлы интерполяции, а сглаживает расчетные или экспериментальные погрешности.

- Необходимо указать правило, позволяющее с заданной степенью точности получить значение функции в промежутках между узлами, в частности, ответить на вопросы, какие узлы использовать для построения приближающей функции и как их расположить.

Таким образом, построение приближающей функции существенно зависит от ответа на перечисленные вопросы.

2.2. Интерполяция

Пусть на отрезке $[a, b]$ задано дискретное множество несовпадающих точек x_i , которые будем называть УЗЛАМИ и в которых известны значения функции $f_i = f(x_i)$, $i = 0, \dots, n$. Потребуем, чтобы приближающая функция $\varphi(x, a)$ совпадала с приближаемой функцией $f(x)$ в $(n + 1)$ узлах таблицы, т. е. потребуем выполнения равенства

$$\varphi(x_i, a_0, \dots, a_n) = f(x_i) = f_i, \quad i = 0, \dots, n. \quad (2.1)$$

Этот способ построения приближающей функции, при котором в узлах значения приближаемой и приближающей функций совпадают, называется интерполяцией или ЛАГРАНЖЕВОЙ ИНТЕРПОЛЯЦИЕЙ. Наиболее распространен способ линейной интерполяции, в случае которой приближающая функция ищется в виде линейной комбинации некоторых базисных функций $\varphi_i(x)$:

$$\varphi(x, a_0, \dots, a_n) = \sum_{i=0}^n a_i \varphi_i(x). \quad (2.2)$$

Система функций $\varphi_i(x)$ должна быть линейно независимой, так как в противном случае число членов в сумме можно было бы уменьшить и, кроме того,

$$\det \begin{vmatrix} \varphi_0(x_0) & \varphi_1(x_0) & \dots & \varphi_n(x_0) \\ \varphi_0(x_1) & \varphi_1(x_1) & \dots & \varphi_n(x_1) \\ \dots & \dots & \dots & \dots \\ \varphi_0(x_n) & \varphi_1(x_n) & \dots & \varphi_n(x_n) \end{vmatrix} \neq 0.$$

Подставляя функцию (2.2) в равенство (2.1), получаем систему линейных уравнений для определения коэффициентов a_i :

$$f_k = \sum_{i=1}^n a_i \varphi_i(x_k), \quad k = 0, \dots, n.$$

В качестве базисных функций можно выбрать любую линейно независимую систему функций, но чаще всего выбираются степенные функции $1, x, x^2, \dots, x^n$. Это объясняется тем, что многочлены легко вычисляются и теория интерполяции многочлена-

ми хорошо разработана. В случае приближения многочленами приближающую функцию ищем в виде многочлена степени n :

$$P_n(x) = \sum_{i=0}^n a_i x^i, \quad (2.3)$$

где нижний индекс n указывает на степень интерполяционного многочлена. Подставляя в функцию (2.3) значения узлов и используя условие $P_n(x_i) = f_i$, получаем систему линейных алгебраических уравнений относительно коэффициентов a_i :

$$\sum_{i=0}^n a_i x^i = f_k, \quad k = 0, \dots, n. \quad (2.4)$$

Определитель системы (2.4) является в случае несовпадения узлов интерполяции отличным от нуля определителем Вандермонда:

$$\begin{vmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{vmatrix} = \prod_{j>i>0}^n (x_j - x_i) \neq 0.$$

Таким образом, решение системы (2.4) существует и единственно. А это значит, что с точностью до формы записи существует единственный интерполяционный многочлен $P_n(x)$.

2.3. Интерполяционный многочлен в форме Лагранжа

Будем искать интерполяционный многочлен в виде

$$L_n(x) = \sum_{i=0}^n f_i l_i(x). \quad (2.5)$$

ВАНДЕРМОНД АЛЕКСАНДР ТЕОФИЛЬ (Vandermonde Alexandre Théophile; 1735—1796) — французский математик, чьи основные работы относятся к алгебре. В. заложил основы и дал логическое изложение теории детерминантов (определитель Вандермонда), а также выделил ее из теории линейных уравнений. Он ввел правило разложения детерминантов с помощью миноров второго порядка.

Здесь $l_i(x)$ — многочлены степени n , так называемые ЛАГРАНЖЕВЫ МНОГОЧЛЕНЫ ВЛИЯНИЯ, удовлетворяющие условию

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

Последнее условие означает, что любой многочлен $l_i(x)$ равен нулю при каждом x_j , кроме x_i , т. е. $x_0, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ — корни этого многочлена. Следовательно, лагранжевы многочлены $l_i(x)$ имеют вид

$$l_i(x) = C_i(x - x_0)(x - x_1)\dots(x - x_{i-1})(x - x_{i+1})\dots(x - x_n).$$

Так как по условию $l_i(x_i) = 1$, то

$$C_i = \frac{1}{(x_i - x_0)(x_i - x_1)\dots(x_i - x_{i-1})(x_i - x_{i+1})\dots(x_i - x_n)}.$$

Таким образом, лагранжевы многочлены влияния запишутся в виде

$$l_i(x) = \prod_{j=0, j \neq i}^n \frac{(x - x_j)}{(x_i - x_j)},$$

а интерполяционный многочлен (2.5) запишется в виде

$$L_n(x) = \sum_{i=0}^n f_i \prod_{j=0, j \neq i}^n \frac{(x - x_j)}{(x_i - x_j)}. \quad (2.6)$$

ЛАГРАНЖ ЖОЗЕФ ЛУИ (Lagrange Joseph Louis; 1736—1813) — выдающийся французский математик и механик, наиболее важные труды которого относятся к вариационному исчислению, к аналитической и теоретической механике. В основу статики Л. положил принцип возможных (виртуальных) перемещений. Он ввел обобщенные координаты и придал уравнениям движения механической системы форму, названную его именем. Л. получил ряд важных результатов в области анализа (формула остаточного члена ряда Тейлора, формула конечных приращений, теория ус洛вных экстремумов); в теории чисел (теорема Лагранжа); в алгебре (теория непрерывных дробей, приведение квадратичной формы к сумме квадратов); в теории дифференциальных уравнений (отыскание частного решения, изучение обыкновенного дифференциального уравнения первого порядка, линейного относительно искомой функции и независимой переменной, с переменными коэффициентами, зависящими от производной от искомой функции); в теории интерполяции (интерполяционная формула Лагранжа).

Интерполяционный многочлен в форме (2.6) называется ИНТЕРПОЛЯЦИОННЫМ МНОГОЧЛЕНОМ ЛАГРАНЖА. Перечислим основные достоинства этой формы записи интерполяционного многочлена.

- Число арифметических операций, необходимых для построения многочлена Лагранжа, пропорционально n^2 и является наименьшим для всех форм записи.
- Формула (2.6) в явном виде содержит значения функций в узлах интерполяции, что бывает удобно при некоторых вычислениях, в частности, при построении формул численного интегрирования.
- Формула (2.6) применима как для равноотстоящих, так и для неравноотстоящих узлов.
- Интерполяционный многочлен Лагранжа особенно удобен, когда значения функций меняются, а узлы интерполяции неизменны, что имеет место во многих экспериментальных исследованиях.

К недостаткам этой формы записи можно отнести то, что с изменением числа узлов приходится все вычисления проводить заново. Это затрудняет проведение апостериорных оценок точности (оценок, получающихся в процессе расчета).

Введем функцию $\omega_{n+1} = (x - x_0)(x - x_1)\dots(x - x_n) = \prod_{i=0}^n (x - x_i)$.

Отметим, что $\omega_{n+1}(x)$ есть многочлен степени $n + 1$. Тогда формулу (2.6) можно записать в виде

$$L_n(x) = \sum_{i=0}^n f_i \frac{\omega_{n+1}(x)}{(x - x_i)\omega'_{n+1}(x_i)}. \quad (2.7)$$

Приведем формулы линейной и квадратичной интерполяции по Лагранжу:

$$L_1(x) = f_0 \frac{(x - x_1)}{(x_0 - x_1)} + f_1 \frac{(x - x_0)}{(x_1 - x_0)}, \quad (2.8)$$

$$\begin{aligned} L_2(x) = & f_0 \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} + f_1 \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} + \\ & + f_2 \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}. \end{aligned} \quad (2.9)$$

Многочлен Лагранжа является в формуле (2.8) многочленом 1-й и в формуле (2.9) — многочленом 2-й степени.

Эти формулы наиболее часто используются на практике. Пусть задан $(n + 1)$ узел интерполяции. На этих узлах можно построить один интерполяционный многочлен n -й степени, $(n - 1)$ многочлен первой степени и большой набор многочленов степени меньше n , опирающихся на некоторые из этих узлов. Теоретически максимальную точность обеспечивает многочлен более высокой степени. Однако на практике наиболее часто используют многочлены невысоких степеней во избежание погрешностей при расчетах коэффициентов при больших степенях многочлена.

2.4. Интерполяционный многочлен в форме Ньютона

Интерполяционный многочлен в форме Ньютона представляет собой иную форму записи интерполяционного многочлена. Она имеет ряд достоинств по сравнению с многочленом Лагранжа.

Для вывода введем понятие РАЗДЕЛЕННОЙ РАЗНОСТИ. Разделенные разности *нулевого порядка* совпадают со значениями функции в узлах. Разделенные разности *первого порядка* $f(x_i, x_j)$ определяются через разделенные разности нулевого порядка:

$$f(x_i, x_j) = \frac{f_i - f_j}{x_i - x_j};$$

разделенные разности *второго порядка* определяются через разделенные разности первого порядка:

$$f(x_i, x_j, x_k) = \frac{f(x_i, x_j) - f(x_j, x_k)}{x_i - x_k}.$$

Разделенная разность порядка n определяется соотношениями

$$\begin{aligned} f(x_i, x_j, x_k, \dots, x_{n-1}, x_n) &= \\ &= \frac{f(x_i, x_j, x_k, \dots, x_{n-1}) - f(x_j, x_k, \dots, x_n)}{x_i - x_n}. \end{aligned} \quad (2.10)$$

Таким образом, для $(n + 1)$ -й точки могут быть построены разделенные разности до n -го порядка; разделенные разности более высоких порядков равны нулю. Из вида разделенных разнос-

тей можно сделать заключение, что они являются аналогом производных соответствующих порядков.

Между разделенными разностями и производными соответствующих порядков существует соотношение

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

Очевидно, что разделенная разность $(n + 1)$ -го порядка от многочлена n -й степени равна нулю.

Часто для таблиц с постоянным шагом h вводят КОНЕЧНЫЕ РАЗНОСТИ $\Delta^n f$, связанные с разделенными соотношением

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

Приведем вывод формулы интерполяционного многочлена Ньютона. Пусть аппроксимируемая функция $f(x)$ — многочлен n -й степени. Выберем точки x_0, x_1, \dots, x_n в качестве узлов и найдем такой интерполяционный многочлен, значения которого в узлах интерполяции совпадают со значениями функции $f(x)$.

Запишем, выбирая в качестве узла точку с координатой x , очевидные соотношения

$$f(x) = f(x_0) + (x - x_0) f(x, x_0). \quad (2.11)$$

Имеем также

$$f(x, x_0) = f(x_0, x_1) + (x - x_1) f(x, x_0, x_1). \quad (2.12)$$

Подставляя соотношение (2.12) в (2.11), получаем

$$f(x) = f(x_0) + (x - x_0) f(x_0, x_1) + (x - x_0)(x - x_1) f(x, x_0, x_1).$$

По определению разделенной разности (2.10) имеем

$$f(x, x_0, x_1) = f(x_0, x_1, x_2) + (x - x_2) f(x, x_0, x_1, x_2). \quad (2.13)$$

Подставляя соотношение (2.13) в выражение (2.12), получаем

$$\begin{aligned} f(x) &= f(x_0) + (x - x_0) f(x_0, x_1) + (x - x_0)(x - x_1) f(x_0, x_1, x_2) + \\ &+ (x - x_0)(x - x_1)(x - x_2) f(x, x_0, x_1, x_2). \end{aligned}$$

Продолжая этот процесс, в итоге приходим для интерполяционной функции к многочлену $P_n(x)$ в виде

$$\begin{aligned} P_n(x) &= f(x_0) + (x - x_0) f(x_0, x_1) + (x - x_0)(x - x_1) f(x_0, x_1, x_2) + \\ &+ \dots + (x - x_0)(x - x_1) \dots (x - x_n) f(x_0, x_1, \dots, x_n). \end{aligned} \quad (2.14)$$

Следующий член в этом ряду будет равен нулю, поскольку он должен содержать произведение функции $\omega_{n+1}(x)$ на разделенную разность $(n+1)$ -го порядка $f(x_0, x_1, \dots, x_n)$.

Так как $\omega_{n+1}(x)$ — многочлен степени $n+1$, а исходный многочлен имеет степень n , то разделенная разность $(n+1)$ -го порядка должна быть равна нулю. В силу условия единственности интерполяционного многочлена $P_n(x)$ существует интерполяционный многочлен n -й степени с узлами x_0, x_1, \dots, x_n .

Запись многочлена в формуле (2.14) является так называемым ИНТЕРПОЛЯЦИОННЫМ МНОГОЧЛЕНОМ НЬЮТОНА. Если функция $f(x)$ не является многочленом n -й степени, то для $P_n(x)$ имеет место формула (2.14), однако она приближает функцию $f(x)$ не точно, а с некоторой погрешностью. Отметим, что при добавлении новых узлов первые члены многочлена Ньютона остаются неизменными, что видно из сравнения формул (2.11) — (2.14).

Если функция задана в точках x_0, x_1, \dots, x_n , то при построении интерполяционного многочлена Ньютона удобно пользоваться таблицей, называемой таблицей разделенных разностей. Последовательность получения разделенных разностей для числа $n = 3$ приведена в таблице 2.1.

Таблица 2.1

x_0	$f(x_0)$	$f(x_0, x_1)$	$f(x_0, x_1, x_2)$	$f(x_0, x_1, x_2, x_3)$
x_1	$f(x_1)$	$f(x_1, x_2)$	$f(x_1, x_2, x_3)$	
x_2	$f(x_2)$			
x_3	$f(x_3)$			

За точностью расчета удобно следить по убыванию величин членов суммы (2.14). Если величина членов убывает быстро, то оставляют только те из них, которые больше допустимой погрешности. Для повышения точности интерполяции в сумму могут быть добавлены новые члены, что требует подключения дополнительных узлов. При этом безразлично, в каком порядке подключаются новые узлы. Этим формула Ньютона выгодно отличается от формулы Лагранжа, при добавлении в которую новых узлов все расчеты надо производить заново. Интерполяционный многочлен Ньютона удобен при построении формул численного дифференцирования, что очевидно из его структуры. Кроме того, он удобен для проведения апостериорных оценок.

Приведем формулы линейной и квадратичной интерполяции в соответствии с многочленом Ньютона. Имеем:

$$P_1(x) = f_0 + (x - x_0) \frac{f_0 - f_1}{x_0 - x_1}; \quad (2.15)$$

$$\begin{aligned} P_2(x) &= f_0 + (x - x_0) \frac{f_0 - f_1}{x_0 - x_1} + \\ &+ (x - x_0)(x - x_1) \frac{\frac{f_0 - f_1}{x_0 - x_1} - \frac{f_1 - f_2}{x_1 - x_2}}{x_0 - x_2}. \end{aligned} \quad (2.16)$$

Нетрудно убедиться, что эти формулы будут совпадать с соответствующими формулами для интерполяционного многочлена Лагранжа после простых преобразований. Очевидно, что это утверждение есть следствие условия единственности интерполяционного многочлена.

2.5. Погрешность и сходимость интерполяции

Построив для функции $f(x)$ интерполяционный многочлен $P_n(x)$ или $L_n(x)$, необходимо выяснить, насколько близко интерполяционный многочлен приближается к функции в других точках отрезка. Ошибку интерполяции можно представить в виде

$$R_n(x) = \varepsilon_n(x) = f(x) - P_n(x) = \omega_{n+1}(x) r(x), \quad (2.17)$$

так как ошибка интерполяции обращается в нуль в узлах сетки. Выражение для $r(x)$ найдем из следующих соображений. Введем функцию

$$q(x) = f(x) - P_n(x) - \omega_{n+1}(x) r(x^*),$$

где x^* — произвольная точка. Очевидно, что $q(x) = 0$ в $(n+2)$ -х точках, в которых $x = x_i$ для $i = 0, \dots, n$, и при $x = x^*$. Если $f(x)$ имеет $(n+1)$ -ю непрерывную производную, то и $q(x)$ можно проанализировать $(n+1)$ раз. А так как между двумя нулями гладкой функции лежит нуль ее производной, то $q'(x)$ обращается в нуль на всем интервале по крайней мере $(n+1)$ раз. Последовательно применяя это правило дальше, получаем, что между крайними нулями функции $q^{(n)}(x)$ лежит, согласно теореме Ролля, по крайней мере один нуль функции $q^{(n+1)}(x)$, т. е. существует

вует такая точка ξ , принадлежащая отрезку интерполирования, в которой выполняется равенство

$$q^{(n+1)}(\xi) = f^{(n+1)}(\xi) - (n+1)! r(x^*) = 0.$$

Следовательно, $r(x^*) = f^{(n+1)}(\xi)/(n+1)!$. Так как x^* — произвольная точка, то погрешность интерполяции запишется в виде

$$R_n(x) = \varepsilon_n(x) = f(x) - P_n(x) = \omega_{n+1}(x) \frac{f^{(n+1)}[\xi(x)]}{(n+1)!}, \quad (2.18)$$

т. е. можно сказать, что существует точка ξ , для которой выполнено равенство (2.18). Так как точка ξ неизвестна, то обычно используют мажорантную оценку. Заменяя величины, входящие в (2.18), их максимальными значениями, получаем при $M_{n+1} = \max |f^{(n+1)}(\xi)|$, $\xi \in [x_0, x_n]$ величину максимальной погрешности

$$|\varepsilon_n(x)| = |f(x) - P_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |\omega_{n+1}(x)|. \quad (2.19)$$

Формула (2.19) дает возможность провести априорную оценку точности, т. е. для случая аналитически заданной функции $f(x)$ провести оценку до начала вычислений. Однако так как величину производных интерполяционной функции бывает оценить достаточно сложно (например, для таблично заданной функции), то априорной оценкой пользуются редко. Удобнее применять апостериорную оценку по первому отброшенному члену интерполяционного многочлена Ньютона. Это удобство является следствием того обстоятельства, что в процессе построения многочлена Ньютона вычисляются разделенные разности, являющиеся, как отмечалось, аналогами производных соответствующих порядков.

На рис. 2.2 представлена функция $\omega_{(n+1)}(x)$ для равноотстоящих узлов. По рис. 2.2 на примере ин-

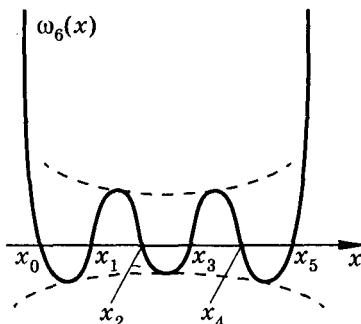


Рис. 2.2

РОЛЛЬ МИШЕЛЬ (Rolle Michel; 1652—1719) — французский математик, автор исследований, относящихся к решению в целых числах неопределенных линейных уравнений с двумя неизвестными. Им выведена теорема Ролля, послужившая основой метода отделения действительных корней алгебраических уравнений.

терполяционного многочлена пятой степени видно, что между центральными узлами интерполяции x_2 и x_3 имеет место наименьший экстремум функции $\omega_6(x)$. К крайним узлам экстремумы возрастают.

Приближение, при котором точка, в которой ищется значение приближающей функции, находится за пределами отрезка интерполяций, называется ЭКСТРАПОЛЯЦИЕЙ. При экстраполяции в точках, далеко отстоящих от концов отрезка интерполяции, ошибка приближения может стать значительной. Ошибка интерполяции в центральном интервале при постоянном шаге таблицы определяется выражением

$$|\varepsilon_n(x)| \leq \sqrt{\frac{2}{n}} M_{n+1} \left[\frac{h}{2} \right]^{n+1}. \quad (2.20)$$

Поэтому первым способом уменьшения ошибки при постоянном шаге таблицы является выбор точки x вблизи середины отрезка интерполирования.

Из выражения (2.20) видно, что если интерполяционный многочлен имеет порядок n , то ошибка интерполяции будет убывать как h^{n+1} при переходе от таблиц с более крупным шагом к таблицам с более мелким шагом. Следовательно, если производная, входящая в оценку погрешности, ограничена, то для фиксированной степени многочлена и изменения шага сетки погрешность $|f(x) - P_n(x)|$ неограниченно убывает. При этом многочлен равномерно сходится к $f(x)$ на ограниченном отрезке. Если значение производной оценивается легко, то из (2.20) легко оценить число узлов, обеспечивающих заданную точность интерполяции.

Другой способ уменьшения ошибки интерполяции заключается в увеличении числа узлов, т. е. степени интерполяции, при постоянном шаге сетки. При этом знаменатель в выражении (2.19) возрастает. Этот способ не всегда приводит к уменьшению ошибки за счет роста числа $(n+1)!$, так как функция $f(x)$ может не иметь $(n+1)$ -й производной. Кроме того, максимум этой функции может возрастать очень быстро.

Трудности, возникающие при многочленной интерполяции, хорошо иллюстрируются примером Рунге: интерполяция на отрезке $[-1, 1]$ функции $f(x) = 1/(1 + 25x^2)$ при равномерном распределении узлов (рис. 2.3). С увеличением порядка n интерпо-

РУНГЕ КАРЛ ДАВИД ТОЛЬМЕ (Runge Carle David Tolmé; 1856—1927) — немецкий математик и физик, основными направлениями исследований которого являлись прикладная математика и спектроскопия. Р. существенно развил численные методы решения дифференциальных уравнений (метод Рунге — Кутты).

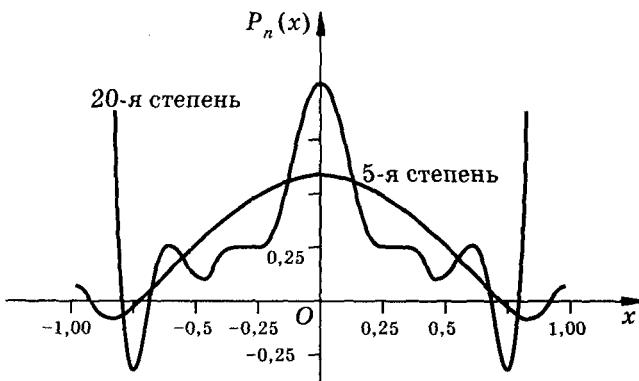


Рис. 2.3

ляционного многочлена возрастают колебательные свойства многочлена, и это приводит к тому, что интерполяционный многочлен при $0,726 \leq x \leq 1$ расходится.

Задача об оптимальном расположении узлов на отрезке интерполяции, приводящем к минимизации ошибки интерполяции, решена П. Л. Чебышевым. Им показано, что если узлы интерполяции не распределены равномерно, а совпадают с корнями так называемых чебышёвских полиномов степени $n + 1$, то ошибка многочленной интерполяции минимальна. При таком выборе узлов интерполяции проблема расходимости многочлена для процедуры Рунге исчезает. Если интерполяция производится на интервале (a, b) , то узлы интерполяционного многочлена следует расположить в нулях многочлена Чебышёва, т. е.

$$x_i = \frac{b+a}{2} + \frac{b-a}{2} z_i.$$

Здесь

$$z_i = \cos \left[\frac{\pi}{2(n+1)} + i \frac{\pi}{n+1} \right], \quad i = 0, \dots, n.$$

ЧЕБЫШЁВ ПАФНУТИЙ ЛЬВОВИЧ (1821—1894) — русский математик и механик, труды которого относятся к математическому анализу, теории чисел, теории вероятностей. В теории вероятностей основная заслуга Ч. состоит во введении в рассмотрение случайных величин, в доказательстве общей формы закона больших чисел и исследовании условий сходимости функций распределения сумм независимых слагаемых к нормальному закону. В области математического анализа ему принадлежат исследования по теории интегрирования, теории интерполяции. Многие открытия Ч. навеяны прикладными задачами.

На практике для повышения точности интерполяционного многочлена, как правило, уменьшают шаг таблицы, оставляя неизменной степень интерполирующего многочлена. Для того чтобы уменьшить влияние колебательных свойств интерполяционного многочлена, используют линейную или квадратичную интерполяцию.

2.6. Интерполяция сплайнами

Использование одной интерполяционной формулы на большом числе узлов, как отмечалось в предыдущем разделе, нецелесообразно. Такой интерполяционный многочлен существенно проявляет свои колебательные свойства, и его значения между узлами могут значительно отличаться от значений интерполируемой функции. Одна из возможностей преодоления этого недостатка заключается в применении СПЛАЙН-ИНТЕРПОЛЯЦИИ. Суть сплайн-интерполяции заключается в определении интерполирующей функции по формулам одного типа для различных подмножеств и в стыковке значений функции и ее производных на границах подмножеств.

Наиболее изученным и широко применяемым является вариант, в котором между любыми двумя точками строится многочлен n -й степени

$$S(x) = \sum_{k=0}^n a_{ik} x^k, \quad x_{i-1} \leq x \leq x_i,$$

который в узлах интерполяции принимает значения интерполируемой функции и непрерывен вместе со своими $(n - 1)$ -й производными. Такой кусочно-непрерывный интерполяционный многочлен называется СПЛАЙНОМ. Его коэффициенты находятся из следующих условий в узлах сетки: должны быть равны значения сплайна и приближаемой функции, а также $n - 1$ производные соответствующих многочленов, при этом подразумеваются производные, взятые на соседних интервалах.

Максимальная по всем частичным отрезкам степень многочленов называется *степенью сплайна*, а разница между степенью сплайна и порядком наивысшей непрерывной производной — *дефектом сплайна**.

* Этую величину принято обозначать через *def* — от *defect* (англ.).

Одним из наиболее распространенных интерполяционных сплайнов является кубический ($\text{def} = 1$). Для вывода уравнения кубического интерполяционного сплайна воспользуемся его интерпретацией как гибкой линейки, изогнутой таким образом, что она проходит через значения функции в узлах сетки, т. е. является упругим бруском в состоянии свободного равновесия. Так как это его состояние описывается уравнением $S''(x) = 0$, где $S''(x)$ — четвертая производная, то между каждой парой соседних узлов интерполяционная формула записывается в виде многочлена третьей степени. Этот многочлен удобно представить как

$$S(x) = a_i + b_i(x - x_{i-1}) + c_i(x - x_{i-1})^2 + d_i(x - x_{i-1})^3, \quad (2.21)$$

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

Для построения кубического сплайна необходимо построить n многочленов третьей степени, т. е. определить $4n$ неизвестных a_i, b_i, c_i, d_i . Эти коэффициенты ищутся из условий в узлах сетки. По определению в узлах сетки сплайн (2.21) должен принимать табличные значения функции (здесь $h_i = x_i - x_{i-1}$):

$$S(x_{i-1}) = a_i = f_{i-1}, \quad (2.22)$$

$$S(x_i) = a_i + b_i h_i + c_i h_i^2 + d_i h_i^3 = f_i. \quad (2.23)$$

Система (2.22), (2.23) содержит $2n$ уравнений. Дополнительные уравнения можно получить, если потребовать непрерывности первой и второй производных функции $S(x)$ во внутренних узлах сетки при $x_{i-1} \leq x \leq x_i$ ($i = 1, \dots, n-1$):

$$S'(x) = b_i + 2c_i(x - x_{i-1}) + 3d_i(x - x_{i-1})^2,$$

$$S''(x) = 2c_i + 6d_i(x - x_{i-1}).$$

Получим еще $2(n-1)$ уравнений:

$$b_{i+1} = b_i + 2c_i h_i + 3d_i h_i^2, \quad (2.24)$$

$$c_{i+1} = c_i + 3d_i h_i. \quad (2.25)$$

Недостающие два уравнения можно получить из граничных условий, предполагая нулевую кривизну сплайна на концах отрезка. Приравнивая значения второй производной в точках x_0 и x_n к нулю, получим два недостающих уравнения:

$$S''(x_0) = c_i = 0, \quad (2.26)$$

$$S''(x_n) = c_n + 3d_n h_n = 0. \quad (2.27)$$

Заметим, что граничные условия (2.26)–(2.27) дают погрешность $O(h^2)$ вблизи границы. Поэтому вместо них можно рекомендовать условие непрерывности $S'''(x)$ в двух внутренних (приграничных) узлах. Если существуют другие условия поведения функции в точках x_0 и x_n , то уравнения (2.26), (2.27) будут другими. Таким образом, система (2.22)–(2.27) для отыскания $4n$ неизвестных a_i, b_i, c_i, d_i ($i = 1, 2, \dots, n$) замкнута. Ее удобно решать, проведя предварительно следующие преобразования. Коэффициент a_i сразу получается из решения уравнения (2.22). Из уравнений (2.25) и (2.27) следует:

$$d_i = (c_{i+1} - c_i)/3h_i, \quad i = 1, 2, \dots, n, \quad (2.28)$$

$$d_n = -c_n/3h_n. \quad (2.28a)$$

Подставляя выражения (2.28), (2.28a) в выражение (2.23) и заменяя в нем a_i на f_{i-1} , получаем

$$b_i = [(f_i - f_{i-1})/h_i] - \frac{1}{3}h_i(c_{i+1} + 2c_i), \quad (2.29)$$

$$b_n = [(f_n - f_{n-1})/h_n] - \frac{2}{3}h_n c_n. \quad (2.29a)$$

Из уравнений (2.24), (2.25) имеем

$$b_{i+1} = b_i + h_i(c_i + c_{i+1})$$

или

$$b_i = b_{i-1} + h_{i-1}(c_{i-1} + c_i). \quad (2.29b)$$

Полагая в выражении (2.29) значение $i = n$ и сравнивая с выражением (2.29a), получаем

$$c_{n+1} = 0.$$

Подставляя теперь выражения (2.29), (2.29a) в (2.29b), приходим к системе уравнений относительно c_i :

$$c_1 = 0, \quad (2.30)$$

$$\begin{aligned} & h_{i-1}c_{i-1} + 2(h_{i-1} + h_i)c_i + h_i c_{i+1} = \\ & = 3[(f_i - f_{i-1})/h_i - (f_{i-1} - f_{i-2})/h_{i-1}], \quad i = 2, \dots, n, \end{aligned} \quad (2.30a)$$

$$c_n = 0. \quad (2.30b)$$

Матрица системы (2.30)–(2.30b) — трехдиагональная, решение находится методом прогонки. Решив ее, находим затем коэффициенты b_i из уравнений (2.29), (2.29a) и d_i — из уравнений (2.28), (2.28a).

Метод сплайн-интерполяции приводит к удовлетворительным результатам в процессе интерполяции непрерывных функций с гладкими первой и второй производными. При этом кубический интерполяционный сплайн, построенный по узлам $f_i = f(x_i)$ ($i = 0, \dots, n$), будет обладать минимальной кривизной по сравнению с кривизной любой интерполяционной функции, имеющей непрерывные первую и вторую производные. Сплайн-интерполяция функций с резким изменением производных сплайна может давать большие ошибки. Сплайны более высокого порядка, чем третий, используются редко, так как при вычислении большого числа коэффициентов может накапливаться ошибка, приводящая к значительным отклонениям.

Рассмотрим сплайн второго порядка. Для него

$$S(x) = a_i + b_i(x - x_{i-1}) + c_i(x - x_{i-1})^2,$$

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

Очевидно, что при $h_i = x_i - x_{i-1}$

$$S(x_{i-1}) = a_i = f_{i-1}, \tag{2.31}$$

$$S(x_i) = a_i + b_i h_i + c_i h_i^2 = f_i.$$

Условие совпадения первой производной $S'(x) = b_i + 2c_i(x - x_{i-1})$ в узлах дает

$$b_{i+1} = \begin{cases} f'_0 & \text{при } i = 0, \\ b_i + 2c_i h_i & \text{при } i = 1, 2, \dots, n-1. \end{cases} \tag{2.32}$$

Аналогично замечанию к уравнениям (2.26), (2.27) следует рекомендовать вместо условия (2.32) условие непрерывности $S''(x)$ в одном узле. Если принять, что в начальной точке заданы f_0 и f'_0 , т. е. a_1 и b_1 , то для определения коэффициентов a_i, b_i, c_i на каждом участке нужно решать систему линейных уравнений третьего порядка (2.31), (2.32). Линейный сплайн или сплайн первого порядка представляет собой ломаную линию, состоящую из отрезков прямых, соединяющих узлы.

2.7. Приближение методом наименьших квадратов

До сих пор рассматривалась интерполяция, т. е. такой способ приближения, когда значения приближаемой и приближающей функций совпадают в узлах некоторой сетки. Однако достаточно

часто, например, при аппроксимации большого числа экспериментальных точек, найденных с некоторой погрешностью, интерполяция становится неразумной. В этом случае целесообразно строить приближающую функцию таким образом, чтобы сгладить влияние погрешности измерения и числа точек эксперимента. Такое сглаживание реализуется при построении приближающей функции по методу наименьших квадратов. Вид приближающей функции может быть произвольным, далее рассмотрен случай, когда приближающая функция является многочленом. При этом добиваются минимизации суммы квадратов отклонений значений приближаемой и приближающей функций в узлах сетки. Эта сумма называется КВАДРАТИЧНЫМ ОТКЛОНЕНИЕМ.

Пусть функция f_j при $j = 0, 1, \dots, N$ задана таблично в узлах x_j . Необходимо построить такой многочлен $F_n(x) = \sum_{i=0}^n a_i x^i$, для которого минимально квадратичное отклонение

$$\Phi = \sum_{j=0}^N [F_n(x_j) - f_j]^2. \quad (2.33)$$

Очевидно, что минимума Φ можно добиться только за счет изменения коэффициента многочлена $F_n(x)$. Необходимые условия экстремума имеют вид

$$\frac{\partial \Phi}{\partial a_k} = 2 \sum_{j=0}^N \left[\sum_{i=0}^n a_i x_j^i - f_j \right] x_j^k = 0, \quad k = 0, 1, \dots, n. \quad (2.34)$$

Эту систему для удобства преобразуют к виду

$$\sum_{i=0}^n a_i \sum_{j=0}^N x_j^{k+i} = \sum_{j=0}^N f_j x_j^k, \quad k = 0, 1, \dots, n. \quad (2.35)$$

Система (2.35) называется НОРМАЛЬНОЙ СИСТЕМОЙ МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ и представляет собой систему линейных алгебраических уравнений относительно коэффициентов a_i . Решив систему, построим многочлен $F_n(x)$, приближающий функцию $f(x)$ и минимизирующий квадратичное отклонение.

Рассмотрим одно важное свойство системы (2.35). Предположим, что точки равномерно распределены на отрезке $[0, 1]$, т. е.

$x_j = j/N$, $j = 0, 1, \dots, N$. В этом случае сумму $\sum_{j=0}^N x_j^{k+i}$ можно приблизенно заменить интегралом

$$\sum_{j=0}^N x_j^{k+i} = N \sum_{j=0}^N \frac{1}{N} x_j^{k+i} \approx N \int_0^1 x_j^{k+i} dx = \frac{N}{k+i+1}.$$

Тогда определитель Δ системы (2.35) составляет

$$\Delta = N^{n+1} \det H_{n+1},$$

где H_{n+1} — матрица Гильберта порядка $n + 1$, элементы которой имеют вид $[1/(k+i+1)]$. Определитель этой матрицы

$$\det H_p = \frac{[1!2!\dots(p-1)!]^3}{p!(p+1)!\dots(2p-1)!}$$

и весьма быстро убывает с ростом p , что приводит к быстрому убыванию величины определителя системы Δ . Так $\det H_2 \approx 10^{-2}$ для $p = 2$ и $\det H_3 \approx 10^{-5}$ для $p = 3$. В первом случае $n = 1$, во втором $n = 2$.

Следовательно, система (2.35) с увеличением степени n приближающего многочлена становится плохо обусловленной и решение ее связано с большой потерей точности. Поэтому в методе наименьших квадратов, как правило, используют приближающий многочлен не выше третьей степени. При необходимости построения многочленов большей степени применяют приемы, позволяющие повысить обусловленность системы (2.35); обычно таких приемов два. В первом используют систему точек, позволяющую разбить систему (2.35) на две подсистемы меньшего порядка, во втором — систему ортогональных многочленов.

Как отмечалось, метод наименьших квадратов широко применяется для сглаживания экспериментальных кривых, полученных с некоторой погрешностью. Если степень аппроксимирующего многочлена равна числу точек, то среднеквадратичный многочлен совпадает с интерполяционным. Поэтому хорошее сглаживание будет при $n \ll N$. Но если n очень мало, то для описания сложной кривой коэффициентов может не хватить. Чтобы выбрать оптимальную степень многочлена, строят многочлен по методу наименьших квадратов некоторой степени n , вычисляют

ГИЛЬБЕРТ, ХИЛЬБЕРТ ДАВИД (Hilbert David; 1862—1943) — немецкий математик. Его основные труды посвящены теории инвариантов, теории алгебраических чисел, основаниям геометрии и др. Построенная им теория интегральных уравнений с симметричным ядром легла в основу современного функционального анализа. Данное Г. решение проблемы Дирихле положило начало разработке прямых методов в вариационном исчислении. Исследования, посвященные логическим основам математики, в большой мере идут по путям, намеченным Г., с использованием созданных им концепций.

квадратичное отклонение Φ и сравнивают его с известной величиной погрешности ϵ .

Если $\Phi \gg \epsilon$, т. е. математическая ошибка существенно превышает ошибку экспериментальных данных, то степень приближающего многочлена недостаточна для описания кривой. Если же $\Phi \ll \epsilon$, то старшие коэффициенты аппроксимации физически недостоверны. Хорошее сглаживание получается в том случае, когда $\Phi = \epsilon$. В этом случае степень приближающего многочлена оптимальна. Обычно начинают построение приближающего многочлена для случая $n = 1$ и увеличивают его степень до тех пор, пока отклонение Φ не станет примерно равным ϵ . Если при этом $n \ll N$, то приближающий многочлен выбран верно. Если это условие не соблюдается, то следует поискать более удачный вид приближающей функции.

Можно показать методами теории вероятностей, что функция $F_n(x)$, найденная по формулам (2.34), (2.35), с наибольшим правдоподобием совпадает с оценкой математического ожидания для заданной выборки (f_0, f_1, \dots, f_n) , т. е. метод наименьших квадратов обладает сглаживающими свойствами.

2.8. Постановка задачи численного дифференцирования

Задача численного дифференцирования возникает при нахождении производных от функции $y = f(x)$, заданной таблично, либо при нахождении производной от аналитической функции, непосредственное дифференцирование которой по каким-либо причинам затруднено. Важное приложение численного дифференцирования — разностная аппроксимация производных, которая широко используется при численном решении обыкновенных дифференциальных уравнений и дифференциальных уравнений в частных производных.

Общий подход к задаче численного дифференцирования состоит в замене функции $y = f(x)$ некоторой приближающей, легко вычисляемой функцией $\phi(x, \bar{a})$, $y = \phi(x, \bar{a}) + R(x)$, где $R(x)$ — остаточный член приближения. Полагают, что $y'(x) = \phi'(x, \bar{a}) + R'(x)$, а $R'(x) = y'(x) - \phi'(x, \bar{a})$ — погрешность вычисления производной.

Наиболее исследованным и широко распространенным является случай, когда в качестве приближающей функции $\varphi(x, \bar{a})$ берется интерполяционный многочлен. При этом производные соответствующих порядков легко определяются дифференцированием многочлена, а погрешность дифференцирования — дифференцированием остаточного члена или погрешности интерполяции, которая, согласно формуле (2.18), равна

$$R_n(x) = \varepsilon_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!} \omega_{n+1}(x).$$

2.9. Дифференцирование интерполяционного многочлена Ньютона

Для вывода формул численного дифференцирования удобно пользоваться интерполяционным многочленом в форме Ньютона, так как он содержит разделившие разности, являющиеся аналогом производных соответствующих порядков. Очевидно, что минимальное число узлов для получения k -й производной равно $(k+1)$, так как дальнейшее дифференцирование многочлена приводит к производной, равной нулю.

Представим функцию $f(x)$ в виде

$$f(x) = P_n(x) + R_n(x). \quad (2.36)$$

Тогда

$$\begin{aligned} f'(x) &= P'_n(x) + R'_n(x), \\ f''(x) &= P''_n(x) + R''_n(x), \\ \dots &\dots \\ f^{(k)}(x) &= P_n^{(k)}(x) + R_n^{(k)}(x). \end{aligned} \quad (2.37)$$

Для равноотстоящих узлов с шагом h имеем:

$$R_n^{(k)}(x) = O(h^{n+1-k}).$$

Таким образом, точность расчета производной при заданной степени интерполяционного многочлена уменьшается с увеличением номера производной. При решении практических задач, в том числе при численном решении обыкновенных дифференци-

альных уравнений и уравнений в частных производных, приходится использовать аппроксимации первых и вторых производных. Значительно реже приходится аппроксимировать производные более высоких порядков.

В связи с этим проведено вычисление первой и второй производной с использованием интерполяционных многочленов Ньютона первой и второй степени.

| Первая производная

Из формул (2.15) и (2.16) с использованием выражений (2.18), (2.36) и (2.37) получаем, дифференцируя по x ,

$$f'(x) = P'_1(x) + R'_1(x) = \frac{f_0 - f_1}{x_0 - x_1} + [2x - (x_0 + x_1)] \frac{f''(\xi)}{2!}, \quad (2.38)$$

$$f'(x) = P'_2(x) + R'_2(x) = \frac{f_0 - f_1}{x_0 - x_1} +$$

$$+ [2x - (x_0 + x_1)] \frac{\frac{f_0 - f_1}{x_0 - x_1} - \frac{f_1 - f_2}{x_1 - x_2}}{x_0 - x_2} + [(x - x_0)(x - x_1) +$$

$$+ (x - x_1)(x - x_2) + (x - x_0)(x - x_2)] \frac{f'''(\xi)}{3!}. \quad (2.39)$$

Формулы (2.38) и (2.39) дают значения первой производной в случае неравноотстоящих узлов, при этом формула (2.38) получена дифференцированием интерполяционного многочлена первой степени, а (2.39) — многочлена второй степени с учетом формул для погрешностей интерполяции. Рассмотрим более подробно эти формулы. Очевидно, что при таблично заданных функциях с использованием интерполяционных многочленов производные вычисляются как производные от интерполяционных многочленов соответствующих порядков. Приближенное значение первой производной по формуле (2.38), полученное дифференцированием $P'_1(x)$, равно

$$f'(x) \approx P'_1(x) = \frac{f_0 - f_1}{x_0 - x_1} = \text{const.} \quad (2.40)$$

При вычислении по формуле (2.40) производной максимальная ошибка, имеющая место в крайних точках интервала (x_0, x_1) , согласно (2.38), пропорциональна первой степени шага $h = (x_0 - x_1)$. В средней точке $\bar{x}_1 = (x_0 + x_1)/2$, производная погрешности $R'_1(\bar{x}_1) = 0$, и значение производной вычисляется с бо-

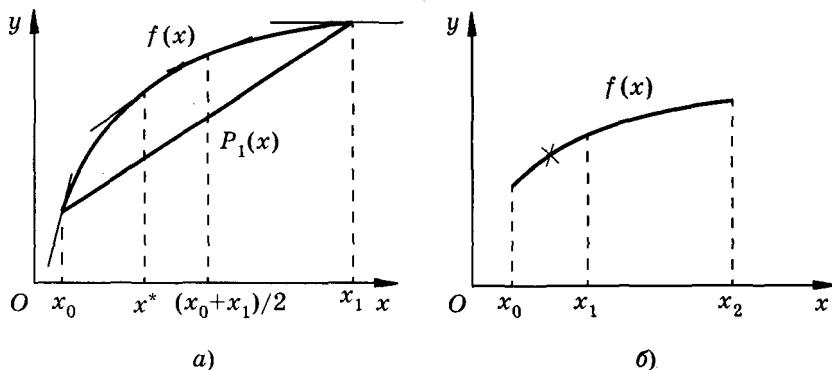


Рис. 2.4

лее высокой точностью, чем в крайних узлах. Естественно, дело не в том, что в точке \bar{x}_1 получается точное значение производной, а в том, что выражение для производной погрешности $R'_n(x)$ является приближенным. Погрешности в точках x_0 и x_1

$$R'_1(x_0) = -(x_1 - x_0) \frac{f''(\xi)}{2!}, \quad R'_1(x_1) = (x_1 - x_0) \frac{f''(\xi)}{2!}$$

равны по величине и противоположны по знаку.

Геометрическое представление формул (2.38), (2.39) и аппроксимация производной показаны на рис. 2.4, а. Видно, что в точке x_0 угол наклона касательной к графику функции $f(x)$ больше, а в точке x_1 меньше, чем угол наклона прямой $P_1(x)$. В некоторой точке x^* точное значение производной и производная $P'_1(x)$ совпадают, и эта точка близка к точке $(x_0 + x_1)/2$.

При использовании интерполяционного многочлена второй степени имеем

$$f'(x) \approx \frac{f_0 - f_1}{x_0 - x_1} + [2x - (x_0 + x_1)] \frac{\frac{f_0 - f_1}{x_0 - x_1} - \frac{f_1 - f_2}{x_1 - x_2}}{x_0 - x_2}. \quad (2.41)$$

Согласно выражению (2.41) производная $f'(x)$ не является величиной постоянной на отрезке $[x_0, x_2]$. Из формул (2.41) и (2.39) следует, что в этом случае при почти равноотстоящих узлах максимальная ошибка в вычислении производной имеет место в крайних точках, а минимальная — в средней, при этом ошибка пропорциональна произведению двучлена $(x_0 - x_1)(x_0 - x_2)$. Если узлы существенно неравноотстоящие, то ошибка может быть пропорциональна первой степени разности $(x - x_i)$.

Если $x_2 \gg (x_1 - x_0)$, то значение функции $f(x)$, полученное в результате интерполяции на отрезке $[x_0, x_1]$, будет иметь погрешность вычисления производной, пропорциональную первой степени шага $(x_0 - x_1)$, а не второй.

Вторая производная

Для вычисления второй производной, очевидно, нужно использовать интерполяционный многочлен Ньютона второй степени. Имеем

$$f''(x) = P_2''(x) + R_2''(x) = \frac{2\left[\frac{f_0 - f_1}{x_0 - x_1} - \frac{f_1 - f_2}{x_1 - x_2}\right]}{x_0 - x_2} + \\ + (3x - x_0 - x_1 - x_2) \frac{f'''(\xi)}{3!}. \quad (2.42)$$

Таким образом, приближенное значение второй производной на отрезке $[x_0, x_2]$ является постоянной, равной

$$f''(x) \approx P_2''(x) = \frac{2\left[\frac{f_0 - f_1}{x_0 - x_1} - \frac{f_1 - f_2}{x_1 - x_2}\right]}{x_0 - x_2}. \quad (2.42a)$$

Из формул (2.42) и (2.42a) следует, что по-прежнему максимальная погрешность в вычислении имеет место в крайних точках и формально по (2.42) равна нулю в точке $\bar{x}_2 = (x_0 + x_1 + x_2)/3$. Однако в связи с приближенностью формулы (2.18) максимальная погрешность не строго равна нулю в этой точке. Отметим, что точки $x_k = (x_0 + x_1 + \dots + x_k)/(k+1)$ называются точками повышенной точности для производной k -го порядка.

2.10. Безразностные формулы численного дифференцирования для равноотстоящих узлов

Если принять узлы равноотстоящими и шаг сетки равным h , то из формул (2.38), (2.39) можно получить выражения для вычисления первой производной в узлах. По формуле (2.38) в точках x_0 и $x_1 = x_0 + h$ производные равны

$$f'(x_0) \approx \frac{f_0 - f_1}{x_0 - x_1} - h \frac{f''(\xi)}{2!}, \quad (2.43a)$$

$$f'(x_1) \approx \frac{f_0 - f_1}{x_0 - x_1} + h \frac{f''(\xi)}{2!}. \quad (2.43b)$$

По формуле (2.39) в точках x_0, x_1, x_2

$$f'(x_0) \approx \frac{1}{2h} [-3f_0 + 4f_1 - f_2] + \frac{h^2}{3} f'''(\xi), \quad (2.44a)$$

$$f'(x_1) \approx \frac{1}{2h} [f_2 - f_0] - \frac{h^2}{6} f'''(\xi)', \quad (2.44b)$$

$$f'(x_2) \approx \frac{1}{2h} [f_0 - 4f_1 + 3f_2] + \frac{h^2}{3} f'''(\xi). \quad (2.44b)$$

Вторые производные в точках x_0, x_1, x_2 по формуле (2.42) равны

$$f''(x_0) \approx \frac{1}{h^2} (f_0 - 2f_1 + f_2) - hf'''(\xi), \quad (2.45a)$$

$$f''(x_1) \approx \frac{1}{h^2} (f_0 - 2f_1 + f_2) - \frac{h^2}{12} f^{IV}(\xi), \quad (2.45b)$$

$$f''(x_2) \approx \frac{1}{h^2} (f_0 - 2f_1 + f_2) + hf'''(\xi). \quad (2.45b)$$

Говорят, что формулы численного дифференцирования (2.43а), (2.43б), (2.45а), (2.45б) — первого порядка точности относительно h , а формулы (2.44а)–(2.44в) и (2.45б) — второго порядка точности. Формулы (2.43а)–(2.45б) носят название БЕЗРАЗНОСТНЫХ ФОРМУЛ ЧИСЛЕННОГО ДИФФЕРЕНЦИРОВАНИЯ. При использовании формул (2.43а), (2.43б) первая производная в узлах x_0, x_1 аппроксимируется с первым порядком точности, а при использовании формул (2.44а)–(2.44в) — со вторым порядком точности.

Вторая производная по формулам (2.45а)–(2.45б) в крайних точках x_0 и x_2 аппроксимируется с первым порядком точности, а в центральной x_1 — со вторым. В центральных точках производные аппроксимируются более точно, чем в крайних, что видно из рис. 2.4 и следует также из приведенных формул. Аппроксимация производных в крайних точках используется при численном решении задачи Коши и краевых задач для обыкновенных дифференциальных уравнений и уравнений в частных производных.

Для повышения точности аппроксимации производных теоретически следует увеличивать степень интерполяционного многочлена.

2.11. Применение ряда Тейлора для численного дифференцирования

На практике формулы численного дифференцирования находят с использованием ряда Тейлора. Разлагая функцию $f(x)$ в ряд Тейлора по степеням h , где h достаточно мало, имеем:

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!} f''(x) + \frac{h^3}{3!} f'''(x) + \frac{h^4}{4!} f^{IV}(x) + \dots, \quad (2.46a)$$

$$f(x-h) = f(x) - hf'(x) + \frac{h^2}{2!} f''(x) - \frac{h^3}{3!} f'''(x) + \frac{h^4}{4!} f^{IV}(x) + \dots. \quad (2.46b)$$

По формулам (2.46a), (2.46b)

$$f'(x) \approx \frac{f(x+h) - f(x)}{h} - \frac{h}{2!} f''(x), \quad (2.47a)$$

$$f'(x) \approx \frac{f(x) - f(x-h)}{h} + \frac{h}{2!} f''(x). \quad (2.47b)$$

Вычитая соотношение (2.46b) из (2.46a), получаем

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h} - \frac{h^2}{6} f'''(x). \quad (2.48)$$

Очевидно, что формулы для первой производной (2.43a) и (2.47a), (2.43b) и (2.47b), (2.44b) и (2.48) представляют собой попарно эквивалентные записи. При сравнении этих формул нужно полагать $x = x_0$ в (2.47a) и $x = x_1$ в (2.47b) и (2.48).

В теории разностных схем производную, вычисленную по формуле (2.47a), называют *односторонней правой производной* или *производной вперед*, вычисленную по формуле (2.47b) — *односторонней левой производной* или *производной назад*, а по формуле (2.48) — *центральной* или *симметричной производной*.

Складывая соотношения (2.46а) и (2.46б), получаем для второй производной:

$$f''(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} - \frac{h^2}{12} f^{IV}(x). \quad (2.49)$$

Очевидно, что формулы (2.45б) и (2.49) эквивалентны, если в (2.49) полагать $x = x_1$.

2.12. Постановка задачи численного интегрирования

В тех случаях, когда при вычислении определенного интеграла

$$F = \int_a^b f(x) dx \quad (2.50)$$

невозможно найти первообразную или она очень сложна для вычислений, то обращаются к формулам численного интегрирования. При этом подынтегральную функцию $f(x)$ заменяют такой приближающей, что она, во-первых, близка в каком-то смысле к $f(x)$:

$$f(x) \approx \varphi(x, \bar{a}),$$

а, во-вторых, интеграл от $\varphi(x, \bar{a})$ легко вычисляется. Чаще всего подынтегральную функцию заменяют интерполяционным многочленом Лагранжа

$$f(x) = L_n(x) + R_n(x), \quad (2.51)$$

в котором $R_n(x)$ — остаточный член интерполяции. Подставляя этот многочлен (2.51) в (2.50), получаем

$$F = \int_a^b L_n(x) dx + R,$$

где $R = \int_a^b R_n(x) dx$ — остаточный член формулы численного интегрирования или ее погрешность.

Заменяя $f(x)$ интерполяционными многочленами различной степени, получают формулы численного интегрирования различного порядка точности.

2.13. Формулы численного интегрирования

Заменим подынтегральную функцию, входящую в (2.50), интерполяционным многочленом Лагранжа нулевой степени, проходящим через середину отрезка — точку $\bar{x} = (a + b)/2$ (рис. 2.5). Площадь криволинейной трапеции можно заменить площадью прямоугольника, т. е.

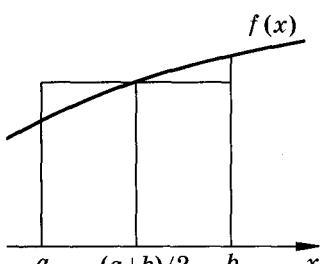


Рис. 2.5

$$F = \int_a^b f(x) dx \approx (b - a)f\left(\frac{a + b}{2}\right). \quad (2.52)$$

Формула (2.52) носит название **ФОРМУЛЫ ПРЯМОУГОЛЬНИКОВ** или **ФОРМУЛЫ СРЕДНИХ**. Ее погрешность составляет

$$R = \int_a^b f(x) dx - (b - a)f(\bar{x}). \quad (2.53)$$

Разложение функции $f(x)$ в ряд относительно середины отрезка имеет вид

$$\begin{aligned} f(x) &= f(\bar{x} - (\bar{x} - x)) = \\ &= f(\bar{x}) - (x - \bar{x})f'(\bar{x}) + \frac{(x - \bar{x})^2}{2!} f''(\bar{x}) + \dots . \end{aligned} \quad (2.54)$$

Подставив выражение (2.54) в (2.53), получим

$$\begin{aligned} R &\approx \int_a^b \left[f(\bar{x}) - (x - \bar{x})f'(\bar{x}) + \frac{(x - \bar{x})^2}{2!} f''(\bar{x}) \right] dx - (b - a)f(\bar{x}) = \\ &= f(\bar{x})(b - a) + \frac{1}{2} f'(\bar{x}) \left[\left(\frac{a - b}{2} \right)^2 - \left(\frac{b - a}{2} \right)^2 \right] + \\ &+ \frac{1}{6} f''(\bar{x}) \left[\left(\frac{b - a}{2} \right)^3 - \left(\frac{a - b}{2} \right)^3 \right] - (b - a)f(\bar{x}) = \\ &= \frac{1}{24} f''(\bar{x})(b - a)^3. \end{aligned} \quad (2.55)$$

При вычислении ошибки интегрирования уничтожился не только первый, но и второй член разложения, что связано с симметричным выбором узла интегрирования. И хотя по построению формула точна для многочленов нулевого порядка, выбор симметричного узла интерполяции привел к тому, что формула точна для любой линейной функции.

Значение остаточного члена в формуле прямоугольников (2.53) может быть велико, так как разность $(b - a)$ может быть достаточно большой. Для повышения точности введем сетку

$$a = x_0 \leq x_1 \leq \dots \leq x_N = b$$

с достаточно мелким шагом $h_i = x_i - x_{i-1}$ и применим формулу прямоугольников на каждом шаге сетки. Тогда получим обобщенную формулу прямоугольников

$$\int_a^b f(x) dx \approx \sum_{i=1}^N h_i f\left(\frac{x_{i-1} + x_i}{2}\right) \quad (2.56)$$

с величиной остаточного члена

$$R \approx \frac{1}{24} \sum_{i=1}^N h_i^3 f''\left(\frac{x_{i-1} + x_i}{2}\right).$$

На равномерной сетке с шагом $h_i = x_i - x_{i-1} = \text{const}$ формула (2.56) упрощается и имеет вид

$$\int_a^b f(x) dx \approx h \sum_{i=1}^N f\left(\frac{x_{i-1} + x_i}{2}\right), \quad (2.57)$$

величина остаточного члена составляет

$$R \approx \frac{1}{24} h^3 \sum_{i=1}^N f''\left(\frac{x_{i-1} + x_i}{2}\right). \quad (2.58)$$

Заменяя в (2.58) сумму интегралом, получаем

$$R \approx \frac{1}{24} h^2 \int_a^b f''(x) dx. \quad (2.59)$$

Для справедливости оценки остаточного члена (2.58) необходимо существование непрерывной второй производной; если вторая производная $f''(x)$ — кусочно-непрерывная, то удается сделать лишь мажорантную оценку, заменяя $f''(x)$ ее максимальной величиной на $[a, b]$. Тогда, если обозначить $M_2 = \max |f''(x)|_{[a, b]}$, остаточный член

$$R \leq \frac{1}{24} h^2 M_2 (b - a). \quad (2.60)$$

В том случае, когда функция $f(x)$ задана в виде таблицы, ее значение в середине интервала неизвестно. Это значение находится, как правило, интерполяцией, что приводит к ухудшению точности формулы.

В случае таблично заданных функций удобно в качестве узлов интерполяции выбрать начало и конец отрезка интегрирования, т. е. заменить функцию $f(x)$ многочленом Лагранжа первой степени. Имеем

$$\int_a^b f(x) dx = \int_a^b L_1(x) dx + \int_a^b R_1(x) dx.$$

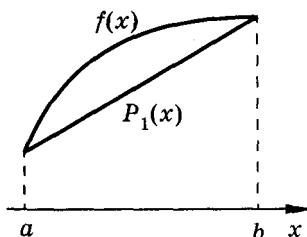


Рис. 2.6

В этом случае величина интеграла, равная площади криволинейной трапеции, приближенно заменяется величиной площади трапеции (рис. 2.6). Поэтому получаем

$$\begin{aligned} \int_a^b f(x) dx &\approx \frac{1}{2}(b-a)[f(a) + f(b)] = \\ &= \int_a^b L_1(x) dx, \end{aligned} \quad (2.61)$$

имея в виду, что $x_0 = a$, $x_1 = b$. Эта формула носит название ФОРМУЛЫ ТРАПЕЦИЙ. При использовании формулы трапеций для оценки погрешности интегрирования вычислим $\int_a^b R_1(x) dx$ по формулам (2.18). Имеем

$$R = \int_a^b R_1(x) dx = -\frac{1}{12}(b-a)^3 f''(\xi). \quad (2.62)$$

Погрешность формулы трапеций вдвое больше погрешности формулы прямоугольников. Это объясняется тем, что выбор в формуле прямоугольников в качестве узла интерполяции симметричного узла приводит к повышению ее точности.

Для повышения точности формулы (2.61) введем на отрезке $[a, b]$ сетку

$$a = x_0 \leq x_1 \leq \dots \leq x_N = b.$$

Подсчитывая значение интеграла для каждого интервала и суммируя эти значения, получаем обобщенную формулу трапеций

$$F(x) = \int_a^b f(x) dx \approx \frac{1}{2} \sum_{i=1}^N (f_{i+1} + f_i)(x_{i+1} - x_i) \quad (2.63)$$

со значением остаточного члена

$$R \approx -\frac{1}{12} \sum_{i=1}^N (x_i - x_{i-1})^3 f''\left(\frac{x_{i+1} + x_i}{2}\right).$$

Эти формулы упрощаются на сетке с постоянным шагом $h = h_i = x_i - x_{i-1} = \text{const}$ ($i = 0, 1, \dots, N - 1$):

$$F = \int_a^b f(x) dx \approx h \left[\frac{1}{2} (f_0 + f_N) + \sum_{i=1}^{N-1} f_i \right], \quad (2.64)$$

$$R \approx -\frac{1}{12} h^3 \sum_{i=1}^N f''\left(\frac{x_i + x_{i-1}}{2}\right) \approx -\frac{1}{12} h^2 \int_a^b f''(x) dx. \quad (2.65)$$

Введем обозначение $M_2 = \max |f''(x)|_{[a, b]}$. На практике пользуются мажорантной оценкой величины остаточного члена

$$|R| \leq \frac{b-a}{12} h^2 M_2. \quad (2.66)$$

Таким образом, формула трапеций (как и формула прямоугольников) имеет второй порядок точности относительно шага сетки, и погрешность асимптотически стремится к нулю при $h \rightarrow 0$ с точностью до членов более высокого порядка малости.

Для повышения порядка точности формулы численного интегрирования заменим подынтегральную кривую параболой — интерполяционным многочленом Лагранжа второй степени, выбрав в качестве узлов интерполяции концы и середину отрезка интегрирования: $x_0 = a$, $x_1 = (a+b)/2$, $x_2 = b$ (рис. 2.7).

В этом случае, проинтегрировав интерполяционный многочлен для равноотстоящих узлов, получим

$$F \approx \int_a^b L_2(x) dx = \frac{h}{3} [f_0 + 4f_1 + f_2]. \quad (2.67)$$

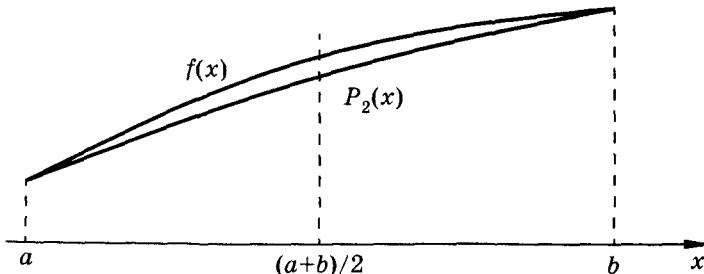


Рис. 2.7

При этом значение остаточного члена $R = \int_a^b R_2(x) dx$ оценивается приближенным соотношением

$$R \approx -\frac{(b-a)^5}{180} f^{IV}\left(\frac{a+b}{2}\right).$$

Формулу (2.67) называют ФОРМУЛОЙ СИМПСОНА. Для неравноотстоящих узлов x_0, x_1, x_2 величина F составляет

$$F \approx \int_a^b L_2(x) dx.$$

Как и в предыдущих двух случаях, для повышения точности формулы (2.67) введем сетку с достаточно малым шагом. Суммируя значения интегралов, полученных по (2.67) для каждого интервала, получаем обобщенную формулу Симпсона (парабол), которая на равномерной сетке имеет вид

$$F \approx \frac{h}{3} (f_0 + 4f_1 + 2f_2 + 4f_3 + 2f_4 + \dots + 2f_{N-2} + 4f_{N-1} + f_N), \quad (2.68)$$

а величина остаточного члена —

$$R \approx -\frac{h^4}{180} \int_a^b f^{IV}(x) dx. \quad (2.69)$$

Таким образом, формула парабол имеет четвертый порядок точности относительно шага сетки. Введем обозначение $M_4 = \max |f^{IV}(x)|_{[a, b]}$. Как правило, для оценки величины погрешности применяют мажорантную оценку

$$|R| \leq \frac{(b-a)}{180} h^4 M_4. \quad (2.70)$$

Количество интервалов N , на которое делится отрезок интегрирования, должно быть всегда четным, так как интерполяционный многочлен второй степени проводится через три точки.

СИМПСОН ТОМАС (Simpson Thomas; 1710—1761) — английский математик, ряд работ которого посвящен элементарной геометрии, тригонометрическому анализу и математическому анализу. В 1743 г. С. вывел формулу приближенного интегрирования (формулу Симпсона).

2.14. Метод Рунге – Ромберга – Ричардсона повышения порядков точности

В вычислениях с помощью формул численного интегрирования результат получается тем точнее, чем с большим числом узлов используется сетка. Однако при этом формулы оказываются достаточно громоздкими. Метод Рунге – Ромберга – Ричардсона позволяет получать более высокий порядок точности без значительного увеличения числа операций.

Пусть для вычисления величины $f(x)$ имеется некоторая функция $\varphi(x, h)$, дающая возможность приближенного расчета величины $f(x)$ на равномерной сетке с шагом h . Значение остаточного члена можно представить в виде

$$f(x) - \varphi(x, h) = \psi(x)h^p + O(h^{p+1}), \quad (2.71)$$

где p — порядок точности расчетной формулы, а $\psi(x)h^p$ — главный член погрешности. Например, для формулы прямоугольников и трапеций $p = 2$, а для формулы парабол $p = 4$.

Если провести расчет на другой сетке с шагом rh , приближенное значение $\varphi(x, rh)$ величины $f(x)$ окажется другим. В этом случае остаточный член примет вид

$$f(x) - \varphi(x, rh) = \psi(x)(rh)^p + O((rh)^{p+1}). \quad (2.72)$$

Значение главного члена погрешности определяется вычитанием выражения (2.72) из выражения (2.71):

$$\psi(x)h^p = \frac{\varphi(x, h) - \varphi(x, rh)}{r^p - 1} + O(h^{p+1}). \quad (2.73)$$

Следовательно, пользуясь расчетом на сетке с шагом rh , удается оценить главный член погрешности расчета на равномерной

РИЧАРДСОН ЛЬЮИС ФРАЙ (Richardson Lewis Fry; 1881–1953) — английский математик, занимавшийся прикладной математикой, в частности, решением уравнений в конечных разностях (дискретными конечномерными аналогами дифференциальных уравнений). Р. был первым ученым, предложившим использовать математические методы (в особенности метод конечных разностей) в предсказании погоды.

сетке с шагом h . Если же подставить найденную погрешность (2.73) в (2.72), то получим результат с более высокой точностью:

$$f(x) = \varphi(x, h) + \frac{\varphi(x, h) - \varphi(x, rh)}{r^p - 1} + O(h^{p+1}). \quad (2.74)$$

Метод Рунге — Ромберга — Ричардсона может быть использован для оценки погрешности расчета и получения результата с более высокой степенью точности при решении различных задач, когда возможен расчет какой-либо величины на сетках с различным шагом, например, при численном интегрировании, дифференцировании, решении дифференциальных уравнений.

В качестве примера рассмотрим полученные формулы численного интегрирования, имеющие более высокий порядок точности. Проведем расчет определенного интеграла по формуле трапеций с шагом $2h$ (точки x_0, x_2), а затем на сетке с шагом h (точки x_0, x_1, x_2). Обозначим результаты этих расчетов через F_{2h} и F_h соответственно. Порядок точности формулы трапеций равен двум, а сам остаточный член имеет такой же вид, как и в формуле (2.71). Обозначая более точное значение через F , проведем уточнение по формуле (2.74):

$$\begin{aligned} F &= F_{2h} + \frac{F_{2h} - F_h}{(1/2)^2 - 1} = \frac{1}{3}[4F_h - F_{2h}] = \\ &= \frac{1}{3}\left[4\frac{h}{2}[f_0 + 2f_1 + f_2] - \frac{2h}{2}[f_0 + f_2]\right] = \frac{h}{3}[f_0 + 4f_1 + f_2]. \end{aligned}$$

Таким образом, уточнение расчета формулы трапеций с использованием выражения (2.74) привело к формуле парабол, имеющей более высокий — четвертый — порядок точности.

2.15. Метод статистических испытаний

При изучении численных методов основной задачей являлось преобразование известной математической модели к виду, допускающему эффективное численное решение, т. е. сведение всех математических понятий (таких, как производная, интеграл, дифференциальное уравнение и т. п.) к последовательности эле-

ментарных арифметических операций. Для дальнейшего существенно, что вместе с алгоритмом метода мы всегда получали так называемую гарантированную оценку погрешности метода. Алгоритм численного метода считался состоятельным лишь в том случае, если малые погрешности, содержащиеся в исходных данных и внесенные в процессе преобразований или вычислений, не влияли неприемлемым образом на результат вычислений. Процесс отыскания решения при таком подходе являлся строго детерминированным, т. е. при безошибочном повторении неизменно приводил к одному и тому же результату.

Однако существует целый ряд задач, для которых разработка и использование детерминистских методов оказываются практически невозможными и нецелесообразными. Примером такой математической задачи может быть задача о вычислении интеграла кратности S . Так, для класса функций $G_1, \dots, _1(M, \dots, M)$, имеющих кусочно-непрерывные первые производные, ограниченные по модулю постоянной M , не существует методов с оценкой погрешности лучшей, чем величина $aM_n^{-1/S}$, где n — число узлов интегрирования, a — некоторая постоянная. Если требуется гарантировать оценку погрешности $0,01aM$, то число узлов интегрирования должно удовлетворять условию $n > 100^S$. При вычислении интеграла кратности $S = 5$ это означает, что требуется более десяти миллиардов узлов. Задачи такого типа не являются экзотическими.

Например, решение задачи о проникновении частиц через толстый экран требует вычисления $S = 3 \cdot k$ кратных интегралов, где k — количество столкновений при прохождении частиц через экран. Для $k = 4$ это требует выполнения $2,4 \cdot 10^{14}$ операций.

Другой проблемой, возникающей при решении практических задач, является то обстоятельство, что на практике часто возникают задачи, для которых формулировка математической модели в виде системы уравнений, допускающей численное решение, представляет значительные трудности или вообще оказывается невозможной. Такая ситуация является обычной при изучении сложных физических явлений или реальных сложных систем, состоящих из большого числа различных по природе и сложным образом взаимодействующих элементов, функционирующих в условиях воздействия большого числа случайных факторов. К сложным системам относятся крупные технологические и производственные комплексы, экономические и социологические объекты, многопроцессорные вычислительные системы и т. п. Одним из существенных признаков сложных систем является принципиально вероятностный характер их функционирования в условиях большого числа случайных факторов.

Таким образом, не всегда можно построить детерминистский метод решения задачи (и, в частности, эффективный численный метод), и даже, имея детерминистский метод, не всегда можно практически получить результат с гарантированной оценкой погрешности.

Достаточно общим способом решения таких проблем являются методы статистических испытаний. С помощью методов статистических испытаний математические задачи и задачи исследования сложных систем решаются путем моделирования случайных реализаций имитации случайных процессов, происходящих в сложных системах, с последующей оценкой их вероятностных характеристик. Принято различать два основных класса методов статистических испытаний: численные методы статистических испытаний (методы МОНТЕ-КАРЛО) и методы имитационного моделирования.

Суть ЧИСЛЕННЫХ МЕТОДОВ статистических испытаний* заключается в установлении связи между величинами, являющимися решениями математической задачи (например, значения интеграла), и вероятностными характеристиками некоторого случайного процесса, такими, как вероятность, математическое ожидание и др. Выбор конкретной случайной величины обуславливается особенностями конкретной задачи и используемой процедурой случайных испытаний. Существенной чертой численных методов статистических испытаний является вероятностная оценка погрешности вычислений.

Суть МЕТОДОВ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ заключается в моделировании физических представлений о реальном

* Идея метода статистических испытаний восходит, по-видимому, к Ж. Л. Л. Бюффону, предложившему случайный процесс для вычисления значения числа π (задача Бюффона). Современными создателями метода считают американских математиков Дж. Неймана и С. Улама, использовавших метод для моделирования поведения нейронов.

БЮФФОН ЖОРЖ ЛУИ ЛЕКЛЕРК (Buffon George Louis Leclerc; 1707—1788) — французский естествоиспытатель и инженер. Б. занимался двенадцатеричной системой счисления; он первым исследовал задачи, посвященные геометрической вероятности.

УЛАМ СТАНИСЛАВ МАРЦИН (Ulam Stanislaw Marcin; 1909—1984) — американский математик. Его основные труды относятся к топологии, функциональному анализу, теории вероятностей и теории множеств. В математической физике У. развил метод Монте-Карло (совместно с Дж. фон Нейманом).

явлении или моделировании процесса функционирования сложной системы. При этом практически снимаются ограничения на степень приближения имитационной модели к реальному объекту.

Перечислим основные математические задачи, для которых построены вероятностные модели: вычисление кратных интегралов, решение систем линейных уравнений, обращение матриц, нахождение собственных значений и собственных векторов матриц, решение краевых задач для уравнений эллиптического типа, решение задачи Коши для уравнений параболического типа, решение задачи интерполяции, решение уравнений Больцмана и др.

| Задача Бюффона

Основные идеи численных методов статистических испытаний рассмотрим на примере задачи Бюффона, в которой предложен случайный процесс, позволяющий вычислить число π . Пусть имеется семейство равноотстоящих параллельных прямых с расстоянием между прямыми, равным d , и игла длины $l \leq d$.

При случайном бросании иглы ее средняя точка может упасть на расстоянии x от какой-либо прямой ($0 \leq x \leq d/2$). Угол между игрой и нормалью к направлению прямых обозначим через ϕ (рис. 2.8). Игla пересекается с одной из прямых при условии $x < (l/2)\cos \phi$ для $\phi \in (-\pi/2, \pi/2)$. Геометрически это означает попадание средней точки иглы в заштрихованную на рис. 2.9

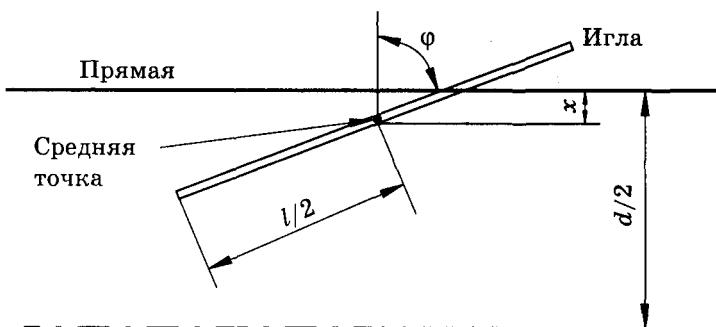


Рис. 2.8

БОЛЬЦМАН ЛЮДВИГ (Boltzmann Ludwig; 1844—1906) — австрийский физик, заложивший основы современной теории излучения. Самыми крупными его работами являются классические исследования по кинетической теории газов и статистическому толкованию 2-го начала термодинамики.

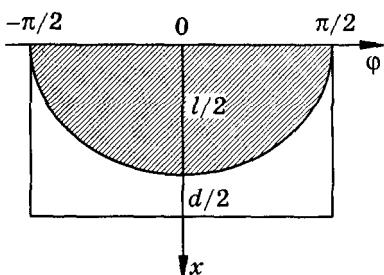


Рис. 2.9

где n — полное число попаданий в прямоугольник возможных точек попадания (полное число бросаний иглы), m — число попаданий в заштрихованную область (число пересечений иглы с одной из параллельных прямых), m/n — относительная частота. Таким образом, приближенное соотношение для вычисления числа π имеет вид

$$\pi \approx 2 \frac{l}{d} \frac{n}{m}.$$

Задача Бюффона интересна во многих отношениях, так как позволяет интуитивно сформулировать основные проблемы и особенности метода статистических испытаний (названного Бюффоном *методом бросания*).

Во-первых, задача наглядно представляет общую схему численных методов статистических испытаний: выбирается некоторая процедура случайного испытания такая, чтобы какая-либо вероятностная характеристика ее равнялась искомой величине или очевидным образом была с ней связана.

Во-вторых, отметим, что в задаче используется совсем не тривиальное понятие *случайного бросания*. В задаче Бюффона нетрудно представить себе неправильно организованное бросание иглы (например, параллельно прямым с *прицеливанием* в середину между прямыми), которое приведет к грубому искажению результата. Далее будут рассмотрены современные методы формирования последовательностей случайных чисел (являющихся ис-

областью. Вероятность попадания иглы в заштрихованную область равна отношению площадей заштрихованной области и всего прямоугольника возможных точек попадания середины иглы, т. е.

$$P = \int_{-\pi/2}^{\pi/2} \frac{(l/2)\cos \phi}{\pi d/2} d\phi = \frac{2l}{\pi d}.$$

По теореме Бернулли

$$P \approx m/n,$$

БЕРНУЛЛИ ЯКОБ старший (Bernoulli Jacob; 1654—1705) — швейцарский ученый, который блестяще применил методы дифференциального исчисления Г. Лейбница к изучению свойств ряда кривых. Совместно со своим братом Иоганом Б. заложил основы вариационного исчисления. В области теории вероятности он доказал теорему Бернулли, относящуюся к закону больших чисел.

ходными для формирования случайных объектов любой природы — случайных событий, величин, процессов) и связанные с этим проблемы. Здесь отметим лишь, что незамеченная неслучайность таких последовательностей может иметь *катастрофические* последствия при использовании метода.

В-третьих, интуитивно ясно, что вычисление числа π с использованием предложенного метода потребует большего числа бросаний для обеспечения вычисления числа π с точностью, например, до двух знаков. При решении практических задач (вычислении интегралов, решении систем линейных уравнений, моделировании сложных систем и др.) длина используемых последовательностей случайных чисел исчисляется десятками и сотнями тысяч и может достигать миллионов чисел. Поэтому практическое использование методов статистических испытаний определяется двумя факторами: наличием простых и экономичных способов формирования последовательностей случайных чисел и возможностью организации эффективных испытаний.

Например, для задачи Бюффона было установлено, что использование двух жестко скрепленных под углом $\pi/2$ игл приводит к уменьшению числа бросаний для достижения одной и той же точности в 12,2 раза (при этом одно брасание креста засчитывается за два брасания иглы). Использование трех жестко скрепленных под углом $\pi/3$ игл приводит к уменьшению числа бросаний в 44,3 раза, использование четырех жестко скрепленных игл — в 107,2 раза. Этот пример показывает, что рациональная организация процедуры испытаний (использование коррелированных случайных величин) может существенно уменьшить объем испытаний для достижения требуемой точности.

Случайные числа. Основные понятия

Случайной называется величина, для которой указано, какие значения она может принимать и каковы вероятности появления этих значений. Случайная величина ξ задается ЗАКОНОМ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ

$$P(\xi < x) = F(x)$$

или для непрерывных случайных величин также ПЛОТНОСТЬЮ ВЕРОЯТНОСТИ (функцией плотности) $f(x)$ такой, что

$$F(x) = \int_{-\infty}^x f(x) dx.$$

Дискретная случайная величина принимает дискретное множество значений, и закон распределения ее задается таблицей

$$\xi = \begin{pmatrix} \xi_1 & \xi_2 & \dots & \xi_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}.$$

Вероятность попадания в интервал (a, b) для непрерывной случайной величины определяется соотношением

$$P(a < \xi < b) = \int_a^b f(x) dx.$$

МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ случайной величины составляет для дискретных случайных величин $M\xi = \sum_{i=1}^n \xi_i p_i$, для непрерывных случайных величин $M\xi = \int_a^b xf(x) dx$.

ДИСПЕРСИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ составляет для дискретных случайных величин $D\xi = M[(\xi - M\xi)^2]$, для непрерывных случайных величин $D\xi = \int_a^b (x - Mx)^2 f(x) dx$.

Для независимых дискретных случайных величин справедливо соотношение

$$D\xi = M(\xi^2) - (M\xi)^2.$$

Математическое ожидание функции случайной величины ξ составляет

$$M\eta = M\varphi(\xi) = \int_a^b \varphi(x)f(x) dx,$$

где $\varphi(x)$ — произвольная функция, $f(x)$ — плотность вероятности случайной величины ξ , η — функция случайной величины ξ , ξ — сама случайная величина.

РАВНОМЕРНО РАСПРЕДЕЛЕННАЯ СЛУЧАЙНАЯ ВЕЛИЧИНА — такая случайная величина, для которой на интервале (a, b) плотность вероятности равна

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{при } x \in (a, b), \\ 0 & \text{при } x \notin (a, b). \end{cases}$$

ФУНКЦИЯ РАСПРЕДЕЛЕНИЯ РАВНОМЕРНО РАСПРЕДЕЛЕННОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ определяется следующим образом:

$$F(x) = \begin{cases} 0 & \text{при } x \in (-\infty, a], \\ \frac{x-a}{b-a} & \text{при } x \in (a, b), \\ 1 & \text{при } x \in [b, \infty). \end{cases}$$

МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ и ДИСПЕРСИЯ РАВНОМЕРНО РАСПРЕДЕЛЕННОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ равны

$$M\xi = \frac{a+b}{2}, \quad D\xi = \frac{(b-a)^2}{12}.$$

Особый интерес представляет равномерно распределенная случайная величина на интервале $(0, 1)$, для которой $M = 1/2$, $D = 1/12$. В этом случае

$$f(x) = \begin{cases} 1 & \text{при } x \in (0, 1), \\ 0 & \text{при } x \notin (0, 1), \end{cases} \quad F(x) = \begin{cases} 0 & \text{при } x \in (-\infty, 0], \\ x & \text{при } x \in (0, 1), \\ 1 & \text{при } x \in [1, \infty). \end{cases}$$

| Способы получения случайных чисел

Для моделирования случайных процессов при решении задач методом статистических испытаний необходимы последовательности из десятков и сотен тысяч и даже миллионов чисел. Для выработки этих последовательностей могут быть использованы так называемые *генераторы случайных чисел* — какие-либо физические устройства, в которых реализуется случайный процесс (игральная кость, рулетка, счетчик Гейгера и т. п.). Сложность технической реализации, неповторяемость результата, неэкономичность и медлительность этих устройств объясняют их малое распространение. Кроме того, при непосредственном использовании генераторов случайных чисел трудно проверить *случайные* свойства вырабатываемых последовательностей, которые могут зависеть от технического состояния устройства.

В основу современных методов получения последовательностей случайных чисел для использования на универсальных вычислительных машинах положены следующие требования: во-первых, для получения используемых компьютерами последовательностей случайных чисел желательно использовать сами компьютеры; во-вторых, количество операций, затрачиваемых на получение значения этого числа, должно быть небольшим; в-третьих, не имеет значения, каким образом получена последовательность случайных чисел, лишь бы она удовлетворяла принятой системе тестов *качества* используемых случайных чисел, которая может зависеть от конкретной задачи. Простейшими тестами являются контроль для каждого разряда частоты появления цифр числа в каждом разряде; вычисление коэффициентов корреляции последовательных разрядов числа $\sum_i (\xi_i - 0,5)(\xi_{i+1} - 0,5)$ (для двоичной системы счисления) или групп разрядов внутри числа, которые должны быть близкими к нулю.

Существенным отклонением от статистической структуры последовательности может быть наличие периодичности. Заметим, что наличие периодичности в последовательностях, генерируемых с помощью компьютеров, неизбежно — это будет пояснено далее. Важно, чтобы период был возможно большим.

Числа, получаемые вычислительной машиной по какому-либо алгоритму, не являются случайными в строгом смысле и их называют *псевдослучайными*. Первый алгоритм, предложенный для получения последовательностей псевдослучайных чисел Дж. Нейманом, называется МЕТОДОМ СЕРЕДИНЫ КВАДРАТОВ. Суть его поясним с помощью примера.

ПРИМЕР. Выберем некоторое четырехзначное число. Возведем его в квадрат и из полученного восьмизначного числа выделим четыре средние цифры. Полученное таким образом четырехзначное число вновь используем для получения следующего числа. При $\xi_0 = 8726$ этот результат даст последовательность чисел, приведенную в таблице 2.2.

Таблица 2.2

i	ξ_i	ξ_{i+1}
0	8726	<u>76</u> <u>1430</u> <u>76</u>
1	1430	<u>02</u> <u>0449</u> <u>00</u>
2	0449	<u>00</u> <u>2016</u> <u>01</u>
3	2016	<u>04</u> <u>0642</u> <u>56</u>
4	0642	<u>00</u> <u>4121</u> <u>64</u>
5	4121	

Процесс порождения чисел последовательности можно записать в виде рекуррентного соотношения $\xi_{i+1} = f(\xi_i)$, в котором

НЕЙМАН ДЖОН (ЯНОШ) ФОН (Neumann John (Janos) von; 1903—1957) — американский математик и инициатор использования ЭВМ при решении различных научных проблем. Работая до 1930 г. в Европе, фон Н. занимался различными разделами математики. Переехав перед Второй мировой войной в США, Н. обратился к вопросу взаимодействия ударных волн (фон Н. принимал участие в работах по созданию первой атомной бомбы). Эта проблема потребовала огромной вычислительной работы, что и стало причиной интереса фон Н. к ЭВМ. В основу проекта построенного им электронного вычислителя были положены особенности работы человеческого мозга.

функция f осуществляет отображение множества чисел на самое себя.

Использование и тщательное исследование метода середины квадратов обнаружило его существенные недостатки: получаемое распределение чисел оказалось недостаточно равномерным, в нем преобладают малые значения; в методе возможно полное вырождение последовательности, когда некоторое значение ξ_i и все последующие обращаются в нуль.

Использование рекуррентных соотношений (указанного или более общего вида) для получения последовательностей псевдослучайных чисел является в настоящее время наиболее распространенным. Приведем краткое описание некоторых из предложенных функций.

Один из алгоритмов связан с выделением дробной части произведения $\xi_{i+1} = \{A\xi_i\}$, в котором A — постоянная, а фигурные скобки обозначают выделение дробной части произведения. Качество генерируемых последовательностей псевдослучайных чисел существенно зависит от выбора значений A и ξ_0 .

Все современные генераторы псевдослучайных чисел делятся на три группы.

- Первая группа содержит генераторы с линейной рекуррентностью

$$Z_i = \sum_{s=1}^p A_s Z_{i-s} \pmod{M},$$

в которых A_s — неотрицательные целые числа в диапазоне от 0 до $(M - 1)$. Введем обозначение $K_i = \left\{ M^{-1} \sum_{s=1}^p A_s Z_{i-s} \right\}$. В качестве вычислительной формулы для псевдослучайных чисел Z_i используют соотношение

$$Z_i = \sum_{s=1}^p A_s Z_{i-s} - K_i M.$$

- Вторая группа состоит из генераторов с линейной рекуррентностью по модулю 2

$$B_j = \sum_{s=1}^p a_s B_{j-s} \pmod{2};$$

здесь $a_s \in \{0, 1\}$, $a_p = 1$, $B_j \in \{0, 1\}$, и псевдослучайное число Z_i имеет вид

$$Z_i = \sum_{j=1}^m 2^{m-j+1} B_{ri+j}, \quad r \geq m.$$

- К третьей группе относятся нелинейные конгруэнтные генераторы псевдослучайных чисел

$$Z_i = f(Z_{i-1}) \pmod{M},$$

в которых в качестве функции f принимается нелинейная целая функция.

Получение случайных чисел
с заданным законом распределения

При решении практических задач требуется получать последовательности случайных объектов различной природы и с различными законами распределения вероятностей. В настоящее время принята следующая схема получения таких последовательностей:

- В качестве исходной рассматривается последовательность случайных чисел с равномерным распределением в интервале $[0, 1]$. Эта последовательность может быть получена на компьютере с наименьшими затратами машинного времени, обеспечивает простоту и удобство дальнейших преобразований, позволяет относительно просто проверить *качество* последовательности.
- Путем функциональных преобразований конструируются случайные последовательности с практически любым законом распределения.

Идея такого подхода может быть проиллюстрирована простым примером. Пусть имеется дискретная случайная величина с законом распределения, заданным таблицей

$$\eta = \begin{pmatrix} \eta_1 & \eta_2 & \dots & \eta_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}.$$

Интервал $(0, 1)$ разобьем на n частей так, чтобы длина каждого из отрезков была равна p_i .

Имея в качестве исходной равномерно распределенную на интервале $[0, 1]$ случайную величину ξ (рис. 2.10), поступаем следующим образом. Выбирая из последовательности значение ξ , фиксируем, в какой интервал попадает это значение. В зависимости от этого полагаем значение случайной величины $\eta = \eta_i$.



Рис. 2.10

Такой метод формирования (разыгрывания) случайных величин для непрерывных случайных величин носит название *метода обратной функции*. Суть метода заключается в следую-

щем. Пусть непрерывная случайная величина η задана законом распределения

$$F(\eta) = \int_{-\infty}^{\eta} f(y) dy,$$

где $f(\eta)$ — плотность распределения вероятностей. В соответствии со свойствами функции распределения (изменяется от 0 до 1, непрерывна и монотонна) уравнение

$$\xi = F(\eta) \quad (\alpha)$$

имеет единственное решение, которое может быть принято за значение η . Рассмотрим график функции $F(\eta)$ (рис. 2.11).

Если случайная величина η , полученная из уравнения (α), т. е. $\eta = F^{-1}(\xi)$, лежит в интервале (a, b) , то очевидно, что $\xi \in (F(a), F(b))$ и наоборот. Вероятность того, что $\eta \in (a, b)$, равна вероятности того, что $\xi \in (F(a), F(b))$, т. е.

$$P(a < \eta < b) = P(F(a) < \xi < F(b)).$$

ПРИМЕР. Пусть требуется получить последовательность случайных чисел с показательным законом распределения

$$F(\eta) = \int_0^{\eta} \lambda e^{-\lambda y} dy, \quad \eta > 0.$$

Для отыскания η_i запишем уравнение

$$\xi_i = \lambda \int_0^{\eta_i} e^{-\lambda \eta} d\eta,$$

в котором ξ_i — случайное число последовательности с равномерным распределением на интервале $(0, 1)$. После интегрирования получим $\xi_i = 1 - e^{-\lambda \eta_i}$.

Разрешая это соотношение относительно η_i , придем к выражению

$$\eta_i = -\frac{1}{\lambda} \ln(1 - \xi_i)$$

для вычисления значений случайной величины η_i с показательным законом распределения по значениям случайной величины ξ_i , с равномерным законом распределения на интервале $(0, 1)$.

| Теоретические основы
метода статистических испытаний

Теоретическую основу метода статистических испытаний составляют фундаментальные теоремы теории вероятностей —

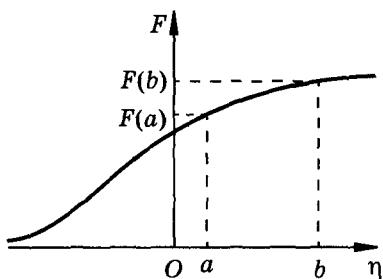


Рис. 2.11

теорема Чебышёва, центральная предельная теорема и теорема Бернулли.

ТЕОРЕМА ЧЕБЫШЁВА. Пусть ξ_1, \dots, ξ_n — независимые случайные величины с математическим ожиданием $M\xi$ и с равномерно ограниченной дисперсией. Тогда для любого $\epsilon > 0$

$$\lim_{n \rightarrow \infty} P\left(\left| \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n} - M\xi \right| < \epsilon \right) = 1.$$

В некоторых случаях будем пользоваться записью теоремы Чебышёва для нормального закона распределения

$$P\left(\left| \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n} - M\xi \right| < \alpha \sqrt{\frac{D\xi}{n}} \right) \approx \Phi(\alpha),$$

где $\Phi(\alpha) = \frac{2}{\sqrt{2\pi}} \int_0^\alpha e^{-x^2/2} dx$ — интеграл вероятностей.

Теорема Чебышёва утверждает сходимость по вероятности среднего арифметического n случайных величин к их математическому ожиданию при $n \rightarrow \infty$.

ТЕОРЕМА БЕРНУЛЛИ. Если некоторое событие A может наступить в каждом m из n независимых испытаний с вероятностью P и при этом $P(\bar{A}) = 1 - P$, то для любого $\epsilon > 0$

$$\lim_{n \rightarrow \infty} P\left(\left| \frac{m}{n} - P \right| < \epsilon \right) = 1,$$

где m/n — относительная частота события A .

Таким образом, по теореме Бернулли при неограниченном увеличении числа испытаний математическая оценка вероятности случайного события сходится к истинному значению вероятности.

Вычисление интегралов методом статистических испытаний

Вычисление определенных интегралов — одна из наиболее изученных и распространенных задач недетерминистской природы, решаемых методом статистических испытаний. Практическая ценность метода для вычисления кратных интегралов обусловле-

на тем, что порядок сходимости метода статистических испытаний не зависит от кратности интеграла и при умеренных требованиях к точности результата упомянутая в начале данного параграфа задача о проникновении частиц через толстый экран может быть решена за минуты (вместо восьми лет). Известно несколько вероятностных схем решения задачи о вычислении интеграла. Рассмотрим основные идеи метода, оценку погрешности и некоторые приемы повышения эффективности процедуры получения решения.

Пусть требуется вычислить интеграл

$$I(u) = \int_G \phi(u) du, \quad (2.75)$$

где $\phi(u)$ — функция, заданная в S -мерной гиперплоскости на области G и u — координаты точки на этой гиперплоскости.

Приведем интеграл (2.75) к виду

$$I(u) = \int_G \frac{\phi(u)}{f(u)} f(u) du = \int_G \psi(u) f(u) du \quad (2.76)$$

и введем случайную величину ξ с плотностью распределения вероятностей $f(\xi)$ и случайную величину $\eta = \psi(\xi)$, являющуюся функцией случайной величины ξ . Математическое ожидание $M\eta$ и дисперсия $D\eta$ функции $\eta = \psi(\xi)$ случайной величины ξ равны

$$M\eta = \int_G \psi(u) f(u) du, \quad (2.77)$$

$$D\eta = M(\eta^2) - (M\eta)^2. \quad (2.78)$$

Пусть имеется последовательность независимых в совокупности случайных величин ξ_1, \dots, ξ_n . Согласно центральной предельной теореме, среднее арифметическое значений функции $\eta = \psi(\xi)$ по всей совокупности точек не зависит от функции распределения случайной величины ξ и является случайной величиной, распределенной при $n \rightarrow \infty$ по нормальному закону. На основании теоремы Чебышёва и из сравнения соотношений (2.76) и (2.77), которые показывают, что интеграл $I(u)$ можно рассматривать как математическое ожидание случайной величины $\eta_i = \psi(\xi_i)$, можно записать

$$P\left(\left|\frac{\eta_1 + \eta_2 + \dots + \eta_n}{n} - I\right| < \alpha \sqrt{\frac{D\eta}{n}}\right) = \Phi(\alpha). \quad (2.79)$$

Полученное соотношение дает способ вычисления интеграла как среднего арифметического случайных величин $\eta_i = \psi(\xi_i)$, где ξ — случайная величина с плотностью распределения $f(\xi)$, т. е.

$$I \approx \frac{1}{n} \sum_{i=1}^n \frac{\psi(\xi_i)}{f(\xi_i)}. \quad (2.80)$$

Кроме того, из соотношения (2.79) следует, что оценка погрешности вычисления интеграла носит вероятностный характер. Значение коэффициента α в соотношении (2.79) зависит от величины доверительной вероятности $\Phi(\alpha)$ и приведено в таблице 2.3 для трех значений этого коэффициента.

Таблица 2.3

$\Phi(\alpha)$	0,5	0,997	0,99999
α	0,675	3,0	5,0

При $P = \Phi(0,675) = 0,5$ ошибку называют вероятной. При $\alpha = 3$ вероятность P приближается к единице. Этот факт известен как правило трех сигм*.

Вероятностная оценка погрешности метода, согласно выражению (2.79), зависит от дисперсии случайной величины η , не зависит от кратности интеграла и убывает пропорционально $n^{-0,5}$. Последнее означает, что для уменьшения погрешности вычисления интеграла методом статистических испытаний в 10 раз необходимо увеличить число испытаний в 100 раз. По этой причине с учетом возможностей современных компьютеров и реальных требований к точности решения практических задач метод статистических испытаний редко используется для получения решения с точностью, лучшей, чем 1,0—0,1%.

Не известная в соотношении (2.79) величина дисперсии $D\eta$ случайной величины может быть оценена в процессе вычислений по соотношению (2.78), если заменить входящие в него математические ожидания на их оценки типа (2.80). Тогда

$$D\eta \approx \frac{1}{n-1} \left(\frac{1}{n} \sum_{i=1}^n (\psi(\xi_i))^2 - \left(\frac{1}{n} \sum_{i=1}^n \psi(\xi_i) \right)^2 \right).$$

Заметим, что возможность сравнительно простой оценки погрешности в процессе вычислений является важным достоинством метода статистических испытаний.

* Трех сигм правило — мнемоническое правило, согласно которому в некоторых задачах теории вероятностей и математической статистики считают практически невозможным событие, заключающееся в отклонении значения нормально распределенной случайной величины от ее математического ожидания больше, чем на три стандартных отклонения.

Другим достоинством метода, проявляющимся для интегралов высокой кратности, является то, что оценка погрешности вычисления значения интеграла методом статистических испытаний не зависит от кратности интеграла. В самом деле, если при интегрировании методом статистических испытаний принимать во внимание только порядок сходимости, то погрешность вычисления интеграла $\epsilon \approx n^{-0.5}$ не зависит от числа измерений. Следовательно, полное число испытаний (число узлов интегрирования) для обеспечения заданной точности составляет $n \approx \epsilon^{-2}$.

При использовании кубатурных формул с порядком точности P погрешность вычисления интеграла $\epsilon \approx k^{-P}$, где k — число узлов интегрирования по каждому измерению. Полное число узлов интегрирования $n = k^S$. Следовательно, при использовании кубатурных формул требуемое для достижения необходимой точности число узлов интегрирования

$$n \approx \epsilon^{-S/P}.$$

Так как порядок точности кубатурных формул для интегралов высокой кратности редко превышает значение $P = 2$, то метод статистических испытаний оказывается предпочтительным при вычислении интегралов, кратность S которых больше четырех.

Получена следующая простая оценка для сопоставления метода статистических испытаний и кубатурных формул. При вычислении интеграла кратности S методом статистических испытаний требуется $S \cdot m \cdot n$ операций, где m — число операций, необходимых для вычисления одного значения подынтегральной функции по одному измерению, n — число испытаний. С использованием кубатурных формул для вычисления интеграла требуется $S \cdot m \cdot k^S$ операций. Здесь k — число узловых точек по одному измерению.

Следует, однако, иметь в виду одно важное обстоятельство. При вычислении интегралов практическое значение имеет относительная погрешность результата

$$\delta I = \sqrt{D\eta} / |I| \sqrt{n},$$

зависящая от величины дисперсии. С увеличением кратности интеграла величина отношения $\sqrt{D\eta} / |I|$ может существенно возрастать. Так, для интеграла

$$I = \int_0^1 \dots \int_0^1 \exp(-32(x_1^2 + \dots + x_S^2)) dx_1 \dots dx_S$$

справедливо соотношение $\sqrt{D\eta} / |I| > 10^{S/2} - 1$.

В таких случаях оценки числа операций и времени счета для метода статистических испытаний при одинаковой относительной погрешности могут существенно ухудшаться.

Решение дифференциальных уравнений
в частных производных
методом статистических испытаний

Математические основы метода статистических испытаний связаны с соответствием задачи о случайному блуждании на многомерной решетке с одномерными марковскими цепями. Основные идеи метода статистических испытаний рассмотрены на примере ЗАДАЧИ ДИРИХЛЕ для УРАВНЕНИЯ ЛАПЛАСА.

Требуется найти решение уравнения

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad (2.81)$$

в области G с кусочно-гладкой границей Γ , на которой функция принимает заданные значения

$$u(x, y)|_{\Gamma} = \varphi(x, y). \quad (2.82)$$

Приблизим область интегрирования G сеточной областью G_h с постоянным шагом h по x и y (рис. 2.12). Темными кружками

ДИРИХЛÉ ПЕТЕР ГУСТАВ ЛЕЖЕН (Dirichlet Peter Gustav Lejeune; 1805—1859) — немецкий математик. Его основные труды относятся к теории чисел и математическому анализу. Д. впервые точно сформулировал и исследовал понятие условной сходимости ряда (признак Дирихле), дал строгое доказательство возможности разложения в ряд Фурье кусочно-монотонной функции с конечным числом точек разрыва. Значительные работы Д. посвящены механике и математической физике (принцип Дирихле в теории гармонических функций).

ЛАПЛАС ПЬЕР СИМОН (Laplace Pierre Simon; 1749—1827) — великий французский астроном, математик и физик. Фундаментальными являются работы Л. по дифференциальному уравнениям, например, по интегрированию методом каскадов уравнений в частных производных. Введенные им шаровые функции имеют разнообразное применение. В алгебре Л. принадлежит важная теорема о представлении определителей суммой произведений дополнительных миноров (теорема Лапласа). Математическая теория вероятностей в значительной мере создана Л.; в частности, он ввел производящие функции и широко применял преобразование, носящее его имя.

здесь обозначены граничные узлы; светлыми — внутренние узлы. Значения функции в граничных точках сеточной области определяем точно так же, как это делалось в разностных методах. Напомним, что при использовании метода сеток значения функции во внутренних узлах (i, j) определяются из системы линейных уравнений

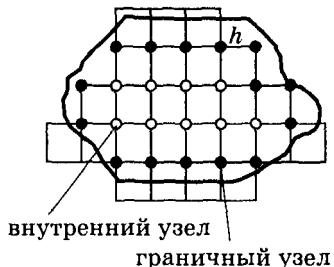


Рис. 2.12

$$u_{i,j} = \frac{1}{4}(u_{i-1,j} + u_{i+1,j} + u_{i,j-1} + u_{i,j+1}), \quad (2.83)$$

дополненной значениями функции в граничных точках

$$u_{pq} = \phi(x_p, y_q) = \varphi_{pq}.$$

Пусть некоторая точка, начав движение во внутреннем узле, совершает равномерное случайное блуждание по узлам сетки. Попав в граничный узел (p, q), точка прекращает движение. Доказывается, что с вероятностью, равной единице, каждая блуждающая точка попадает в один из граничных узлов (p, q). Тогда можно записать, что

$$\sum_{p,q} P(i,j,p,q) = 1. \quad (2.84)$$

Здесь суммирование проводится по всем граничным узлам, а $P(i,j,p,q)$ — вероятность того, что точка, начавшая движение во внутреннем узле (i, j), закончит движение в граничном узле (p, q). При этом для $p' \neq p$ и (или) $q' \neq q$

$$\begin{aligned} P(p, q, p, q) &= 1, \\ P(p', q', p, q) &= 0. \end{aligned} \quad (2.85)$$

Если в сумме

$$V_{i,j} = \sum_{p,q} P(i,j,p,q) \cdot \varphi_{pq} \quad (2.86)$$

рассматривать $\varphi(x, y)$ как случайную величину, принимающую значения на границе сеточной области Γ_h , то $V_{i,j}$ есть математическое ожидание функции $\varphi(x, y)$ на границе для точек, начавших движение во внутреннем узле (i, j).

Найти вероятности $P(i, j, p, q)$ в явном виде сложно, однако можно получить соотношение, связывающее искомые вероятности во внутренних узлах.

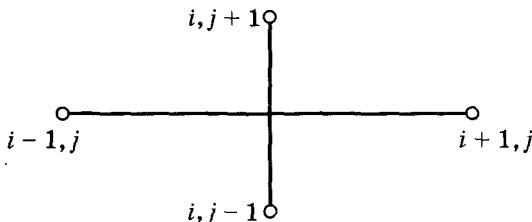


Рис. 2.13

Движение точки, начавшееся в некотором внутреннем узле (i, j) и закончившееся в граничном узле (p, q) , можно представить как движение, распадающееся с равной вероятностью на четыре новых движения, начинающихся соответственно в точках $(i - 1, j)$, $(i + 1, j)$, $(i, j - 1)$, $(i, j + 1)$ (рис. 2.13).

По формуле полной вероятности можно записать

$$\begin{aligned} P_{i, j, p, q} = & \frac{1}{4} [P(i - 1, j, p, q) + P(i + 1, j, p, q) + \\ & + P(i, j - 1, p, q) + P(i, j + 1, p, q)]. \end{aligned} \quad (2.87)$$

Умножая обе части последнего равенства на φ_{pq} и суммируя по всем граничным узлам, получим

$$V_{i, j} = \frac{1}{4} (V_{i-1, j} + V_{i+1, j} + V_{i, j-1} + V_{i, j+1}). \quad (2.88)$$

В соответствии со значениями (2.85) имеем $V_{p, q} = \varphi_{pq}$.

Из сравнения выражений (2.88) и (2.83) следует, что формулы совпадают с точностью до обозначений и, следовательно, исключенные неизвестные $u_{i, j}$ можно рассматривать как математическое ожидание $V_{i, j}$ функции $\varphi(x, y)$ на границе для точки, начавшей движение во внутреннем узле (i, j) .

Если многократно промоделировать движение точки из внутреннего узла (i, j) и при каждом испытании фиксировать граничный узел, в который попадает ближайшая точка, то приближенное значение математического ожидания можно вычислить по соотношению

$$V_{i, j} \approx \frac{1}{n} \sum_{l=1}^n \varphi_{pq}^{(l)}. \quad (2.89)$$

Приближенное решение задачи (2.81)–(2.82) во внутреннем узле сеточной области G_h получим, положив $u_{i, j} = V_{i, j}$.

Соотношение (2.89) содержит в неявном виде вероятности $P(i, j, p, q)$, так как суммирование идет по всем испытаниям и,

следовательно, каждое значение Φ_{pq} входит в сумму $m_{p,q}$ раз, где $m_{p,q}$ — число попаданий блуждающей точки в граничный узел (p, q) .

Получаемый алгоритм соответствует принципу максимума

$$\min u(x, y)|_{\Gamma} \leq u(x, y) \leq \max u(x, y)|_{\Gamma}.$$

Основные особенности применения метода статистических испытаний к решению краевых задач заключаются в следующем.

- Решение задачи определяется в единственном узле (i, j) области интегрирования, в котором начинается блуждание точки. Часто для практических приложений эта особенность оказывается весьма полезной. Дополнительная информация о решении задачи может быть получена, если использовать то обстоятельство, что каждый внутренний узел сеточной области, через которую проходит блуждающая точка, может считаться узлом, в котором начинается движение.

- Для вычисления решения в выбранном узле (i, j) сеточной области не требуется помнить информацию об остальных узлах сеточной области (кроме граничных) — алгоритм метода статистических испытаний обладает высокой эффективностью по памяти.

Метод статистических испытаний обобщается на более общие краевые условия, виды уравнений и задачи, например, нестационарные. Предложены вероятностные схемы для задач, в которых вероятность перехода из узла в узел зависит от истории движения и не равна нулю вероятность выхода блуждающей точки из граничного узла.

ПРИМЕР. Методом статистических испытаний найти решение уравнения

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

в точке $(2, 2)$ квадратной области $G = [0 \leq x \leq 4] \times [0 \leq y \leq 4]$ с граничными условиями, заданными на схеме (рис. 2.14). Шаг сетки $h = 1,0$.

Для определения направления движения блуждающей точки используем таблицу случайных чисел по следующему правилу: цифры 0 или 1 соответствуют движению вправо на шаг, цифры 2 или 3 — вверх на

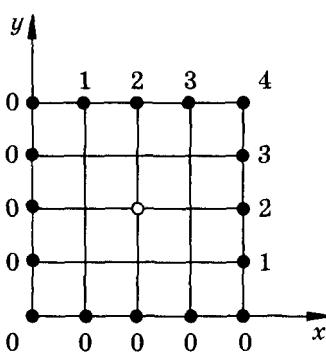


Рис. 2.14

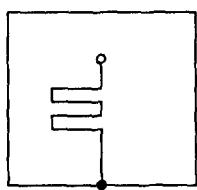


Рис. 2.15

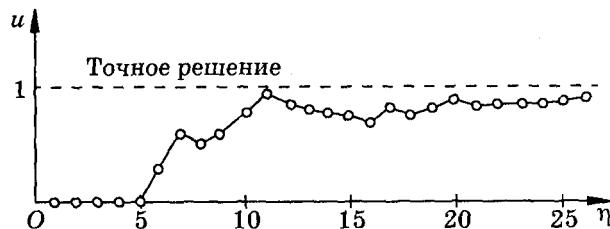


Рис. 2.16

шаг, цифры 4 или 5 — влево на шаг, цифры 6 или 7 — вниз на шаг, цифры 8 или 9 — точка остается на месте.

Траектория движения первой выпущенной точки (первое испытание) показана на рис. 2.15 (последовательность случайных цифр 8—6—5—1—5—9—0—7). Результаты первых тридцати испытаний приведены на графике рис. 2.16.

Глава 3

ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

В главе рассматриваются различные методы решения задачи Коши и краевой задачи для обыкновенных дифференциальных уравнений. Подробно рассмотрены метод Эйлера и ошибки, возникающие при его реализации на компьютерах; семейство одношаговых методов Рунге — Кутты, многошаговые методы. В качестве примеров методов решения краевой задачи излагаются методы стрельбы и разностный.

3.1. Основы теории обыкновенных дифференциальных уравнений

ОБЫКНОВЕННЫМ ДИФФЕРЕНЦИАЛЬНЫМ УРАВНЕНИЕМ называется уравнение вида

$$F(x, y, y', \dots, y^{(n)}) = 0, \quad (3.1)$$

связывающее одну независимую переменную x , исковую функцию $y(x)$ и ее производные вплоть до n -го порядка. ПОРЯДКОМ обыкновенного дифференциального уравнения называется порядок старшей производной от исковой функции.

Дифференциальное уравнение называется ЛИНЕЙНЫМ, если оно имеет вид

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

Так, например, $y'' - x^2y + x^2 = 0$ — линейное дифференциальное уравнение второго порядка; $y'' + e^y = 0$ — нелинейное уравнение.

ОБЩИМ ИНТЕГРАЛОМ уравнения (3.1) называют функцию

$$\Phi(x, y, c_1, \dots, c_n), \quad (3.2)$$

связывающую независимую переменную x , искомую функцию $y(x)$ и n постоянных интегрирования c_1, \dots, c_n с помощью уравнения $\Phi(x, y, c_1, \dots, c_n) = 0$. Таким образом, функция $y(x)$ входит в функцию (3.2) неявным образом, причем число постоянных интегрирования равно порядку уравнения.

ОБЩИМ РЕШЕНИЕМ обыкновенного дифференциального уравнения называется функция

$$y(x) = \varphi(x, c_1, \dots, c_n), \quad (3.3)$$

связывающая независимую переменную и n постоянных интегрирования, иначе говоря, уравнение (3.3) определяет функцию $y(x)$ явным образом.

Для определения постоянных интегрирования задаются **ДОПОЛНИТЕЛЬНЫЕ УСЛОВИЯ**. Их число равно числу постоянных интегрирования. Если в дополнительные условия подставить функцию (3.2) и решить полученную систему относительно постоянных интегрирования c_1, \dots, c_n , а затем полученные решения подставить в (3.2), то получим **ЧАСТНЫЙ ИНТЕГРАЛ** $\Phi_1(x, y(x)) = 0$.

Аналогичные процедуры с общим решением (3.3) дают частное решение $y(x) = \varphi_1(x)$.

Если все дополнительные условия задаются более чем в одной точке x_0 , то совокупность обыкновенного дифференциального уравнения и дополнительных условий называют **ЗАДАЧЕЙ КОШИ** для рассматриваемого дифференциального уравнения. В этом случае дополнительные условия называют **НАЧАЛЬНЫМИ УСЛОВИЯМИ**.

Если дополнительные условия задаются более чем в одной граничной точке расчетной области, то совокупность обыкновенного дифференциального уравнения и дополнительных условий называют **КРАЕВОЙ ЗАДАЧЕЙ** для рассматриваемого уравнения, а дополнительные условия — **ГРАНИЧНЫМИ** или **КРАЕВЫМИ УСЛОВИЯМИ**.

КУТТА МАРТИН ВИЛЬГЕЛЬМ (Kutta Martin Wilhelm; 1867—1944) — немецкий физик и математик. В математике К. развел метод Рунге (метод Рунге — Кутты) численного решения задачи Коши для системы обыкновенных дифференциальных уравнений.

Таким образом, если задать числа y_0, y_1, \dots, y_{n-1} , то задача Коши для уравнения (3.1) имеет вид

$$\begin{cases} F(x, y, y', \dots, y^{(n)}) = 0, \\ y(x_0) = y_0, \\ y'(x_0) = y_1, \\ \dots \dots \dots \dots \\ y^{(n-1)}(x_0) = y_{n-1}. \end{cases}$$

Порядок старшей производной в начальных условиях не превышает $(n-1)$, где n — порядок дифференциального уравнения.

Если старшая производная в уравнении выражается как функция переменных x, y и производных более низкого порядка, то уравнение называют РАЗРЕШЕННЫМ ОТНОСИТЕЛЬНО СТАРШЕЙ ПРОИЗВОДНОЙ и оно записывается в виде $y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$.

Задачу Коши для обыкновенного дифференциального уравнения n -го порядка $y^{(n)} = f(x, y, y', \dots, y^{(n-1)})$ можно свести к задаче Коши для следующей нормальной системы уравнений

$$\begin{cases} y' = z_1(x), \\ z_1' = z_2(x), \\ \dots \dots \dots \dots \\ z_{n-2}' = z_{n-1}(x), \\ z_{n-1}' = f(x, y(x), z_1(x), \dots, z_{n-1}(x)), \\ y(x_0) = y_0, \\ z_1(x_0) = y_1, \\ z_2(x_0) = y_2, \\ \dots \dots \dots \dots \\ z_{n-1}(x_0) = y_{n-1}. \end{cases}$$

КОШИ ОГЮСТЕН ЛУИ (Cauchy Augustin Louis; 1789—1857) — французский математик. В его курсах анализа, основанных на систематическом применении понятия предела, сформулировано представление о непрерывности функции, дано четкое построение теории сходящихся рядов (критерий Коши). В теории аналитической функции комплексного переменного К. дал выражение аналитической функции в виде интеграла Коши, получил разложение функции в степенной ряд. В теории дифференциальных уравнений К. поставил одну из важнейших общих задач, развел метод интегрирования уравнений с частными производными первого порядка. К. является одним из создателей теории упругости.

Здесь $z_1(x), \dots, z_{k-1}$ вводятся следующим образом:

$y'(x) = z_1(x)$ — первое дифференциальное уравнение,

$y''(x) = z'_1(x) = z_2(x)$ — второе дифференциальное уравнение,

.....

$y^{(n-1)}(x) = \dots = z'_{n-2}(x) = z_{n-1}(x)$ — $(n-1)$ -е дифференциальное уравнение.

Актуальность задачи Коши для многих областей науки и техники явилась причиной разработки для ее решения большого количества методов. Рассматриваются две группы численных методов решения задачи Коши.

- **ОДНОШАГОВЫЕ МЕТОДЫ**, в которых для нахождения решения в некоторой точке отрезка используется информация лишь в одной предыдущей точке. К этим методам относятся методы Эйлера и Рунге — Кутты.

- **МНОГОШАГОВЫЕ МЕТОДЫ**, в которых для отыскания решения в некоторой точке используется информация о решении в нескольких предыдущих точках. К таким методам относится метод Адамса.

Рассмотрим применение этих методов к дифференциальным уравнениям первого порядка. Алгоритм для случая системы дифференциальных уравнений легко получается из алгоритма для решения одного уравнения.

Необходимо отметить, что подобные методы можно применять только к хорошо обусловленным задачам, т. е. к таким задачам, решение которых мало изменяется в случае малого изменения начальных условий. В противном случае небольшие ошибки, которые неизбежны при решении задачи, могут существенно изменить решение.

ПРИМЕР. В качестве плохо обусловленной задачи рассмотрим решение следующей задачи Коши на отрезке $[0, 100]$:

$$u'(x) = u - x, \quad u(0) = 1.$$

Общим решением этого дифференциального уравнения является

$$u(x, c) = 1 + x + ce^x.$$

Из начального условия находим, что постоянная $c = 0$, поэтому $u(100) = 101$.

Предположим, что начальное условие изменилось и $u(0) = 1,000001$. Следовательно, изменится и постоянная c в общем решении. В этом случае $c = 10^{-6}$ и решение имеет вид $u(100) =$

$\approx 2,7 \cdot 10^{37}$, т. е. изменение начального условия на 10^{-6} изменило решение на 35 порядков.

Постановка краевых задач для обыкновенных дифференциальных уравнений

Постановку краевых задач для обыкновенных дифференциальных уравнений рассмотрим на примере уравнения второго порядка с переменными коэффициентами:

$$y'' + p(x)y' + q(x)y = f(x), \quad x \in [0, 1].$$

- Если на границах $x = 0$ и $x = 1$ заданы значения искомой функции, то такие граничные условия называются ГРАНИЧНЫМИ УСЛОВИЯМИ ПЕРВОГО РОДА, а задачу

$$\begin{cases} y'' + p(x)y' + q(x)y = f(x), \\ y(0) = y_0, \\ y(1) = y_1, \end{cases}$$

где y_0, y_1 — заданные числа, называют ПЕРВОЙ КРАЕВОЙ ЗАДАЧЕЙ.

- Если на границах $x = 0$ и $x = 1$ заданы значения производных искомой функции, то такие граничные условия называются ГРАНИЧНЫМИ УСЛОВИЯМИ ВТОРОГО РОДА:

$$\begin{cases} y'(0) = y_0, \\ y'(1) = y_1, \end{cases}$$

а соответствующая задача называется ВТОРОЙ КРАЕВОЙ ЗАДАЧЕЙ:

$$\begin{cases} y'' + p(x)y' + q(x)y = f(x), \\ y'(0) = y_0, \\ y'(1) = y_1. \end{cases}$$

- Если на границах задана линейная комбинация первой производной и самой искомой функции, то такие граничные условия называются ГРАНИЧНЫМИ УСЛОВИЯМИ ТРЕТЬЕГО РОДА:

$$\begin{cases} \alpha_0 y'(0) + \beta_0 y(0) = y_0, \\ \alpha_1 y'(1) + \beta_1 y(1) = y_1. \end{cases}$$

Такая задача называется ТРЕТЬЕЙ КРАЕВОЙ ЗАДАЧЕЙ:

$$\begin{cases} y'' + p(x)y' + q(x)y = f(x), \\ \alpha_0 y'(0) + \beta_0 y(0) = y_0, \\ \alpha_1 y'(1) + \beta_1 y(1) = y_1. \end{cases}$$

Очевидно, что если порядок обыкновенного дифференциального уравнения равен n , то порядок старшей производной в краевых условиях не может превышать значения $(n - 1)$.

В заключение приведем одну из постановок краевой задачи для обыкновенного дифференциального уравнения четвертого порядка, встречающегося в теории упругости:

$$\begin{cases} y^{(IV)} + p(x)y = f(x), \\ y(0) = y_0, \\ y'(0) = y_1, \\ y(1) = y_2, \\ y'(1) = y_3. \end{cases}$$

Здесь y_0, y_1, y_2, y_3 — заданные числа.

3.2. Понятие о методе конечных разностей.

Порядок точности разностной схемы

Численное решение обыкновенных дифференциальных уравнений и уравнений в частных производных во многих случаях осуществляется методом конечных разностей. МЕТОД КОНЕЧНЫХ РАЗНОСТЕЙ сводит решение дифференциальных уравнений к решению линейных или нелинейных уравнений с достаточно разреженными матрицами. При этом построение решения в методе сеток осуществляется в три этапа.

- Область непрерывного изменения аргумента (или аргументов) заменяется конечным дискретным множеством точек, называемых РАЗНОСТНОЙ СЕТКОЙ. В разностной сетке выделяются внутренние и граничные узлы. Решение разыскивается во внутренних узлах, а в граничных узлах значение искомой функции задается при аппроксимации граничных условий исходной дифференциальной задачи. Функция дискретного аргумента, определенная на разностной сетке, называется СЕТОЧНОЙ ФУНКЦИЕЙ.

- Дифференциальные уравнения и граничные условия заменяются по определенным правилам своими разностными аналогами. Разностные операторы, соответствующие дифференциальному уравнению, записываются во внутренних узлах сетки. Разностные операторы, соответствующие граничным условиям, записываются в граничных узлах. В результате получается система

алгебраических уравнений, число которых пропорционально числу внутренних узлов разностной сетки.

- Осуществляется решение системы алгебраических уравнений каким-либо из известных методов. В большинстве случаев получаемая система уравнений является системой линейных алгебраических уравнений очень большого порядка (как правило, $N \geq 100$), но с весьма разреженной матрицей.

В случае нелинейных систем итерационные процедуры, как правило, сводят их к линейным системам.

Продемонстрируем эти этапы на примере решения задачи Коши для простейшего линейного уравнения первого порядка. Пусть требуется найти решение задачи Коши на интервале $(0, 1)$ для уравнения

$$\frac{du}{dx} + Au = 0, \quad u(0) = u_0. \quad (3.4)$$

Непрерывный интервал можно заменить множеством точек $0, x_1, x_2, \dots, x_{n-1}, x_n, \dots, 1$. Расстояние между точками h называется *шагом разностной сетки*. В общем случае эти шаги могут быть переменными. Рассмотрим случай постоянного шага и примем $x_n = nh$. Заменим дифференциальное уравнение разностным в точке x_n :

$$\frac{du}{dx} + Au \approx \frac{u(x_{n-1} + h) - u(x_{n-1})}{h} + Au(x_{n-1}) = 0. \quad (3.5)$$

Таким образом, первая производная аппроксимирована односторонней разностью вперед с первым порядком точности. Перепишем уравнение (3.5) в виде

$$\frac{u_n - u_{n-1}}{h} + Au_{n-1} = 0. \quad (3.6)$$

Здесь $u(x_{n-1} + h) = u_n$, а $u(x_{n-1}) = u_{n-1}$. Из (3.6) следует, что

$$u_n = (1 - Ah)u_{n-1}. \quad (3.7)$$

Из рекуррентного соотношения (3.7) имеем точное решение разностного уравнения (3.6)

$$u_n = u_0(1 - Ah)^n = u_0(1 - Ah)^{\frac{x_n}{h}}. \quad (3.8)$$

Очевидно, что точное решение разностных уравнений можно получить лишь для простейших случаев. Точное решение задачи (3.4) имеет вид

$$u(x_n) = u_0 e^{-Ax_n}. \quad (3.9)$$

Определим погрешность разностного решения. Имеем

$$\delta(x_n) = [e^{-Ax_n} - (1 - Ah)^{\frac{x_n}{h}}] u_0. \quad (3.10)$$

Предполагая шаг разностной схемы малым, представим член $(1 - Ah)^{\frac{x_n}{h}}$ в виде

$$\begin{aligned} (1 - Ah)^{\frac{x_n}{h}} &= e^{\frac{x_n}{h} \ln(1 - Ah)} = e^{\frac{x_n}{h} \left[-Ah - \frac{A^2 h^2}{2} + O(h^2) \right]} = \\ &= e^{-Ax_n} e^{-\frac{A^2 h^2}{2} x_n} e^{O(h^2)} = e^{-Ax_n} \left[1 - \frac{A^2 h x_n}{2} + O(h^2) \right] [1 + O(h^2)] = \\ &= e^{-Ax_n} - h \frac{A^2 x_n}{2} e^{-Ax_n} + O(h^2). \end{aligned}$$

Подставляя это выражение в (3.10), получим

$$\delta(x_n) = h u_0 \frac{A^2 x_n}{2} e^{-Ax_n} + O(h^2) = O(h). \quad (3.11)$$

Таким образом, величина погрешности имеет первый порядок точности и совпадает в данном случае с порядком аппроксимации производной. В общем случае порядок погрешности решения определяется не только порядком аппроксимации дифференциального уравнения, но и порядком аппроксимации граничных условий.

Говорят, что разностная схема (3.5) имеет первый порядок точности на интервале. Можно показать, что на шаге эта разностная схема имеет второй порядок точности. Представим точное решение в виде ряда Тейлора в точке x_n . Имеем

$$u_n = u_{n-1} + h \frac{du}{dx} \Big|_n + \frac{h^2}{2!} \frac{d^2 u}{dx^2} \Big|_n + O(h^3). \quad (3.12)$$

Поскольку

$$\frac{du}{dx} \Big|_n = -A u_{n-1}, \quad \frac{d^2 u}{dx^2} \Big|_n = +A^2 u_{n-1},$$

то

$$u_n = (1 - Ah) u_{n-1} + A^2 u_{n-1} \frac{h^2}{2} + O(h^3). \quad (3.13)$$

Из соотношений (3.7) и (3.13) следует, что погрешность на шаге имеет порядок $O(h^2)$. Очевидно, что погрешность на интер-

вале выше на порядок, так как она накапливается по мере увеличения числа шагов. Отметим еще, что погрешность решения стремится к нулю при $h \rightarrow 0$, т. е. решение разностной задачи сходится к точному. Однако для сходимости необходимы не только аппроксимация дифференциального уравнения, но и устойчивость разностной схемы. Не углубляясь в проблему устойчивости, отметим лишь, что свойство устойчивости можно трактовать как равномерную зависимость решения разностной задачи относительно возмущений правых частей и граничных условий от шага h .

В теории разностных схем доказывается, что *разностная схема является сходящейся*, если она обладает свойством аппроксимации и устойчива.

3.3. Метод Эйлера. Метод Эйлера с пересчетом

Метод Эйлера является простейшим методом решения задачи Коши и имеет невысокую точность, поэтому на практике его используют достаточно редко. Однако на его основе в дальнейшем легче объяснить алгоритмы более эффективных (и, как правило, более сложных) методов и способы построения и исследования этих алгоритмов.

Рассмотрим задачу Коши для дифференциального уравнения первого порядка

$$\begin{aligned} \frac{du}{dx} &= f(x, u), \quad x \in [a, b], \\ u(a) &= u_0. \end{aligned} \tag{3.14}$$

ЭЙЛЕР ЛЕОНАРД (Euler Leonhard; 1707—1783) — великий математик, механик и физик, швейцарец по происхождению, академик Петербургской АН. Необыкновенно широк был круг исследований Э., охватывавших все отделы современной ему математики и механики, теорию упругости, оптику, теорию музыки, баллистику, теорию машин, морскую науку и т. д. Основные достижения Э. как математика заключались в разработке математического анализа. Он стал создателем вариационного исчисления, теории дифференциальных уравнений, теории специальных функций. По выражению П. Лапласа, Э. явился общим учителем математиков 2-й половины XVII в. Мировое признание принесли Э. его труды по механике.

Обозначим границы отрезка a и b через x_0 и x_N соответственно и введем на отрезке $[a, b]$ сетку (в общем случае неравномерную) значений аргумента x такую, чтобы выполнялось соотношение $x_0 < x_1 < \dots < x_N$. Выполним разложение решения $u(x)$ в окрестности узла сетки x_n по формуле Тейлора. Обозначив шаг сетки через $h_n = x_{n+1} - x_n$ и $u(x_n)$ через u_n , получаем

$$u_{n+1} = u_n + h_n u'_n + \frac{1}{2!} h_n^2 u''_n + \frac{1}{3!} h_n^3 u'''_n + \dots . \quad (3.15)$$

Если функция f имеет непрерывную p -ю производную, то в соотношении (3.15) можно оставлять члены вплоть до $O(h^{p+1})$. Эти производные можно найти, дифференцируя правую часть уравнения (3.14) требуемое число раз. Например, для первой и второй производных имеем

$$u'(x) = f(x, u),$$

$$u''(x) = \frac{d}{dx} u'(x) = \frac{d}{dx} f(x, u) = f'_u u' + f'_x = f f'_u + f'_x = f \frac{\partial f}{\partial u} + \frac{\partial f}{\partial x}.$$

Аналогично можно получить производные более высоких порядков.

Однако использование формулы (3.15) с большим числом членов имеет ряд недостатков: во-первых, с ростом порядка производной выражение для нее может оказаться очень сложным; кроме того, если функция f известна лишь приближенно или задана таблично, ее производные находятся с большой ошибкой. В связи с этим в разложении (3.15) оставляют только два члена. При такой замене вместо точного решения $u(x_{n+1})$ получается его приближенное значение $u(x_{n+1})$, которое находится по формуле

$$u_{n+1} = u_n + h_n f(x_n, u_n). \quad (3.16)$$

Так как значение $u(x_0) = u_0$ известно из начального условия, то, последовательно пользуясь формулой (3.16), находим приближенные решения u_1, u_2, \dots, u_N . Формула (3.16) записана для случая неравномерной сетки. Полагая шаг сетки $h_n = h$ постоянным, получим

$$u_{n+1} = u_n + h f(x_n, u_n). \quad (3.16a)$$

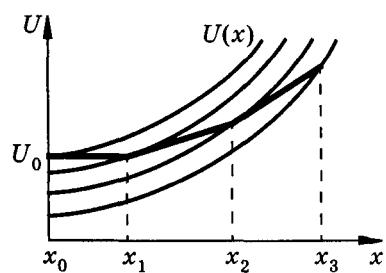


Рис. 3.1

Формула (3.16а) является ОСНОВНОЙ ФОРМУЛОЙ МЕТОДА ЭЙЛЕРА или МЕТОДА ЛОМАНЫХ. Последнее название становится понятным из геометрической интерпретации схемы, представленной

на рис. 3.1, на которой изображено семейство интегральных кривых уравнения (3.14).

Как видно, решение задачи Коши по методу Эйлера дает решение, не совпадающее ни с одной из интегральных кривых, и является ломаной линией, совпадающей на каждом шаге с касательной к соответствующей интегральной кривой. Из рис. 3.1 следует, что метод Эйлера дает одностороннее приближение к точному решению $u(x)$. Сравнение формул (3.15) и (3.16) показывает, что метод Эйлера обладает вторым порядком точности на шаге, являясь в то же время методом первого порядка аппроксимации на интервале.

Существует несколько модификаций метода Эйлера. Остановимся на одной из них — методе Эйлера с пересчетом, который называют также МЕТОДОМ ЭЙЛЕРА — КОШИ. В этом методе значение u_{n+1} находится по формуле

$$u_{n+1} = u_n + \frac{h}{2} [f(x_n, u_n) + f(x_{n+1}, u_{n+1})], \quad (3.17)$$

т. е. вместо тангенса угла наклона касательной к интегральной кривой в точке (x_n, u_n) , который имеет место в формуле (3.16a), выражающей метод Эйлера, используется полусумма значений тангенсов углов наклона касательных в известной (x_n, u_n) и искомой (x_{n+1}, u_{n+1}) точках. Поскольку, однако, значение u_{n+1} неизвестно, то (3.17) есть в общем случае нелинейное уравнение относительно u_{n+1} , которое можно решить различными методами, изложенными в главе 1.

В рассматриваемом случае логично использовать метод простой итерации, представленный формулой (1.10), поскольку нелинейное уравнение уже разрешено относительно u_{n+1} . Тогда, если номер итерации обозначить верхним индексом, итерационный процесс запишется в виде

$$u_{n+1}^{(k+1)} = u_n + \frac{h}{2} [f(x_n, u_n) + f(x_n + h, u_{n+1}^{(k)})]. \quad (3.18)$$

В качестве значения $u_{n+1}^{(0)}$ можно принять либо u_n , либо $u_{n+1}^{(0)} = u_n + hf(x_n, u_n)$, т. е. использовать значение, вычисленное по формуле Эйлера (3.16a). В этом случае в первой итерации имеем

$$u_{n+1}^{(1)} = u_n + \frac{h}{2} \{f(x_n, u_n) + f[(x_n + h, u_n + hf(x_n, u_n))]\}. \quad (3.19)$$

Формула (3.19) и есть ОСНОВНАЯ ФОРМУЛА МЕТОДА ЭЙЛЕРА С ПЕРЕСЧЕТОМ. Подобные схемы часто называют схемами типа ПРОГНОЗ-КОРРЕКТОР (ПРЕДИКТОР-КОРРЕКТОР). Сна-

чала по формуле (3.16) определяется прогнозируемое приближение решения, а затем по формуле (3.19) это решение уточняется.

Метод Эйлера с пересчетом обладает третьим порядком точности на шаге. Действительно, из соотношения (3.15) следует

$$u_{n+1} = u_n + hf(x_n, u_n) + \frac{h^2}{2} \left[\frac{\partial f}{\partial x} + f(x_n, u_n) \frac{\partial f}{\partial u} \right] + O(h^3). \quad (3.20)$$

Разлагая в ряд второй член в квадратных скобках (3.19), имеем

$$f[x_n + h, u_n + hf(x_n, u_n)] = f(x_n, u_n) + \frac{\partial f}{\partial x} h + \frac{\partial f}{\partial u} hf(x_n, u_n). \quad (3.21)$$

Подставляя правую часть соотношения (3.21) в последний член выражения (3.19), получаем

$$u_{n+1} = u_n + hf(x_n, u_n) + \frac{h^2}{2} \left[\frac{\partial f}{\partial x} + f(x_n, u_n) \frac{\partial f}{\partial u} \right]. \quad (3.22)$$

Сравнивая соотношения (3.20) и (3.22), видим, что метод Эйлера с пересчетом обладает третьим порядком точности на шаге и, соответственно, вторым на интервале. Метод Эйлера с пересчетом дает двустороннее приближение к решению.

Отметим, что поскольку нулевое приближение итерационного процесса задано, то итерационный процесс (3.18) можно продолжать до достижения заданной точности. Однако последующие итерации не повышают порядка точности схемы, и поэтому на практике используют одну-две итерации. Будет показано, что метод Эйлера и метод Эйлера с пересчетом являются частными случаями двухстадийных схем Рунге — Кутты.

3.4. Метод Рунге — Кутты

Метод Рунге — Кутты позволяет строить схемы различного порядка точности. Основная идея метода состоит в построении специального алгоритма — такого, чтобы приращение функции на шаге $\Delta u = u_{n+1} - u_n$ совпадало с приращением Δu , которое определяется из ряда Тейлора (3.15) с учетом возможно большего числа членов. При этом вторые и следующие производные определяются не дифференцированием, а путем многократного вычисления функции $f(x, u)$ в некоторых промежуточных точках между x_n и x_{n+1} .

Проиллюстрируем основные идеи метода на примере получения схемы второго порядка точности.

Оставим в разложении (3.15) члены вплоть до $O(h^2)$, имея в виду, что последний член разложения соответствует предполагаемому порядку точности схемы. Имеем

$$u_{n+1} = u_n + hu'_n + \frac{1}{2}h^2u''_n. \quad (3.23)$$

Чтобы избежать явного дифференцирования, заменим вторую производную u''_n в соотношении (3.23) разностью

$$u''_n = \frac{d}{dx}[f(x, u)] \approx \frac{f(\bar{x}, \bar{u}) - f(x_n, u_n)}{\Delta x}. \quad (3.24)$$

Величины \bar{x} , \bar{u} и Δx подбираются так, чтобы обеспечить нужный порядок точности. Подставляя выражение (3.24) в (3.23), получаем

$$u_{n+1} = u_n + hu'_n + \frac{1}{2}\frac{h^2}{\Delta x}[f(\bar{x}, \bar{u}) - f(x_n, u_n)]. \quad (3.25)$$

Полагая $\bar{x} = x_n + \gamma h$, $\bar{u} = u_n + \delta h$, обозначая $\alpha = h/2\Delta x$ и имея в виду, что $\Delta x = \bar{x} - x_n$, и, следовательно, $\gamma = 1/2\alpha$, а также учитывая, что $u'_n = f(x_n, u_n)$, переписываем выражение (3.25) в виде

$$u_{n+1} = u_n + h \left[(1 - \alpha)f(x_n, u_n) + \alpha f \left(x_n + \frac{h}{2\alpha}, u_n + \delta h \right) \right]. \quad (3.26)$$

Параметры α и δ определяются из условия наилучшего соответствия выражения (3.26) ряду (3.15). Принимая

$$f \left(x_n + \frac{h}{2\alpha}, u_n + \delta h \right) = f(x_n, u_n) + \frac{h}{2\alpha} \frac{\partial f}{\partial x} + \delta h \frac{\partial f}{\partial u} \quad (3.27)$$

и подставляя соотношение (3.27) в (3.26), имеем

$$u_{n+1} = u_n + hf(x_n, u_n) + \frac{1}{2}h^2 \left(\frac{\partial f}{\partial x} + 2\delta\alpha \frac{\partial f}{\partial u} \right)_n. \quad (3.28)$$

Перепишем ряд (3.15) в виде

$$u_{n+1} = u_n + hf(x_n, u_n) + \frac{1}{2}h^2 \left(\frac{\partial f}{\partial x} + f \frac{\partial f}{\partial u} \right)_n + O(h^3). \quad (3.29)$$

Из сравнения формул (3.27) и (3.29) видно, что первые три члена в этих формулах совпадают, если $\alpha\delta = f(x_n, u_n)/2$. Выражая δ через α и подставляя в (3.26), получаем однопараметрическое се-

мейство разностных схем Рунге — Кутты второго порядка точности. В результате

$$\begin{aligned} u_{n+1} = u_n + h \left\{ (1 - \alpha) f(x_n, u_n) + \right. \\ \left. + \alpha f \left(x_n + \frac{h}{2\alpha}, u_n + \frac{h}{2\alpha} f(x_n, u_n) \right) \right\}. \end{aligned} \quad (3.30)$$

Схема (3.30) при $\alpha \neq 0$ имеет третий порядок точности на шаге и второй на интервале. Из (3.30) можно получить все рассмотренные разностные схемы. Так, при $\alpha = 0$ имеем метод Эйлера, выражаемый формулой (3.16а), при $\alpha = 1/2$ — первую итерацию метода Эйлера с пересчетом, которому соответствует формула (3.19), а при $\alpha = 1$ — так называемый метод хорд, которому соответствует соотношение

$$u_{n+1} = u_n + h f \left(x_n + \frac{h}{2}, u_n + \frac{h}{2} f(x_n, u_n) \right). \quad (3.31)$$

Аналогичным образом могут быть получены и схемы Рунге — Кутты более высокого порядка точности. В настоящее время наиболее распространены схемы четвертого порядка точности. Например, схема Рунге — Кутты четвертого порядка точности, которая используется в большинстве стандартных программ на компьютерах, имеет вид

$$\begin{aligned} u_{n+1} = u_n + \frac{h}{6} (K_1 + 2K_2 + 2K_3 + K_4), \\ K_1 = f(x_n, u_n), \quad K_2 = f \left(x_n + \frac{h}{2}, u_n + \frac{h}{2} K_1 \right), \\ K_3 = f \left(x_n + \frac{h}{2}, u_n + \frac{h}{2} K_2 \right), \quad K_4 = f(x_n + h, u_n + h K_3). \end{aligned}$$

3.5. Многошаговый метод Адамса

В предыдущих схемах решение в точке x_{n+1} вычисляется с использованием решения только в одной точке x_n . Логично предположить, что можно повысить точность метода, если использовать информацию о поведении решения в предыдущих точках x_n, x_{n-1}, \dots . Такие методы получили название МНОГОШАГОВЫХ.

Общая схема построения многошаговых методов выглядит следующим образом. Пусть нам известно приближенное решение

в нескольких узлах сетки $x_n, x_{n-1}, \dots, x_{n-p}$. Следовательно, в этих точках отрезка известно и значение $f(x_i, u_i)$ правой части дифференциального уравнения (3.14) при $i = n, n-1, \dots, n-p$, причем $f(x, u(x))$ будет уже функцией только одной переменной $f(x, u(x)) = F(x)$. Заменим функцию $F(x)$ интерполяционным многочленом Лагранжа $L_p(x)$ и вычислим значение u_{n+1} , проинтегрировав (3.14) на отрезке $[x_n, x_{n+1}]$. Находим

$$u_{n+1} = u_n + \int_{x_n}^{x_{n+1}} L_p(x) dx. \quad (3.32)$$

Проводя интегрирование, находим разностную схему для решения дифференциального уравнения. Порядок схемы определяется величиной остаточного члена интерполяционного полинома.

В случае, когда для построения интерполяционного многочлена используются четыре узла $x_n, x_{n-1}, x_{n-2}, x_{n-3}$, получается ФОРМУЛА АДАМСА, которая на сетке с постоянным шагом записывается в виде

$$u_{n+1} = u_n + \frac{h}{24} (55F_n - 59F_{n-1} + 37F_{n-2} - 9F_{n-3}) \quad (3.33)$$

с локальной ошибкой дискретизации

$$\epsilon_n \leq \frac{251}{750} h^5 u^{(5)}(x).$$

Таким образом, метод Адамса имеет четвертый порядок точности на интервале. Чтобы начать счет по схеме Адамса, необходимо знать решение в четырех начальных точках x_0, x_1, x_2, x_3 .

По существу, интерполяционный многочлен $L_p(x)$ в формуле (3.32) используется вне области интерполяции, т. е. в данном случае это экстраполяционный многочлен. Однако, поскольку интервал (x_n, x_{n+1}) мал, ошибка за счет экстраполяции невелика. Недостающие значения функции вычисляются в точках x_1, x_2, x_3 , как правило, по методу Рунге — Кутты соответствующего порядка. Это является недостатком метода, так как увеличивает объем программы для компьютера. Преимущество многошаговых методов заключается в том, что на каждом шаге правая часть

ÁДАМС ДЖОН КАУЧ (Adams John Couch; 1819—1892) — английский астроном и математик, автор работ по небесной механике. А. известен теоретическим открытием планеты Нептун, сделанным им в 1845 г. независимо от У. Леверье. В математике используется метод Адамса численного интегрирования обыкновенных дифференциальных уравнений.

дифференциального уравнения вычисляется только один раз, а в методе Рунге — Кутты четвертого порядка точности на каждом шаге функция $f(x, u)$ вычисляется четыре раза. Здесь, естественно, необходимо учитывать соотношение шагов двух методов, обеспечивающих заданную точность. В случае, когда правая часть дифференциального уравнения является функцией только одной переменной x , остаточный член в формуле Рунге — Кутты четвертого порядка точности на равномерной сетке равен остаточному члену ϵ_n формулы трапеций численного интегрирования с шагом $h/2$, составляющему

$$\epsilon_n \leq \frac{b-a}{2880} h^4 M_4.$$

В формуле Адамса остаточный член

$$\epsilon_n \leq \frac{b-a}{3} h^4 M_4.$$

Таким образом, остаточный член в формуле Рунге — Кутты примерно на три порядка меньше, чем в формуле Адамса. Следовательно, для получения результата с одной и той же точностью можно шаг в методе Рунге — Кутты брать примерно в шесть раз большим, чем в формуле Адамса. Поэтому в методе Рунге — Кутты функция $f(x, u(x))$ вычисляется меньшее число раз.

Достоинство метода Адамса по сравнению с методом Рунге — Кутты заключается в простоте оценки остаточного члена метода.

3.6. Неявные схемы. Понятие о жестких системах

Рассмотрим следующую разностную схему для численного решения уравнения (3.14) при условии, что вычислено значение функции u в точке x_n . Имеем

$$u_{n+1} = u_n + h[sf(x_n, u_n) + (1-s)f(x_{n+1}, u_{n+1})], \quad (3.34)$$

где s — параметр разностной схемы такой, что $0 \leq s \leq 1$. Говорят, что разностная схема является *явной*, если правая часть дифференциального уравнения записывается в известной точке (x_n, u_n) . Схема называется *неявной*, если правая часть дифференциаль-

го уравнения записывается в неизвестной точке (x_{n+1}, u_{n+1}) . Схема называется *полуявной* (полунеявной), если правая часть с определенными весами записывается в известной и неизвестной точках. В соответствии с этими терминами схема (3.34) является явной при $s = 1$, неявной при $s = 0$ и полуявной (полунеявной) при $0 < s < 1$. В явных схемах решение в точке x_{n+1}, u_{n+1} определяется непосредственно из алгебраических соотношений с известными коэффициентами. Все рассмотренные выше схемы методов Эйлера, Рунге — Кутты, Адамса (кроме схемы метода Эйлера с пересчетом) являются явными. В неявных и полунеявных схемах при $s \neq 1$ для нахождения решения в точке (x_{n+1}, u_{n+1}) необходимо решить в общем случае нелинейное уравнение (3.34).

Нелинейное уравнение (3.34) можно разрешить различными методами, например, методами простой итерации, дихотомии, Ньютона и т. д. В рассматриваемом случае логично использовать метод простой итерации, поскольку нелинейное уравнение (3.34) уже разрешено относительно u_{n+1} . Тогда, если номер итерации обозначить верхним индексом, итерационный процесс можно записать в виде

$$u_{n+1}^{(k+1)} = u_n + h[sf(x_n, u_n) + (1 - s)f(x_{n+1}, u_{n+1}^{(k)})]. \quad (3.35)$$

Известно, что сходимость метода простой итерации зависит от выбора начального приближения. В рассматриваемом случае в качестве $u_{n+1}^{(0)}$ можно выбрать значение u_n или какое-либо другое, например, значение, вычисленное по методу Эйлера, достаточно близкое к искомому значению u_{n+1} . Вычисления по формуле (3.35) продолжаются до получения заданной точности $\epsilon = |u^{(k+1)} - u^{(k)}|$.

Таким образом, очевидно, что при расчете на шаге явные схемы требуют значительно меньшего числа операций, чем неявные. Однако неявные схемы значительно более устойчивы, чем явные. Поэтому в неявных схемах ограничения на шаг разностной схемы, связанные с устойчивостью, как правило, менее жесткие, чем в явных. В результате может оказаться, что общее число операций при заданной точности и при расчетах на интервале интегрирования в неявных схемах меньше, чем в явных.

В конечном итоге о качестве разностной схемы следует судить по числу операций, необходимых для получения решения с заданной точностью. Этот критерий и является основным при сравнении явных и неявных схем.

| Понятие о жестких системах уравнений

Во многих задачах управления, химической кинетики, неравновесной динамики газов возникают системы уравнений, содержащих некоторые уравнения с малым параметром при старшей производной. Соответственно матрица системы дифференциальных уравнений

$$\frac{du}{dx} = Au \quad (3.36)$$

имеет существенно различные собственные числа. Последнее обстоятельство приводит к тому, что в решении содержатся быстро и медленно убывающие члены, которые нужно рассчитывать с одинаковой точностью, что накладывает значительные ограничения на шаги разностной схемы.

Обозначим через λ_k собственные числа матрицы A и через l так называемое число жесткости.

СИСТЕМА называется ЖЕСТКОЙ, если

$$\operatorname{Re} \lambda_k < 0, \quad l = \frac{\max |\operatorname{Re} \lambda_k|}{\min |\operatorname{Re} \lambda_k|} \gg 1.$$

Такое определение не пригодно, если система сводится к одному дифференциальному уравнению, так как для него $l = 1$. В данном случае полагают, что уравнение является жестким, если содержит малый параметр при старшей производной.

Покажем на примере модельного уравнения для химической кинетики, что численное решение жесткого уравненияrationально строить с использованием неявных или полуяявных разностных схем, имея в виду условия устойчивости и ограничения на шаги. Рассмотрим модельное жесткое линейное уравнение, характерное для задач химической кинетики:

$$\frac{d\alpha}{dt} = \frac{1}{\tau} (\alpha^* - \alpha), \quad (3.37)$$

где $\alpha(t)$ — концентрация химического компонента; τ — время релаксации; α^* — равновесное значение химического компонента, достигаемое при $t \rightarrow 0$. Примем здесь, что τ и α^* — постоянные и τ мало, что характерно для условий приближения системы к равновесию.

Если обозначить $\alpha_0 = \alpha(0)$, то точное решение уравнения (3.37) имеет вид

$$\alpha = (\alpha_0 - \alpha^*) e^{-\frac{t}{\tau}} + \alpha^*. \quad (3.38)$$

Заменим уравнение (3.37) разностным уравнением

$$\frac{\alpha_{n+1} - \alpha_n}{h} = \frac{s}{\tau} (\alpha^* - \alpha_n) + \frac{1-s}{\tau} (\alpha^* - \alpha_{n+1}). \quad (3.39)$$

Из соотношения (3.39) следует рекуррентное соотношение

$$\alpha_{n+1} = \frac{1 - \alpha s}{1 + \alpha(1-s)} \alpha_n + \frac{\alpha^* \alpha}{1 + \alpha(1-s)}, \quad \alpha = \frac{h}{\tau}. \quad (3.40)$$

Используя (3.40), получаем точное решение разностного уравнения:

$$\alpha_{n+1} = (\alpha_0 - \alpha^*) \left[\frac{1 - \alpha s}{1 + \alpha(1-s)} \right]^{n+1} + \alpha^*, \quad (3.41)$$

а точное решение дифференциального уравнения в точке $(n+1)$ имеет вид

$$\alpha_{n+1} = (\alpha_0 - \alpha^*) e^{-\alpha(n+1)} + \alpha^*. \quad (3.42)$$

Из сравнения выражений (3.41) и (3.42) следует, что решение разностного уравнения (3.41) стремится к точному при $\alpha \rightarrow 0$ для всех s . Однако порядок точности схемы и устойчивость решения разностного уравнения зависят от величины этого параметра. При $s = 1/2$ разностная схема является схемой второго порядка точности, а при $s = 0$ и $s = 1$ — первого. При $s = 1$ (явная схема типа Эйлера или Рунге—Кутты) и $\alpha \gg 1$ (что имеет место, когда τ мало) решение разностного уравнения существенно отличается от точного.

Действительно, при $s = 1$ и $\alpha \gg 1$ член в квадратных скобках в уравнении (3.41) равен $(-\alpha)^{n+1}$, т. е. очень велик при больших α , в то время как в точном решении соответствующий член $e^{-\alpha(n+1)}$ очень мал. Из решения (3.41) следует, что максимальный шаг интегрирования при $s = 1$ для обеспечения устойчивого счета равен 2τ . Для задач химической кинетики характерны значения времени релаксации τ порядка 10^{-6} — 10^{-9} . Поэтому явные схемы позволяют численно решать жесткие уравнения лишь с очень малым шагом $h \approx 2\tau$, что делает их абсолютно непригодными даже при использовании достаточно мощных компьютеров.

В связи с этим более предпочтительной является неявная схема с $s = 0$, которая при $\alpha \gg 1$ дает решения, близкие к точному. Действительно, при $s = 0$ член в квадратных скобках в уравнении (3.41) равен $[1/\alpha]^{n+1}$ и качественно верно отражает поведение экспоненты $e^{-\alpha(n+1)}$ в точном решении. Применение неявной схемы позволяет существенно увеличить шаг интегрирования по сравнению с явными схемами при сохранении устойчивости и необходимой точности расчета.

3.7. Постановка краевой задачи для обыкновенного дифференциального уравнения

КРАЕВАЯ ЗАДАЧА — это задача отыскания частного решения системы обыкновенных дифференциальных уравнений

$$\frac{du_i}{dx} = f_i(x, u_1, u_2, \dots, u_n), \quad i = 1, 2, \dots, n \quad (3.43)$$

с дополнительными условиями, налагаемыми на значения функций $u_i(x)$ не менее чем в двух точках отрезка $[a, b]$. Следовательно, краевая задача ставится для системы дифференциальных уравнений порядка не менее второго (или одного дифференциального уравнения порядка не ниже второго).

Свое название краевая задача получила по случаю, в котором дополнительные условия заданы на концах (краях) отрезка $[a, b]$. Естественно, дополнительные условия могут задаваться и во внутренних точках отрезка. Такие условия называются внутренними краевыми условиями. Краевые условия могут связывать между собой значения нескольких функций, производных функций или комбинаций функций и производных в одной или нескольких точках отрезка, на котором ищется решение.

Точное аналитическое решение краевой задачи удается найти крайне редко, так как для этого надо найти общее решение системы дифференциальных уравнений (3.43) и выразить из краевых условий значения входящих в него констант. Поэтому широкое распространение получили численные методы решения краевых задач таких, как метод стрельбы и разностный метод.

3.8. Метод стрельбы

Рассмотрим метод стрельбы на примере решения дифференциального уравнения второго порядка

$$u'' = f(x, u, u'), \quad x \in [a, b] \quad (3.44)$$

с краевыми условиями

$$u(a) = \alpha, \quad (3.44a)$$

$$u(b) = \beta. \quad (3.446)$$

Основная идея метода стрельбы заключается в сведении решения краевой задачи (3.44)–(3.44б) к решению серии задач Коши.

Чтобы поставить задачу Коши для уравнения (3.44), необходимо в какой-либо одной точке отрезка задать два дополнительных условия. В точке a известно одно дополнительное условие $u(a) = \alpha$. Поэтому зададим в этой точке значение производной функции $u'(a)$. Так как это значение заранее неизвестно, то зададим его равным некоторому произвольному значению η . В результате получим задачу Коши

$$u'' = f(x, u, u'), \quad (3.45)$$

$$u(a) = \alpha, \quad (3.45a)$$

$$u'(a) = \eta. \quad (3.45b)$$

Решая эту задачу Коши каким-либо численным методом, получаем ее решение $u(x, \eta)$, зависящее от η как от параметра. Так как значение η выбрано произвольно, то решение задачи Коши удовлетворяет условию краевой задачи в точке a и не удовлетворяет ее условию в точке b . Таким образом, необходимо менять параметр η таким образом, чтобы решение задачи Коши в точке b совпадало с условием (3.44б). Следовательно, решение краевой задачи сводится к нахождению корня нелинейного алгебраического уравнения

$$F(\eta) = u(b, \eta) - \beta = 0. \quad (3.46)$$

При этом функция $F(\eta)$ задана не аналитически, а в виде таблицы чисел, которая составляется при решении серии задач Коши. Решение уравнения (3.46) можно искать методом дихотомии. Делают пробные *выстрелы*, т. е. решают задачу Коши с разными значениями η до тех пор, пока среди величин $(u(b, \eta) - \beta)$ не окажется двух разных по знаку. Пара соответствующих значений $F(\eta)$ делится пополам и находится новое значение η , с которым решается задача Коши. Такая процедура повторяется до получения условия (3.45б) с необходимой точностью.

Для ускорения сходимости к корню уравнения (3.46) можно применять другие методы, например метод секущих. Для этого делают два расчета с произвольными значениями $\eta^{(0)}$ и $\eta^{(1)}$, а следующие значения η вычисляют по формуле

$$\eta^{(k+1)} = \eta^{(k)} - \frac{(\eta^{(k)} - \eta^{(k-1)})[u(b, \eta^{(k)}) - \beta]}{u(b, \eta^{(k)}) - u(b, \eta^{(k-1)})}.$$

Простота алгоритма метода стрельбы и возможность использования стандартных программ решения задачи Коши позволяют успешно применять его при решении как линейных, так и нелинейных дифференциальных уравнений.

3.9. Конечно-разностный метод решения краевых задач

Идея метода заключается в сведении краевой задачи к решению системы алгебраических уравнений путем замены производных в дифференциальном уравнении и краевых условий конечно-разностными соотношениями.

Рассмотрим идею метода на примере краевой задачи для уравнения второго порядка

$$u'' + A(x)u' + B(x)u = f(x), \quad (3.47)$$

$$u(a) = \alpha, \quad (3.47a)$$

$$u(b) = \beta. \quad (3.47b)$$

Введем на $[a, b]$ сетку $a = x_0 < x_1 < \dots < x_n = b$. Для простоты будем полагать шаг сетки постоянным. Будем искать решение не на всем отрезке $[a, b]$, а лишь в $(N - 1)$ узлах сетки. Таким образом, нам необходимо найти $(N - 1)$ неизвестных значений функции u в узлах сетки. Воспользуемся для аппроксимации производных функций формулами численного дифференцирования (см. § 2.9—2.12):

$$u' \approx \frac{u_{i+1} - u_{i-1}}{2h} + O(h^2),$$

$$u'' \approx \frac{1}{h^2}(u_{i-1} - 2u_i + u_{i+1}) + O(h^2), \quad i = 1, 2, \dots, N - 1. \quad (3.48)$$

Такую аппроксимацию можно записать для каждого внутреннего узла сетки и получить систему линейных уравнений $(N - 1)$ -го порядка, если подставить формулы (3.48) в уравнение (3.47). Два недостающих уравнения получаем из краевых условий (3.47a) и (3.47b). Введя обозначения $a_i = 1/h^2 - A_i/2h$,

$b_i = 2/h^2 - B_i$, $c_i = 1/h^2 + A_i/2h$, запишем эту линейную систему в виде

$$\begin{aligned} a_i u_{i-1} - b_i u_i + c_i u_{i+1} &= d_i, \quad i = 1, 2, \dots, N-1. \\ u_0 &= \alpha, \\ u_n &= \beta. \end{aligned} \tag{3.49}$$

Матрица этой линейной системы — трехдиагональная, и поэтому решать ее можно методом прогонки (см. § 1.5). При этом, как обычно, на прямом ходе определяются прогоночные коэффициенты, а на обратном — неизвестные u_i . Устойчивость и корректность прогонки, как известно, обеспечивается условием преобладания диагональных элементов.

Описанный метод можно применить и в случае, если в краевые условия (3.47а) и (3.47б) входят линейные комбинации функций и первых производных.

3.10. Интегральные уравнения

ИНТЕГРАЛЬНЫМ называют уравнение, в котором неизвестная функция $u(x)$ находится под знаком интеграла. Одномерное нелинейное интегральное уравнение имеет вид

$$\int_a^b K(x, \xi, u(\xi)) d\xi = F(x, u(x)), \quad a \leq x \leq b, \tag{3.50}$$

где ядро $K(x, \xi, u)$ и правая часть $F(x, u)$ — заданные функции.

К интегральным уравнениям приводят многие физические задачи. Так, задача восстановления переданного радиосигнала $u(t)$ по принятому сигналу $f(t)$ сводится к решению *интегрального уравнения типа свертки*

$$\int_0^t K(t-\tau) u(\tau) d\tau = f(t), \tag{3.51}$$

где ядро $K(t-\tau)$ зависит от свойств приемной аппаратуры и среды, через которую проходит сигнал. К интегральным уравнениям сводятся также задачи образования заданной новой фазы при конденсации и кристаллизации.

Интегральные уравнения в некоторых отношениях удобнее дифференциальных. Во-первых, интегральное уравнение содер-

жит в себе полную постановку задачи. Например, интегральное уравнение

$$u(x) = u_0 + \int_{x_0}^x f(\xi, u(\xi)) d\xi \quad (3.52)$$

эквивалентно задаче Коши для дифференциального уравнения

$$\frac{du(x)}{dx} = f(x, u), \quad u(x_0) = u_0, \quad (3.53)$$

однако для уравнения (3.52) не требуется задавать никаких дополнительных условий, начальных или граничных.

Во-вторых, в интегральных уравнениях переход от одной переменной ко многим является естественным. Так, многомерным аналогом уравнения (3.50) является уравнение

$$\int_G K(x, \xi, u(\xi)) d\xi = F(x, u(x)), \quad x = \{x_1, x_2, \dots, x_p\} \in G(x), \quad (3.54)$$

отличающееся от уравнения (3.50) только тем, что интегрирование проводится по многомерной области G . Поскольку оба уравнения не требуют дополнительных условий и полностью определяют задачу, аналогия является полной. Тем самым теоретическое обоснование постановок и методов решения одномерных задач непосредственно обобщается на случай многих переменных.

Наоборот, в дифференциальных уравнениях переход от одной переменной к нескольким, т. е. от обыкновенных дифференциальных уравнений к уравнениям в частных производных, является принципиальным усложнением, приводит к новым постановкам задач и требует новых методов для их обоснования.

Далее мы ограничимся рассмотрением одномерного уравнения (3.50) и некоторых его частных случаев.

| Линейные задачи

Лучше всего изучены уравнения, в которые неизвестная функция $u(x)$ входит линейно. Их можно записать в виде

$$u(x) - \lambda \int_a^b K(x, \xi) u(\xi) d\xi = f(x), \quad a \leq x \leq b. \quad (3.55)$$

ФРÉДГОЛЬМ ЭРИК ИВАР (Fredholm Erik Ivar; 1866—1927) — шведский математик, известный своими исследованиями по теории линейных интегральных уравнений. Им разработаны общие методы решения некоторых их видов (уравнений Фредгольма).

Такое уравнение называют УРАВНЕНИЕМ ФРЕДГОЛЬМА ВТОРОГО РОДА: ядро $K(x, \xi)$ этого уравнения определено на квадрате $a \leq x \leq b, a \leq \xi \leq b$.

Если ядро $K(x, \xi)$ отлично от нуля только на треугольнике $a \leq \xi \leq x \leq b$, т. е. $K(x, \xi) = 0$ при $x < \xi$, то уравнение (3.55) переходит в УРАВНЕНИЕ ВОЛЬТЕРРА ВТОРОГО РОДА

$$u(x) - \lambda \int_a^x K(x, \xi) u(\xi) d\xi = f(x), \quad a \leq x \leq b. \quad (3.56)$$

Это уравнение теоретически исследовать или численно решить много проще, чем уравнение Фредгольма.

Если в уравнениях (3.55) и (3.56) отбросить член $u(x)$, оставив только $u(\xi)$ под знаком интеграла, то получим уравнения Фредгольма и Вольтерра первого рода. Задачи для уравнений первого рода являются некорректно поставленными. Для уравнений второго рода задачи корректно поставлены; остановимся на этих задачах. Изложим численный метод, позволяющий получать решение одномерных задач с хорошей точностью, а двумерных — с удовлетворительной. Он рассчитан на применение компьютера, хотя оценки с небольшим числом узлов сетки можно производить вручную.

Рассмотрим одномерное нелинейное уравнение (3.50). Возьмем на отрезке $[a, b]$ какую-либо квадратную формулу*, например, линейную формулу с узлами x_n и весами c_n

$$\int_a^b \Phi(\xi) d\xi \approx \sum_{n=1}^N c_n \Phi(x_n). \quad (3.57)$$

Введем в квадрате $[a \leq x \leq b, a \leq \xi \leq b]$ сетку $[x_n, \xi_m]$, где x_n и ξ_m являются узлами формулы (3.57). Заменяя интеграл в уравнении (3.50) суммой (3.57), получим систему алгебраических уравнений для определения приближенных значений в узлах $y_n = u(x_n)$

$$\sum_{m=1}^N c_m K(x_n, x_m, y_m) = F(x_n, y_n), \quad 1 \leq n \leq N. \quad (3.58)$$

Эту систему целесообразно решать методом Ньютона.

* Нелинейные квадратурные формулы почти никогда не используются.

ВОЛЬТЕРРА ВИТО (Volterra Vito; 1860—1940) — итальянский математик. Наиболее известны его работы в областях дифференциальных уравнений с частными производными, интегральных (уравнения Вольтерра) и интегро-дифференциальных уравнений, функционального анализа и теории упругости.

Неоднородное уравнение Фредгольма (3.55) приводит к линейной неоднородной алгебраической системе

$$y_n - \lambda \sum_{m=1}^N c_m K_{nm} y_m \approx f_n, \quad 1 \leq n \leq N, \quad f_n = f(x_n). \quad (3.59)$$

Разностное решение легко определяется методом исключения Гаусса.

Уравнение Вольтерра (3.56) получают из уравнения Фредгольма (3.55), полагая $K(x, \xi) = 0$ при $x < \xi$. Алгебраическая система (3.59) становится при этом треугольной

$$y_n - \lambda \sum_{m=1}^N c_m K_{nm} y_m = f_m, \quad 1 \leq n \leq N \quad (3.60)$$

и решается обратным ходом метода Гаусса всего за $(3/2)N^2$ действий. Поэтому здесь объем вычислений остается умеренным даже при $N \approx 1000$, что позволяет проводить расчеты с очень высокой точностью.

| Выбор квадратурной формулы

Большинство задач приходится решать, используя сравнительно небольшое число узлов N . Поэтому для получения хорошей точности целесообразно выбирать квадратурные формулы высокого порядка точности, разумеется, если $K(x, \xi)$ и $f(x)$ имеют достаточно большое число непрерывных производных.

Обычно наилучшие результаты для достаточно гладких решений дают квадратурные формулы Гаусса или Гаусса — Кристоффеля; при числе узлов k их порядок точности $p = 2k$. Можно также использовать простейшую формулу трапеций, последовательно сгущая сетки вдвое от $N_1 = 2$ до $N_k = 2^k$ и уточняя решение способом Рунге; это также дает результат с порядком точности* $p = 2k$, но требует использования существенно большего числа узлов, чем в формулах Гаусса.

* Каждая лишняя сетка позволяет повысить порядок точности на 2, поскольку погрешность формулы трапеций разлагается по четным степеням шага h .

КРИСТОФЕЛЬ ЭЛЬВИН БРУНО (Christoffel Elwin Bruno; 1829—1900) — немецкий математик, известный работами в области теории функций, теории дифференциальных уравнений с частными производными, теории инвариантов алгебраических форм и теории дифференциальных квадратичных форм, где им введен (1869) символ, носящий его имя. Этот символ используется, в частности, при записи элементов тензора кривизны, который иногда называют вследствие этого тензором Римана — Кристоффеля.

Глава 4

ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ УРАВНЕНИЙ В ЧАСТНЫХ ПРОИЗВОДНЫХ

Уравнения в частных производных описывают широкий круг физических задач в таких областях, как механика сплошных сред, термодинамика, квантовая механика, электродинамика, теория упругости и многих других. Аналитическое решение уравнения в частных производных удается получить лишь в единичных случаях, и поэтому роль численных методов для решения задач, описываемых уравнениями в частных производных, особенно велика.

В главе излагаются основные методы численного решения уравнений в частных производных такие, как метод сеток, метод прямых, метод характеристик, метод установления, методы сквозного счета, а также некоторые специальные методы, нашедшие применение в последнее время.

Особое внимание уделяется методу сеток как наиболее развитому и употребительному методу решения уравнений в частных производных. При изложении метода сеток вводятся такие основные понятия, как аппроксимация, устойчивость и сходимость разностной схемы, имеющие первостепенное значение при рассмотрении любого численного метода.

Представлены также некоторые сведения из теории уравнений в частных производных.

4.1. Некоторые сведения из теории уравнений в частных производных

Постановку задач для уравнений математической физики дадим на основе классических уравнений параболического, гиперболического и эллиптического типов.

| Основные определения

Если искомая функция зависит не от одной, а от нескольких переменных, то при рассмотрении многих важных задач возникают уравнения с частными производными, связывающие между собой независимые переменные x_1, x_2, \dots, x_k , а также искомую функцию $u(x_1, x_2, \dots, x_k)$ и ее частные производные $\frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_k}$, $\frac{\partial^2 u}{\partial x_1^2}, \frac{\partial^2 u}{\partial x_1 \partial x_2}, \frac{\partial^2 u}{\partial x_2^2}, \frac{\partial^2 u}{\partial x_2 \partial x_3}$ и т. д.

В общем случае уравнение в частных производных может быть записано в виде

$$F\left(x_1, \dots, x_k, u, \frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_k}, \frac{\partial^2 u}{\partial x_1^2}, \frac{\partial^2 u}{\partial x_1 \partial x_2}, \dots, \frac{\partial^n u}{\partial x_1^{p_1} \partial x_2^{p_2} \dots \partial x_k^{p_k}}\right) = f(x_1, \dots, x_k). \quad (4.1)$$

Здесь $p_1 + p_2 + \dots + p_k = n$. Наивысший порядок входящей в дифференциальное уравнение производной называется ПОРЯДКОМ ДИФФЕРЕНЦИАЛЬНОГО УРАВНЕНИЯ.

Если функция F линейна относительно функции u и ее производных, то уравнение (4.1) называется ЛИНЕЙНЫМ уравнением в частных производных. Если $f(x_1, \dots, x_k) \neq 0$, то уравнение (4.1) называется НЕОДНОРОДНЫМ, а если $f(x_1, \dots, x_k) = 0$, то ОДНОРОДНЫМ.

Если функция F линейна по высшим производным (n -го порядка), т. е. коэффициенты при высших производных могут зависеть лишь от функции u и производных до $(n - 1)$ -го порядка, то дифференциальное уравнение называется КВАЗИЛИНЕЙНЫМ.

КЛАССИЧЕСКИМ РЕШЕНИЕМ называется функция $u(x_1, \dots, x_k)$, имеющая частные производные до требуемого порядка и обращающая уравнение (4.1) в тождество. Решение уравнений с частными производными, как правило, не единственное. Если в случае обыкновенного дифференциального уравнения общее решение зависит от конечного числа параметров, то для уравнений в частных производных обычно имеется общее решение, зависящее от производных функций с числом аргументов на единицу меньше, чем у функции $u(x_1, \dots, x_k)$. Например, уравнение

$$\frac{\partial u}{\partial x_1} + \frac{\partial u}{\partial x_2} = 0$$

имеет решение, зависящее от произвольной функции

$$u(x_1, x_2) = \varphi(x_1 - x_2).$$

Таким образом, так же как и для обыкновенного дифференциального уравнения, для выделения единственного решения не-

обходится задавать дополнительные условия, в которые будут входить функции с числом аргументов $k - 1$. В качестве таких условий можно потребовать, чтобы искомая функция $u(x_1, \dots, x_k)$ принимала заданные значения на поверхности размерности $k - 1$ в пространстве переменных x_1, x_2, \dots, x_k . К этому типу условий относятся начальные и граничные условия, являющиеся наиболее часто встречающимися при постановке краевых задач математической физики.

Уравнения математической физики

Рассмотрим классические уравнения математической физики. В этих уравнениях в качестве независимых переменных присутствуют время t и пространственные координаты: декартовы, цилиндрические и сферические.

Задача называется СТАЦИОНАРНОЙ, если решение не зависит от времени, и НЕСТАЦИОНАРНОЙ или ЭВОЛЮЦИОННОЙ, если оно зависит от времени. Задачи с одной пространственной переменной называются одномерными, с двумя — двумерными, с тремя — трехмерными.

Приведем каноническую форму уравнений с двумя независимыми переменными*. Пусть

$$\begin{aligned} & A(x, y) \frac{\partial^2 u}{\partial x^2} + 2B(x, y) \frac{\partial^2 u}{\partial x \partial y} + C(x, y) \frac{\partial^2 u}{\partial y^2} + \\ & + F\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) = 0 \end{aligned} \quad (4.2)$$

— уравнение с дважды непрерывно дифференцируемыми коэффициентами A, B, C , которые не обращаются в нуль одновременно. Приведем классификацию дифференциальных уравнений в частных производных. Обозначим $D = B^2 - AC$, тогда при $D \equiv 0$ уравнение (4.2) называется ПАРАБОЛИЧЕСКИМ, при $D > 0$ — ГИПЕРБОЛИЧЕСКИМ, а при $D < 0$ — ЭЛЛИПТИЧЕСКИМ дифференциальным уравнением. Тип одного и того же уравнения может меняться в зависимости от области пространства.

Классическим уравнением гиперболического типа является волновое уравнение, параболического типа — уравнение теплопроводности, а эллиптического типа — уравнение Лапласа.

В математической физике различают задачу Коши (или задачу с начальными данными); краевые или граничные задачи; смешанные краевые задачи или краевые (граничные) задачи с на-

* В общем случае для более чем двух независимых переменных таких канонических форм не существует.

чальными данными. Задачи, в которых в качестве переменной выступает время t , иногда называют эволюционными.

В теории уравнений в частных производных важную роль играют характеристические поверхности, которые определяются как поверхности, в окрестности которых решения задачи Коши не существует, либо решение это не единственно. Можно показать, что в действительной плоскости характеристические поверхности существуют для гиперболических и параболических уравнений и не существуют для эллиптических уравнений.

Общая теория характеристик позволяет сформулировать для гиперболических уравнений некоторые специфические задачи такие, как задача Гурса, смешанная задача и ряд других. Их решение можно построить, используя как аналитический, так и численный методы характеристик.

Постановка задач для уравнений параболического типа

Классическим примером уравнения параболического типа является УРАВНЕНИЕ ТЕПЛОПРОВОДНОСТИ или ДИФФУЗИИ в области $[0, l]$ на отрезке времени $[0, t]$ (рис. 4.1)



Рис. 4.1

$$\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < l, \quad (4.3)$$

где a — коэффициент теплопроводности (если u — температура) и массопроводности (если u — концентрация, давление в задачах фильтрации и т. п.). Поскольку в уравнение (4.3) входит производная по времени, то необходимо задавать начальные условия при $t = 0$ и граничные условия при $x = 0$, $x = l$, $t > 0$.

Для уравнения теплопроводности при краевых условиях

$$u(0, t) = \varphi_1(t), \quad t > 0,$$

$$u(l, t) = \varphi_2(t), \quad t > 0$$

и начальном условии

$$u(x, 0) = \psi(x), \quad 0 \leq x \leq l$$

ГУРСА ЭДУАРД ЖАН БАТИСТ (Goursat Edouard Jean Baptiste; 1858—1936) — французский математик. Ему принадлежат труды по дифференциальным уравнениям с частными производными 2-го порядка и по теории аналитических функций. Г. поставлена задача интегрирования дифференциальных уравнений при граничных условиях, данных вдоль характеристик (задача Гурса).

имеем ПЕРВУЮ НАЧАЛЬНО-КРАЕВУЮ или КРАЕВУЮ ЗАДАЧУ С НАЧАЛЬНЫМИ ДАННЫМИ — иначе, задачу Коши с краевыми (граничными) условиями.

Для уравнения теплопроводности с краевыми условиями

$$\frac{\partial u}{\partial x} = \varphi_1(t), \quad x = 0, t > 0,$$

$$\frac{\partial u}{\partial x} = \varphi_2(t), \quad x = l, t > 0$$

и начальными данными

$$u(x, 0) = \psi(x), \quad 0 \leq x \leq l$$

имеем ВТОРОЮ КРАЕВУЮ ЗАДАЧУ С НАЧАЛЬНЫМИ ДАННЫМИ.

Для уравнения теплопроводности с краевыми условиями

$$\alpha_0 \frac{\partial u}{\partial x} + \beta_0 u = \varphi_1(t), \quad x = 0, t > 0,$$

$$\alpha_1 \frac{\partial u}{\partial x} + \beta_1 u = \varphi_2(t), \quad x = l, t > 0$$

и начальными условиями

$$u(x, 0) = \psi(x)$$

имеем ТРЕТЬЮ КРАЕВУЮ ЗАДАЧУ С НАЧАЛЬНЫМИ ДАННЫМИ.

В бесконечной области ставится следующая ЗАДАЧА КОШИ для уравнения теплопроводности:

$$u = \varphi(x), \quad -\infty < x < +\infty, t = 0.$$

Постановка задач для уравнений гиперболического типа

Классическим примером уравнения гиперболического типа является ВОЛНОВОЕ УРАВНЕНИЕ

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < l, t > 0 \quad (4.4)$$

или более общее, описывающее малые продольные колебания стержня и поперечные колебания струны. Здесь функция $u(x, t)$ характеризует отклонение струны от положения равновесия, а параметр c — скорость распространения возмущений. К волновому уравнению сводятся также уравнения распространения малых акустических возмущений.

Поскольку в уравнение (4.4) входит вторая производная по времени, старший порядок производной по времени в начальных

условиях не превышает единицы. Тогда первая, вторая и третья начально-краевые или краевые задачи с начальными данными для волнового уравнения будут иметь соответственно следующий вид:

$$\begin{cases} u(0, t) = \varphi_1(t), & t > 0, \\ u(l, t) = \varphi_2(t), & t > 0, \\ u(x, 0) = \psi_1(x), & 0 \leq x \leq l, \\ \frac{\partial u(x, 0)}{\partial t} = \psi_2(x), & 0 \leq x \leq l; \end{cases}$$

$$\begin{cases} \frac{\partial u}{\partial x} = \varphi_1(t), & x = 0, t > 0, \\ \frac{\partial u}{\partial x} = \varphi_2(t), & x = l, t > 0, \\ u(l, t) = \psi_1(t), & t > 0, \\ u(x, 0) = \psi_2(x), & 0 \leq x \leq l; \end{cases}$$

$$\begin{cases} \alpha_0 \frac{\partial u}{\partial x} + \beta_0 u = \varphi_1(t), & x = 0, t > 0, \\ \alpha_1 \frac{\partial u}{\partial x} + \beta_1 u = \varphi_2(t), & x = l, t > 0, \\ u(x, 0) = \psi_1(x), & 0 \leq x \leq l, \\ \frac{\partial u(x, 0)}{\partial t} = \psi_2(x), & 0 \leq x \leq l. \end{cases}$$

В бесконечной области можно поставить задачу Коши:

$$\begin{cases} u(x, 0) = \psi_1(x), & -\infty < x < +\infty, \\ \frac{\partial u(x, 0)}{\partial t} = \psi_2(x), & -\infty < x < +\infty. \end{cases}$$

ПУАССОН СИМЕОН ДЕНИ (Poisson Siméon Denis; 1781—1840) — французский математик и механик, труды которого относятся к небесной механике, математике и математической физике. Получили известность его исследования об устойчивости движения планет Солнечной системы. В работах по теории притяжения П. вывел уравнение, носящее его имя. Собственно в математике П. обращался к определенным интегралам (интеграл Пуассона), уравнениям в конечных разностях (формула суммирования Пуассона), теории дифференциальных уравнений с частными производными, теории вероятностей. В гидромеханике П. обобщил уравнение Навье — Стокса на случай движения сжимаемой вязкой жидкости с учетом теплопередачи. Им решен ряд задач теории упругости и обобщены уравнения теории упругости на анизотропные тела. В сопротивлении материалов П. ввел коэффициент поперечной деформации μ , носящий в названии имя автора.

Метод характеристик позволяет решать задачи Коши, Гурса и смешанные при задании условий в ограниченной области.

Постановка задач для уравнений эллиптического вида

Классическими примерами уравнений эллиптического типа являются УРАВНЕНИЕ ЛАПЛАСА

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad x, y \in \Omega \quad (4.5)$$

и УРАВНЕНИЕ ПУАССОНА

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y), \quad (4.6)$$

описывающие течение идеальной (без вязкости и теплопроводности) жидкости в стационарных потоках, стационарное распределение температуры, стационарное распределение напряженности электрического или магнитного полей. При этом уравнение (4.5) описывает все эти явления при отсутствии источников или стоков (нет правых частей), а уравнение (4.6) — с распределенными по области Ω источниками, задаваемыми правой частью $f(x, y)$ (рис. 4.2).

Поскольку уравнения (4.5), (4.6) — стационарные, то начальные условия не задаются; на границах же Γ расчетной области $\Omega + \Gamma$ задаются граничные условия 1, 2 и 3-го родов.

- Первая краевая задача для уравнения Лапласа (задача Дирихле):

$$u \Big|_{\Gamma} = \varphi(x, y).$$

- Вторая краевая задача для уравнения Лапласа (задача Неймана):

$$\frac{\partial u}{\partial n} \Big|_{\Gamma} = \varphi(x, y).$$

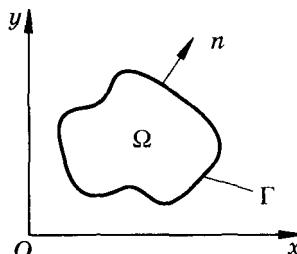


Рис. 4.2

НÉЙМАН КАРЛ ГОТФРИД (Neumann Karl Gottfried; 1832—1925) — немецкий математик, основные исследования которого относятся к теории логарифмического потенциала. Н. предложил метод (метод Неймана) решения задачи Дирихле для случая выпуклых контуров и выпуклых поверхностей. Н. исследовал вторую краевую задачу (задачу Неймана). Ему принадлежат также работы в области математической физики, теории функций Бесселя, римановой теории и абелевых интегралов.

- Третья краевая задача для уравнения Лапласа (смешанная задача):

$$\alpha \frac{\partial u}{\partial n} \Big|_{\Gamma} + \beta u \Big|_{\Gamma} = \varphi(x, y), \quad \alpha \neq 0.$$

Во второй и третьей краевой задаче производные берутся в направлении внешней нормали по отношению к границе области Ω .

Постановка многомерных задач
математической физики

В заключение приведем примеры постановок многомерных по пространственным переменным задач для уравнений параболического и эллиптического типов (рис. 4.3).

- Первая начально-краевая задача для двумерного уравнения теплопроводности:

$$\begin{cases} \frac{\partial u}{\partial t} = c \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right), & (x, y) \in \Omega, t > 0, \\ u \Big|_{\Gamma} = \varphi(x, y, t), & t > 0, \\ u(x, y, 0) = \psi(x, y), & (x, y) \in \Omega + \Gamma. \end{cases}$$

- Вторая начально-краевая задача для двумерного уравнения теплопроводности:

$$\begin{cases} \frac{\partial u}{\partial t} = c \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right), & (x, y) \in \Omega, t > 0, \\ \frac{\partial u}{\partial n} \Big|_{\Gamma} = \varphi(x, y, t), & t > 0, \\ u(x, y, 0) = \psi(x, y), & (x, y) \in \Omega + \Gamma. \end{cases}$$

- Третья начально-краевая задача для двумерного уравнения теплопроводности:

$$\begin{cases} \frac{\partial u}{\partial t} = c \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right), & (x, y) \in \Omega, t > 0, \\ \alpha \frac{\partial u}{\partial n} \Big|_{\Gamma} + \beta u \Big|_{\Gamma} = \varphi(x, y, t), & t > 0, \\ u(x, y, 0) = \psi(x, y), & (x, y) \in \Omega + \Gamma. \end{cases}$$

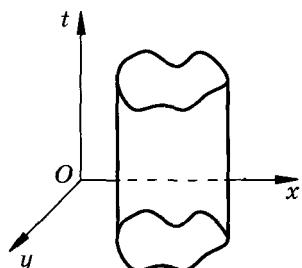


Рис. 4.3

- Первая начально-краевая задача для двумерного волнового уравнения:

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} = c \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right), & (x, y) \in \Omega, t > 0, \\ u \Big|_{\Gamma} = \varphi(x, y, t), & t > 0, \\ u(x, y, 0) = \psi_1(x, y), & (x, y) \in \Omega + \Gamma, \\ \frac{\partial u(x, y, 0)}{\partial t} = \psi_2(x, y), & (x, y) \in \Omega + \Gamma. \end{cases}$$

- Вторая начально-краевая задача для двумерного волнового уравнения:

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} = c \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right), & (x, y) \in \Omega, t > 0, \\ \frac{\partial u}{\partial n} \Big|_{\Gamma} = \varphi(x, y, t), & t > 0, \\ u(x, y, 0) = \psi_1(x, y), & (x, y) \in \Omega + \Gamma, \\ \frac{\partial u(x, y, 0)}{\partial t} = \psi_2(x, y), & (x, y) \in \Omega + \Gamma. \end{cases}$$

- Третья начально-краевая задача для двумерного волнового уравнения:

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} = c \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right), & (x, y) \in \Omega, t > 0, \\ \alpha \frac{\partial u}{\partial n} \Big|_{\Gamma} + \beta u \Big|_{\Gamma} = \varphi(x, y, t), & t > 0, \\ u(x, y, 0) = \psi_1(x, y), & (x, y) \in \Omega + \Gamma, \\ \frac{\partial u(x, y, 0)}{\partial t} = \psi_2(x, y), & (x, y) \in \Omega + \Gamma. \end{cases}$$

4.2. Основные понятия метода сеток. Задача Дирихле для уравнения Лапласа

Метод сеток (или метод конечных разностей) сводит решение систем уравнений в частных производных к решению систем линейных, как правило, алгебраических уравнений с достаточно разре-

женными матрицами. При этом построение решения в методе сеток осуществляется в три этапа.

- Область непрерывного изменения аргумента (или аргументов) заменяется конечным дискретным множеством точек, называемых разностной сеткой. В разностной сетке выделяются внутренние и граничные узлы. Решение ищется во внутренних узлах, а в граничных узлах значение искомой функции задается при аппроксимации граничных условий исходной дифференциальной задачи. Функция дискретного аргумента, определенная на разностной сетке, называется СЕТОЧНОЙ ФУНКЦИЕЙ.

- Дифференциальное уравнение и граничные условия заменяются по определенным правилам своими разностными аналогами. Разностные операторы, соответствующие дифференциальному уравнению, записываются во внутренних узлах сетки. Разностные операторы, соответствующие граничным условиям, записываются в граничных узлах. В результате получается система алгебраических уравнений, число которых пропорционально числу внутренних узлов разностной сетки.

- Осуществляется решение системы алгебраических уравнений каким-либо из известных методов. В большинстве случаев получаемая система уравнений есть система линейных алгебраических уравнений достаточно большого порядка, но с весьма разреженной матрицей. В случае нелинейных систем итерационные процедуры, как правило, сводят их к линейным системам.

Основные понятия и этапы метода сеток продемонстрируем на примере решения задачи Дирихле для классического уравнения эллиптического типа — уравнения Лапласа. Итак, имеем уравнение Лапласа

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \quad (4.7)$$

и граничное условие

$$u(x, y) \Big|_{\Gamma} = u_0(x, y), \quad (4.8)$$

где Γ — граница области G (рис. 4.4), в которой ищется решение $u(x, y)$, удовлетворяющее уравнению (4.7) и граничному условию (4.8).

На первом этапе метода сеток область G непрерывного изменения аргумента с границей Γ заменяют приближающей ее сеточной областью G_h с границей Γ_h . Для этого проведем линии $x = \text{const}$ и $y = \text{const}$ так, что $x = mh_x$ при $m = 0, \dots, M$ и $y = nh_y$ при $n = 0, \dots, N$. Величины h_x и h_y , называемые шагами разностной сетки, в общем случае могут быть различными.

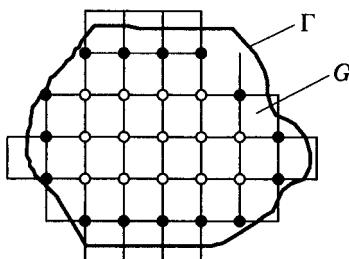


Рис. 4.4

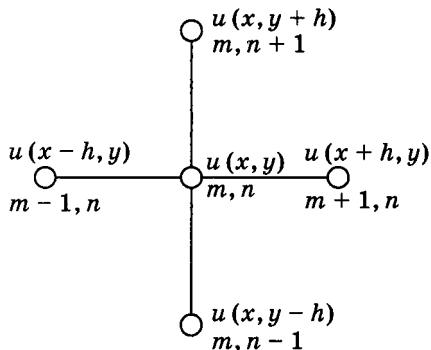


Рис. 4.5

Точки пересечения линий $x = \text{const}$ и $y = \text{const}$ называют УЗЛАМИ РАЗНОСТНОЙ СЕТКИ. Различают два типа узлов — ВНУТРЕННИЕ и ВНЕШНИЕ. Внутренними называются такие узлы, для которых четыре соседних узла (по два в каждом направлении) принадлежат области $\bar{G} = G + \Gamma$. Сеточной функции приписываются нижние индексы. На рис. 4.4 внутренние узлы обозначены пустыми кружочками, а граничные — заливочными.

На втором этапе заменим дифференциальный оператор Лапласа разностным оператором. С этой целью выберем шаблон разностной схемы — набор (конфигурацию) узлов, с использованием которых производится замена производных конечными разностями. Шаблон, содержащий p точек, называется p -точечным. Для аппроксимации вторых производных, входящих в оператор Лапласа, применим пятиточечный шаблон, показанный на рис. 4.5. Не теряя общности рассуждений, рассмотрим далее случай, когда шаги разностной сетки по направлению x и y одинаковы, т. е. $h_x = h_y = h$. С использованием разложения функции на точном решении в ряд Тейлора в окрестности точки (x, y) имеем

$$u(x \pm h, y) = u(x, y) \pm h \frac{\partial u}{\partial x} + \frac{h^2}{2!} \frac{\partial^2 u}{\partial x^2} \pm \frac{h^3}{3!} \frac{\partial^3 u}{\partial x^3} + O(h^4),$$

$$u(x, y \pm h) = u(x, y) \pm h \frac{\partial u}{\partial y} + \frac{h^2}{2!} \frac{\partial^2 u}{\partial y^2} \pm \frac{h^3}{3!} \frac{\partial^3 u}{\partial y^3} + O(h^4).$$

Из этих соотношений получим аппроксимации вторых производных со вторым порядком точности

$$\frac{\partial^2 u}{\partial x^2} = \frac{u(x+h, y) - 2u(x, y) + u(x-h, y)}{h^2} + O(h^2),$$

$$\frac{\partial^2 u}{\partial y^2} = \frac{u(x, y+h) - 2u(x, y) + u(x, y-h)}{h^2} + O(h^2).$$

Введем для сокращения записи общепринятые индексные обозначения узлов в соответствии с рис. 4.5 и соответствующие обозначения сеточных функций в узлах: $u_{m,n}$, $u_{m-1,n}$ и т. д. Тогда соотношения для производных примут вид

$$\frac{\partial^2 u}{\partial x^2} = \frac{u_{m+1,n} - 2u_{m,n} + u_{m-1,n}}{h^2} + O(h^2), \quad (4.9)$$

$$\frac{\partial^2 u}{\partial y^2} = \frac{u_{m,n+1} - 2u_{m,n} + u_{m,n-1}}{h^2} + O(h^2). \quad (4.10)$$

При этом разностный оператор Лапласа Δu_h на пятиточечном шаблоне может быть выписан как

$$\Delta u_h = \frac{u_{m+1,n} + u_{m-1,n} + u_{m,n+1} + u_{m,n-1} - 4u_{m,n}}{h^2} + O(h^2). \quad (4.11)$$

Аппроксимация граничных условий осуществляется с помощью различных приемов. Для случая прямоугольной области с границами, совпадающими с линиями $x = \text{const}$ и $y = \text{const}$, граничные значения известны в узлах точно, так как $u_{0,h} = u_0(mh_x, nh_y)$. Для области произвольной формы значения сеточной функции в граничных узлах принимаются равными значениям функции u_0 в ближайших (по какому-либо критерию) точках кривой Γ (рис. 4.4). Для задачи Неймана или смешанной задачи, когда на границе области задается значение производной функции u , производится аппроксимация производной.

Записав разностный оператор (4.11) во всех внутренних узлах, получим систему сеточных уравнений для нахождения значений сеточной функции $u_{m,n}$ в узлах разностной сетки.

На третьем этапе осуществляется решение системы сеточных уравнений. В общем случае сеточные уравнения могут быть нелинейными. Однако в рассматриваемой задаче, в силу линейности основного дифференциального уравнения Лапласа, система уравнений (4.11) является системой линейных алгебраических уравнений для нахождения значений неизвестных во внутренних узлах. При этом число уравнений системы точно равно числу внутренних узлов разностной сетки. Отметим, что число уравнений может быть весьма велико. Так, для нахождения достаточно точного решения рассматриваемой задачи требуется задать числа $N = M$ порядка 50—100, и, следовательно, количество уравнений достигает нескольких сотен или даже тысяч.

Перепишем систему сеточных решений в виде

$$u_{m+1,n} + u_{m-1,n} + u_{m,n+1} + u_{m,n-1} - 4u_{m,n} = 0. \quad (4.12)$$

Отметим, что каждое из уравнений системы содержит лишь пять неизвестных $u_{m+1,n}, u_{m-1,n}, u_{m,n+1}, u_{m,n-1}, u_{m,n}$, хотя в полной системе таких неизвестных содержится порядка N^2 . Таким образом, матрица системы (4.12) является сильно разреженной. Системы с такими матрицами хорошо решаются с помощью итерационных методов: простой итерации, Зейделя, релаксационных. Рассмотрим их применение к решению системы (4.12).

Запишем систему (4.12) в виде

$$u_{m,n} = \frac{1}{4} \left(u_{m+1,n} + u_{m-1,n} + u_{m,n+1} + u_{m,n-1} \right). \quad (4.13)$$

Значение функции в центральной точке есть среднее арифметическое значений функции в четырех соседних по направлениям x и y узлах. Уравнение (4.13) можно интерпретировать как разностный аналог теоремы о среднем для гармонических функций.

Рассмотрим применение метода простой итерации для решения системы (4.13). Для простоты предположим, что область G имеет прямоугольную форму (рис. 4.6). Построим итерационный процесс

$$u_{m,n}^{(k)} = \frac{1}{4} \left(u_{m+1,n}^{(k-1)} + u_{m-1,n}^{(k-1)} + u_{m,n+1}^{(k-1)} + u_{m,n-1}^{(k-1)} \right). \quad (4.14)$$

Здесь верхний индекс означает номер итерации. В граничных точках значения известны точно. Для нахождения решения по соотношению (4.14) необходимо знать начальное приближение — значения функции на нулевой итерации $u_{m,n}^{(0)}$. Можно доказать, что для любого шага разностной сетки итерационный процесс (4.14) сходится независимо от начального приближения $u_{m,n}^{(0)}$. На практике желательно выбирать начальное приближение возможно более близким к точному решению. Например, на основании теоремы о среднем для гармонических функций в качестве *хорошего* начального приближения можно принять значение функции, полученное интерполяцией на область G значений

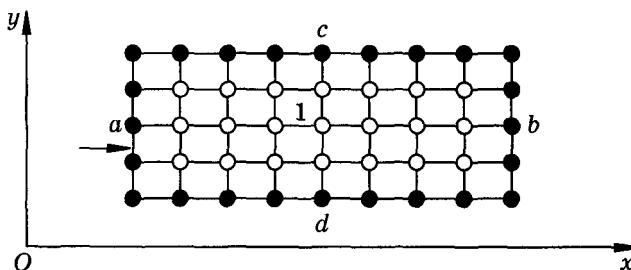


Рис. 4.6

функции в граничных узлах. Так, значение функции $u^{(0)}$ в центральной точке 1 (рис. 4.6) может быть взято как среднее арифметическое значений функции в граничных узлах a, b, c, d .

Сходимость итерационного процесса можно улучшить, если использовать итерационный процесс Зейделя в виде, например,

$$u_{m,n}^{(k)} = \frac{1}{4} \left(u_{m+1,n}^{(k-1)} + u_{m-1,n}^{(k-1)} + u_{m,n+1}^{(k-1)} + u_{m,n-1}^{(k-1)} \right). \quad (4.15)$$

Члены с индексами $(m-1, n)$ и $(m, n-1)$ берутся из k -й итерации, т. е. в вычислениях используются уточненные значения функции в этих точках. Как легко понять из рис. 4.6, расчет по формуле (4.15) при движении, например, слева направо от границы области позволяет последовательно вычислить значения функции на слоях $y = \text{const}$ (или $x = \text{const}$). Использование уточненных значений функции улучшает сходимость метода итераций и позволяет уменьшить требуемый для реализации метода объем оперативной памяти компьютера, так как в методе Зейделя не требуется одновременно хранить значение функции в каждой точке на двух итерациях. Итерации заканчиваются при выполнении условия $\max_{m,n} |u_{m,n}^{(k)} - u_{m,n}^{(k-1)}| \leq \varepsilon$, где ε — заданная погрешность определения решения.

В настоящее время распространены релаксационные методы решения системы (4.13), в которых итерационный процесс строится по формуле

$$u_{m,n}^{(k)} = \omega u_{m,n}^{(k-1)} + \frac{1-\omega}{4} \left(u_{m+1,n}^{(k-1)} + u_{m-1,n}^{(k-1)} + u_{m,n+1}^{(k-1)} + u_{m,n-1}^{(k-1)} \right),$$

в которой ω — параметр релаксации. В методе релаксации член $\omega u_{m,n}^{(k-1)}$ учитывает информацию о величине $u_{m,n}$ из предыдущей итерации. При $\omega > 1$ итерационный метод называется методом ВЕРХНЕЙ РЕЛАКСАЦИИ, при $\omega = 1$ — методом ПОЛНОЙ РЕЛАКСАЦИИ и при $\omega < 1$ — методом НИЖНЕЙ РЕЛАКСАЦИИ.

Можно показать, что число итераций I , необходимых для достижения заданной точности в рассматриваемой задаче, определяется соотношениями, приведенными в табл. 4.1.

Таблица 4.1

Название метода	Число итераций I
Метод простой итерации	$0,2N^2 \ln \varepsilon^{-1}$
Метод итераций Зейделя	$0,1 N^2 \ln \varepsilon^{-1}$
Метод верхней релаксации	$0,7N \ln \varepsilon^{-1}$

Видно, что в методе верхней релаксации число итераций для достижения заданной точности линейно зависит от числа узлов N , тогда как в остальных методах зависимость квадратичная.

4.3. Явные и неявные разностные схемы

При решении задач для уравнений параболического и гиперболического типов используются различные разностные схемы, среди которых важное место занимают так называемые явные и неявные разностные схемы.

Рассмотрим их на примере классического уравнения параболического типа — уравнения теплопроводности (нестационарного уравнения Лапласа)

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \quad (4.16)$$

со следующими начальными

$$u(x, y, 0) = u_0(x, y) \quad (4.17)$$

и граничными

$$u(x, y, t) \Big|_{\Gamma} = u^0(x, y, t) \quad (4.18)$$

условиями. Решение ищется внутри полубесконечного цилиндра (рис. 4.7). Задачу (4.16)–(4.18) называют эволюционной, имея в виду построение эволюции решения во времени. Используются также названия: ЗАДАЧА КОШИ С ГРАНИЧНЫМИ УСЛОВИЯМИ или КРАЕВАЯ ЗАДАЧА С НАЧАЛЬНЫМИ ДАННЫМИ.

Построим разностную схему. Для этого проведем линии

$$x = mh_x, \quad m = 0, \dots, M,$$

$$y = nh_y, \quad n = 0, \dots, N,$$

$$t = k\tau, \quad k = 0, \dots, K,$$

где h_x и h_y — шаги по пространственным координатам x и y соответственно и τ — шаг по времени t . Плоскости $k = \text{const}$ называют t -слоями или слоями по времени. Сеточную функцию

$$u(x + mh_x, y + nh_y, t + k\tau)$$

обозначим через $u_{m,n}^k$.

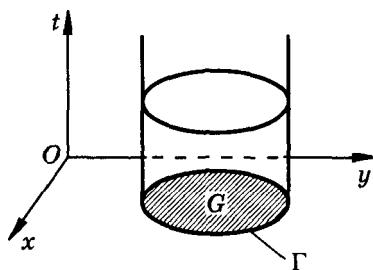


Рис. 4.7

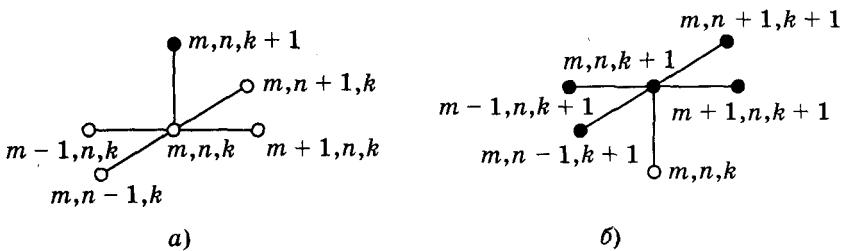


Рис. 4.8

Запишем разностную аппроксимацию уравнения (4.16) с использованием шеститочечных шаблонов двух типов, показанных на рис. 4.8. Здесь пустыми кружочками обозначены узлы на k -м слое с известными значениями функции, а зачеркнутыми — узлы на $(k+1)$ -м слое, на котором параметры неизвестны и должны быть определены. Шаблон на рис. 4.8, *a* соответствует так называемой *явной схеме*, а шаблон на рис. 4.8, *б* — *неявной схеме*. В явной схеме оператор Лапласа аппроксимируется с использованием известных значений функции на k -м слое, а в неявной схеме — с использованием неизвестных значений функции на $(k+1)$ -м слое.

Без потерь общности будем далее полагать, что шаги по пространственным координатам одинаковы, т. е. $h_x = h_y = h$. Тогда явная разностная схема, соответствующая шаблону на рис. 4.8, *а*, записывается следующим образом:

$$\begin{aligned} \frac{u_{m,n}^{k+1} - u_{m,n}^k}{\tau} &= \\ &= \frac{u_{m+1,n}^k + u_{m-1,n}^k + u_{m,n+1}^k + u_{m,n-1}^k - 4u_{m,n}^k}{h^2} + O(\tau, h^2). \end{aligned}$$

Из этого соотношения следует, что искомое значение определяется явным образом через известные значения на k -м слое по соотношению

$$\begin{aligned} u_{m,n}^{k+1} &= u_{m,n}^k + \frac{\tau}{h^2} (u_{m+1,n}^k + u_{m-1,n}^k + u_{m,n+1}^k + \\ &+ u_{m,n-1}^k - 4u_{m,n}^k). \end{aligned} \quad (4.19)$$

Для неявной разностной схемы, соответствующей шаблону на рис. 4.8, *б*, имеем

$$\frac{u_{m,n}^{k+1} - u_{m,n}^k}{\tau} = \frac{u_{m+1,n}^{k+1} + u_{m-1,n}^{k+1} + u_{m,n+1}^{k+1} + u_{m,n-1}^{k+1} - 4u_{m,n}^{k+1}}{h^2}. \quad (4.20)$$

Соотношение (4.20), записанное для всех внутренних узлов $(k+1)$ -го слоя, порождает систему линейных алгебраических

уравнений, с помощью которых определяются неизвестные значения функции в узлах. Каждое уравнение этой системы содержит только пять неизвестных $u_{m+1,n}^{k+1}$, $u_{m-1,n}^{k+1}$, $u_{m,n+1}^{k+1}$, $u_{m,n-1}^{k+1}$, $u_{m,n}^{k+1}$, хотя общее количество неизвестных равно числу внутренних узлов на $(k+1)$ -м слое и может быть весьма велико (число узлов составляет порядка N^2 , при этом N может принимать значения от 50 до 100). Следовательно, матрица системы линейных уравнений является пятидиагональной и сильно разреженной. Для решения этой системы используются, как правило, итерационные методы, некоторые из которых были описаны.

Очевидно, что при одинаковых шагах разностной сетки число операций, необходимых для отыскания решения на $(k+1)$ -м слое, в явных схемах значительно меньше, чем в неявных. Однако в дальнейшем будет показано, что условие устойчивости разностной схемы накладывает значительные ограничения на величины шагов в явных схемах. Поскольку в конечном счете качество разностной схемы при одинаковой требуемой точности должно оцениваться количеством операций, необходимых для получения решения на всем временном интервале, в ряде случаев неявные разностные схемы оказываются более предпочтительными, чем явные.

4.4. Апроксимация, устойчивость, сходимость разностных схем. Основные понятия

Рассмотрим некоторую задачу, определяемую дифференциальным уравнением и граничными условиями. Запишем ее в операторной форме

$$Lu = f, \quad (4.21)$$

где через L обозначен некоторый дифференциальный оператор, действующий на исковую функцию u , через f — правая часть. Примем, что оператор L включает как дифференциальные уравнения, так и граничные условия. На некоторой разностной сетке строим разностный оператор L_h , действующий на сеточную функцию u_h . Примером дифференциального оператора является оператор Лапласа Δ , а соответствующий ему разностный оператор возникает при аппроксимации вторых производных разностными отношениями (4.9) и (4.10).

Величина δf_h называется НЕВЯЗКОЙ, и при подстановке точного решения уравнения (4.21) в оператор L_h имеем

$$L_h u = f_h + \delta f_h.$$

РАЗНОСТНАЯ СХЕМА называется АППРОКСИМИРУЮЩЕЙ на решении, если норма невязки стремится к нулю при стремлении к нулю шага разностной схемы, т. е. $\|\delta f_h\| \rightarrow 0$ при $h \rightarrow 0$.

Если при этом постоянная C_p не зависит от шага h и норма невязки удовлетворяет условию $\|\delta f_h\| \leq C_p h^p$, то говорят, что разностная схема имеет p -й порядок аппроксимации или разностная схема аппроксимирует дифференциальный оператор с порядком p . Например, разностный оператор

$$L_h = \frac{u_{m+1, n} + u_{m-1, n} + u_{m, n+1} + u_{m, n-1} - 4u_{m, n}}{h^2}$$

аппроксирует оператор Лапласа $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ со вторым порядком точности.

РАЗНОСТНАЯ СХЕМА называется СХОДЯЩЕЙСЯ, если норма разности точного и приближенного решений $\|u - u_h\|$ стремится к нулю при стремлении к нулю шага разностной сетки.

Если при этом $\|u - u_h\| \leq C_q h^q$, то говорят, что разностная схема имеет q -й порядок точности или имеет место сходимость с порядком q .

Порядок аппроксимации дифференциального уравнения не всегда совпадает с порядком точности разностной схемы, так как порядок точности разностной схемы зависит как от аппроксимации дифференциального уравнения, так и от аппроксимации граничных условий.

Не всякая аппроксимирующая схема является сходящейся. В связи с этим важным является понятие устойчивости разностной схемы.

УСТОЙЧИВОСТЬЮ РАЗНОСТНОЙ СХЕМЫ называется непрерывная зависимость решения разностной задачи от правых частей и граничных условий.

Для линейного оператора схема устойчива, если $\|u_n\| \leq C \|f_n\|$, где C — постоянная, не зависящая от шага разностной сетки и входных условий f .

Поясним понятие устойчивости. Ошибки при вычислении начальных и граничных условий и правых частей уравнений из-за ошибок округления и других причин можно рассматривать как возмущения начальных и граничных условий и правых частей уравнений. Очевидно, что разностная краевая задача (или задача с начальными данными) будет корректной и устойчивой, если решение разностной краевой задачи будет незначительно изменяться при малом изменении начальных и граничных условий и правых частей, связанном со случайными ошибками. В противном случае разностная краевая задача является неустойчивой. Важно отметить, что для неустойчивых разностных схем уменьшение шага сетки не приводит к устойчивости, поскольку любые малые возмущения решения со временем неограниченно возрастают.

ТЕОРЕМА. Пусть разностная задача $L_h u_h = f_h$ аппроксимирует оператор $Lu = f$ на решении с порядком p и устойчива. Тогда решение u_h сходится к точному решению и имеет место оценка $\|u - u_h\| \leq C_p h^p$, где C_p — постоянная, не зависящая от шага h и входных данных f .

4.5. Примеры неустойчивых разностных схем

Впервые явление неустойчивости при решении уравнений в частных производных возникло при решении уравнения теплопроводности. Рассмотрим более простое уравнение — уравнение переноса, являющееся модельным для некоторых процессов в механике сплошной среды.

Рассмотрим задачу Коши для уравнения переноса

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \quad (4.22)$$

$$u(x, 0) = u_0(x).$$

Уравнение (4.22) является простейшим линейным уравнением гиперболического типа. Задача (4.22) имеет точное решение, которое строится с использованием метода характеристик. Сравним точное решение дифференциального уравнения с разностным, полученным с использованием различных схем, и дадим оценку аппроксимации, устойчивости и сходимости этих схем.

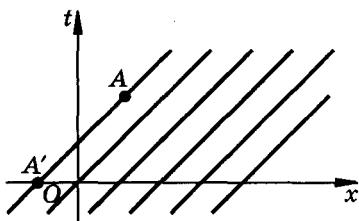


Рис. 4.9

Уравнение (4.22) обладает семейством прямолинейных характеристик $x = t + \text{const}$, которые заполняют всю плоскость (t, x) (рис. 4.9). На каждой характеристике выполняется условие совместности $u = C$, где C — величина постоянная вдоль рассматриваемой характеристики, но меняющаяся от характеристики к характеристике. Отмеченные

свойства позволяют легко построить решение задачи (4.22). Например, в некоторой точке A значение функции u_A равно значению функции u_0 в точке A' , расположенной на пересечении характеристики, выходящей из точки A , и оси абсцисс, на которой заданы данные Коши (рис. 4.9).

Далее рассмотрим некоторую специальную начальную функцию (рис. 4.10, а) вида

$$u_0(x, 0) = u_0(x) \quad \text{при } x < 0,$$

$$u_0(x, 0) = 0 \quad \text{при } x \geq 0.$$

В силу свойств начальной функции правее характеристики AB , выходящей из начала координат (рис. 4.10, б), точное решение задачи таково, что $u \equiv 0$. Левее этой характеристики точное решение определяется по значениям функции $u_0(x)$ указанным способом.

Рассмотрим теперь две разностные схемы, с помощью которых попытаемся построить численное решение задачи (4.22). Трехточечные шаблоны этих схем изображены на рис. 4.11. Шаблон, изображенный на рис. 4.11, а, называют *уголок вперед* или *правый уголок*, а шаблон, изображенный на рис. 4.11, б, — *уголок назад* или *левый уголок*. Разностные схемы на этих шаблонах являются явными, поскольку для аппроксимации пространственной производной $\partial u / \partial x$ используются известные значения функции на k -м слое.

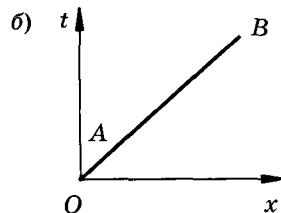
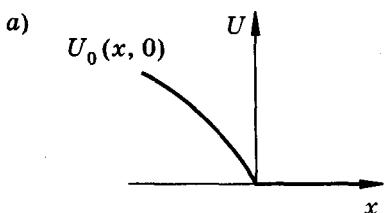


Рис. 4.10

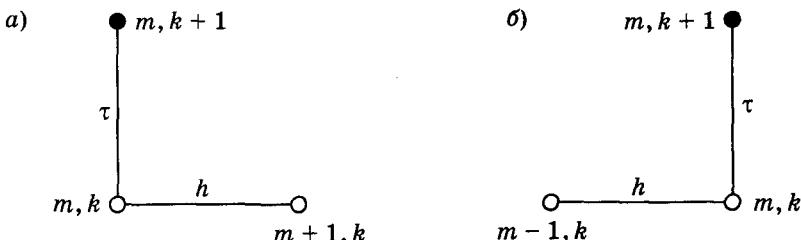


Рис. 4.11

Для шаблона, показанного на рис. 4.11, *a*, разностная схема *уголок вперед* запишется в виде

$$\frac{u_m^{k+1} - u_m^k}{\tau} + \frac{u_{m+1}^k - u_m^k}{h} = 0. \quad (4.23)$$

Так как при аппроксимации производных использовались односторонние разности, разностная схема (4.23) имеет первый порядок аппроксимации по τ и h . Из соотношения (4.23) решение на $(k+1)$ -м слое запишется в виде

$$u_m^{k+1} = u_m^k \left(1 + \frac{\tau}{h} \right) - u_{m+1}^k \frac{\tau}{h}. \quad (4.24)$$

Пусть $k = 0$, тогда в силу граничных условий $u_m^0 = u_{m+1}^0 \equiv 0$ для $x \geq 0$ и из уравнения (4.24) следует, что при всех m и $x \geq 0$ решение на слое $k = 1$ тождественно равно нулю, т. е. $u_m^1 \equiv 0$. Аналогично для всех последующих слоев $u_m^k \equiv 0$ при $x \geq 0$. Таким образом, из (4.24) следует, что в правом верхнем квадранте $u_h \equiv 0$, хотя, согласно точному решению, $u \equiv 0$ только в области, лежащей правее характеристики AB . Следовательно, разностная схема (4.23) дает решение, сколь угодно отличающееся от точного независимо от шагов разностной сетки τ и h ; т. е. при стремлении шагов τ и h к нулю не будет иметь места сходимость решения разностной задачи к точному решению. Вместе с тем разностная схема (4.23) является аппроксимирующей.

В рассмотренном случае для схемы *уголок вперед* малое изменение граничных условий при $x < 0$ приводит к сколь угодно большим различиям в решении, полученном по этой схеме в области между осью ординат и характеристикой, выходящей из начала координат. Таким образом, отсутствует непрерывная зависимость решения от граничных условий.

Для шаблона, показанного на рис. 4.11, *б*, разностная схема *уголок назад* запишется в виде

$$\frac{u_m^{k+1} - u_m^k}{\tau} + \frac{u_m^k - u_{m-1}^k}{h} = 0. \quad (4.25)$$

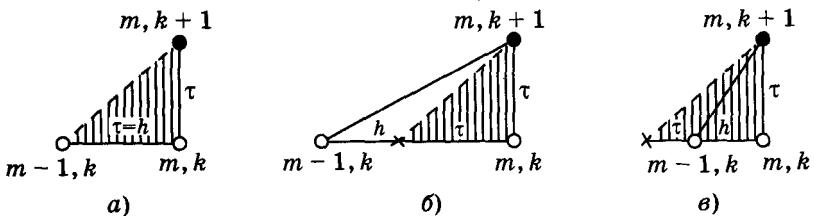


Рис. 4.12

Эта схема также имеет первый порядок аппроксимации. По соотношению (4.25) решение на $(k + 1)$ -м слое имеет вид

$$u_m^{k+1} = u_m^k \left(1 - \frac{\tau}{h}\right) + u_{m-1}^k \frac{\tau}{h}. \quad (4.26)$$

Рассмотрим три случая, соответствующие различным соотношениям шагов τ и h : $\tau/h = 1$, $\tau/h < 1$ и $\tau/h > 1$. В случае $\tau/h = 1$ (рис. 4.12, а) разностная схема дает точное решение $u_m^{k+1} = u_{m-1}^k$, поскольку значение функции по характеристике, помеченной на рисунке пунктиром, переносится из точки $(m - 1, k)$ в точку $(m, k + 1)$. Такое совпадение является, однако, свойством конкретных задач и разностных схем, и в других случаях условие $\tau = h$ может и не обеспечивать точного решения.

В случае $\tau/h < 1$ (рис. 4.12, б) правая часть формулы (4.26) является формулой линейной интерполяции функции u_h между точками $(m - 1, k)$ и (m, k) на слое k , которая дает значение функции в точке, отстоящей от точки (m, k) на расстояние τ . Это значение переносится по характеристике (пунктирная линия на рис. 4.12, б) в точку $(m, k + 1)$ и дает решение, близкое к истинному с такой точностью, какой обладает линейная интерполяция. При этом, если $u_0(x)$ — линейная функция, при $\tau/h < 1$ формула (4.26) дает точное решение.

При $\tau/h > 1$ (рис. 4.12, в) формула (4.26) является формулой линейной экстраполяции функции u_h , и в общем случае можно получить решение, существенно отличающееся от точного.

РАЗНОСТНАЯ СХЕМА называется АБСОЛЮТНО УСТОЙЧИВОЙ, если она устойчива при любом соотношении шагов. Схема называется АБСОЛЮТНО НЕУСТОЙЧИВОЙ, если она неустойчива при любом соотношении шагов. Разностная схема называется УСЛОВНО (частично) УСТОЙЧИВОЙ, если она устойчива при некотором соотношении шагов.

Проведенное рассмотрение двух схем *уголок вперед* и *уголок назад* для уравнения переноса показывает, что схема (4.23) *уголок вперед* является абсолютно неустойчивой, а схема (4.25) *уголок*

лок назад — условно устойчивой при $\tau/h \leq 1$. Условие $\tau/h < 1$ сформулируем в виде условия Куранта — Фридрихса — Леви.

| Условие Куранта — Фридрихса — Леви

Пусть область влияния разностной схемы содержит в себе область влияния дифференциального уравнения. Разностная схема является неустойчивой, если это условие не выполнено.

На рис. 4.12 область влияния дифференциального уравнения заштрихована для всех трех случаев отношения τ/h . Видно, что для случая $\tau/h > 1$ разностная схема неустойчива.

4.6. Практические правила исследования устойчивости

Установить устойчивость разностной схемы с использованием данного выше определения на практике весьма затруднительно. Поэтому предложен ряд способов исследования устойчивости, позволяющих получить достаточные, а в ряде случаев необходимые и достаточные условия устойчивости разностных схем. Из методов исследования устойчивости упомянем спектральный признак устойчивости, метод гармоник Фурье, принцип максимума.

Исследование устойчивости разностных схем, аппроксимирующих сложные и в общем случае нелинейные уравнения, про-

КУРАНТ РИХАРД (Courant Richard; 1888—1972) — немецкий математик, чьи основные труды посвящены развитию и применению принципа Дирихле к теории конформных отображений и к краевым задачам математической физики для уравнений эллиптического типа. Основные результаты изложены К. в сочинении «Методы математической физики», написанной им самим, хотя на титульном листе указан еще и его учитель — Д. Гильберт.

ФРИДРИХС КУРТ ОТТО (Friedrichs Kurt Otto; 1901—1982) — немецкий математик. Главными областями его интересов являлись математическая физика, функциональный анализ и теория дифференциальных уравнений с частными производными.

ЛЁВИ ГАНС (Lewy Hans; 1904—1988) — немецкий математик. К основным направлениям его исследований относятся общие квазилинейные уравнения второго порядка и нелинейные эллиптические и гиперболические дифференциальные уравнения.

водится следующим образом. Первоначально исходное дифференциальное уравнение заменяется более простым модельным уравнением. Это достигается либо *замораживанием* коэффициентов в уравнениях с переменными коэффициентами, либо, в случае нелинейных уравнений, путем линеаризации последних. Для полученного модельного уравнения выписывается исследуемая разностная схема и рассматривается задача Коши при возмущенных граничных условиях и возмущенных правых частях. Исследуется эволюция решения во времени, и, если возмущения с течением времени затухают, разностная схема считается устойчивой, в противном случае — неустойчивой. Действительно, возмущения граничных условий и правых частей можно рассматривать как ошибки округления или аппроксимации, неизбежно возникающие при расчетах, и, если разностная схема будет их усиливать, решение будет неустойчивым.

4.7. Спектральный признак устойчивости

Пусть решение u_m^{k+1} имеет вид $u_m^{k+1} = Su_p^k$, где $p = i, i + 1, i - 1$. Оператор перехода с одного временного слоя на другой S , действующий на вектор u_p^k , позволяет в результате последовательного применения трансформировать начальные данные в разностное решение на любом шаге по времени. Для линейных схем оператор S линеен. Так, для схемы *уголок назад* (4.25) имеем

$$u_m^{k+1} = u_m^k \left(1 + \frac{\tau}{h} \right) - u_{m-1}^k \frac{\tau}{h} = S[u_m^k, u_{m-1}^k]^T,$$

где S — матрица $\begin{vmatrix} 1 - \tau/h & 0 \\ 0 & \tau/h \end{vmatrix}$. Для линейных задач условием устойчивости по начальным данным разностной схемы является

$$\|S\| \leq 1. \quad (4.27)$$

Очевидно, $\|u^{k+1}\| = \|Su^k\| \leq \|S\| \cdot \|u^k\|$ или $\|u^{k+1}\| \leq \|u^k\| \leq \dots \leq \|u^0\|$, и, следовательно, начальные возмущения затухают. Условие (4.27) есть условие того, что спектр оператора S лежит внутри круга единичного радиуса на комплексной плоскости. Это означает, что $\max \|\lambda\| < 1$, где λ — собственное число оператора S . Достаточным условием устойчивости является $|\lambda| < 1$. В связи с этим условие (4.27) называют спектральным признаком устойчи-

вости. Сформулируем более общее условие Неймана: существует постоянная C такая, что $|\lambda| < 1 + o(\tau)$.

4.8. Принцип максимума

Перепишем уравнение (4.26) в виде $u_m^{k+1} = (1 - \gamma)u_m^k + \tau u_{m-1}^k$, где $\gamma = \tau/h$. Так как $\gamma \leq 1$, то в этом соотношении коэффициенты при значениях функций в правой части являются положительными числами и тогда соотношение (4.26) можно переписать в виде

$$\begin{aligned}|u_m^{k+1}| &= |(1 - \gamma)u_m^k + \gamma u_{m-1}^k| \leq [(1 - \gamma) + \gamma] \max_m (|u_m^k|, |u_{m-1}^k|) = \\&= \max_m (|u_m^k|, |u_{m-1}^k|) \leq \max_m |u_m^k|.\end{aligned}$$

Так как правая часть этого неравенства не зависит от значения m , то можно написать, что $\max_m |u_m^{k+1}| \leq \max_m |u_m^k|$. Применяя последовательно это неравенство, получаем неравенство

$$\max_m |u_m^{k+1}| \leq \max_m |u_m^k| \leq \dots \leq \max_m |u_m^0|,$$

называемое ПРИНЦИПОМ МАКСИМУМА для разностной схемы. Сформулируем его.

Максимальное значение модуля решения разностной задачи достигается на границе области.

Для задачи Коши, которую мы рассматривали, максимальное значение достигается при $t = 0$. Таким образом, начальные возмущения затухают. Для схемы *уголок вперед* условие устойчивости по спектральному признаку и принципу максимума не будут выполнены.

4.9. Метод гармоник Фурье исследования устойчивости разностных схем

Метод гармоник Фурье занимает одно из центральных мест среди методов исследования устойчивости разностных схем. Он используется для линейных разностных уравнений с постоянными ко-

эффициентами на равномерной сетке. В методе гармоник Фурье при изучении устойчивости разностных схем рассматривается задача Коши относительно линейного модельного уравнения для возмущений (добавок) начальных данных и правых частей. При этом рассматриваются возмущения (гармоники) специального вида.

Исследуем метод гармоник Фурье на примере задачи Коши для уравнения переноса

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0, \quad (4.28)$$

$$u(x, 0) = u_0(x).$$

Точное решение задачи можно записать в виде

$$u(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} u_0(\theta) e^{i\theta(x-t)} d\theta,$$

где $i = \sqrt{-1}$. Решение задачи (4.28) представляется как суперпозиция бегущих монохроматических волн или гармоник вида $u_0(\theta) e^{i\theta(x-t)}$. Поскольку при разностной аппроксимации $t = kt$, $x = mh$, то соответствующую разностную гармонику можно представить в виде $u_m^k = \lambda^k e^{i\omega m h}$, а решение разностного уравнения — в виде $u_m^k = \sum_{\omega=-\infty}^{+\infty} C_\omega \lambda^k e^{i\omega m h}$. Здесь λ — множитель, который выбирается таким образом, чтобы u_m^k было решением разностной задачи, а ω — номер гармоники в разложении в ряд Фурье — определяется из разложения начальной функции (начального возмущения) в рассматриваемый ряд.

Множитель λ , называемый *модулем перехода*, в общем случае может быть комплексным числом и является показателем роста или затухания соответствующей гармоники. При значении $|\lambda| > 1$ начальные возмущения возрастают со временем, а при $|\lambda| < 1$ они затухают. Очевидно, что разностная схема будет устойчивой при условии $|\lambda| < 1$, поскольку любые возмущения правых частей и

ФУРЬЕ ЖАН БАТИСТ ЖОЗЕФ (Fourier Jean Baptiste Joseph; 1768—1830) — французский математик, труды которого относятся к алгебре, методам численного решения уравнений и математической физике. Ф. вывел дифференциальное уравнение теплопроводности, разработал для решения этого уравнения метод разделения переменных (метод Фурье), в основе которого лежит представление функций тригонометрическими рядами Фурье.

начальных условий (ошибки округления, аппроксимации и т. д.) будут затухать, так как $|e^{i\omega x}| < 1$.

Дж. фон Нейманом было сформулировано УСЛОВИЕ УСТОЙЧИВОСТИ: $|\lambda| \leq 1 + O(\tau)$, которое в ряде случаев является необходимым, но не достаточным условием сходимости. Далее, однако, при изучении разностных схем будем использовать более жесткое условие: $|\lambda| \leq 1$.

4.10. Применение метода гармоник Фурье для исследования устойчивости разностных схем в уравнениях переноса

Для уравнения (4.22) рассмотрим пять разностных схем, шаблоны которых показаны на рис. 4.13: явный уголок вперед (рис. 4.13, а), явный уголок назад (рис. 4.13, б), явная схема второго порядка точности (рис. 4.13, в), явная схема Лакса (рис. 4.13, г) и неявный уголок назад (рис. 4.13, д). Светлыми точками выделены узлы, в которых решение известно, темными — узлы, в которых требуется определить решение.

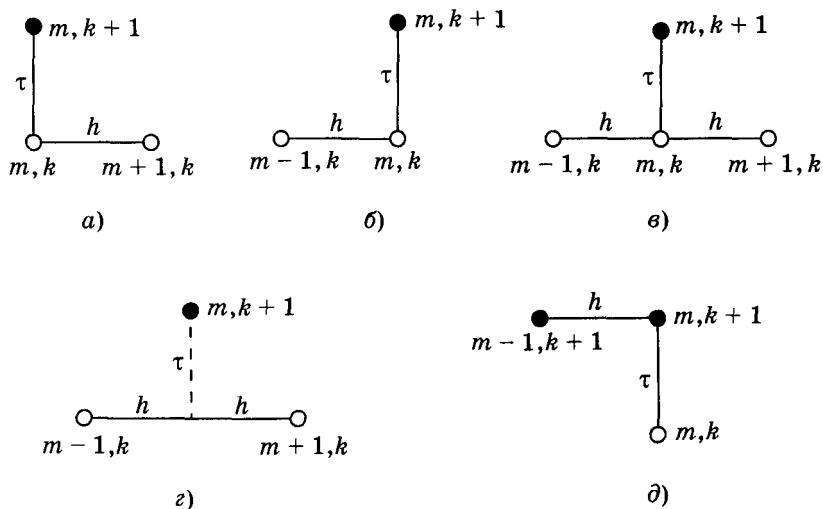


Рис. 4.13

Выпишем некоторые вспомогательные формулы:

$$\begin{cases} e^{i\omega x} = \cos(\omega x) + i \cdot \sin(\omega x), \\ 1 - \cos(\omega x) = 2 \sin^2\left(\frac{\omega x}{2}\right), \\ \sin(\omega x) = 2 \sin\left(\frac{\omega x}{2}\right) \cos\left(\frac{\omega x}{2}\right), \end{cases} \quad (4.29)$$

$$\begin{cases} u_{m-1}^{(k+1)} = \lambda^{k+1} \cdot e^{i\omega(x-h)}, \\ u_m^{(k+1)} = \lambda^{k+1} \cdot e^{i\omega(x+h)}, \\ u_m^{(k)} = \lambda^k \cdot e^{i\omega x}, \\ u_{m-1}^{(k)} = \lambda^k \cdot e^{i\omega(x-h)}, \\ u_{m+1}^{(k)} = \lambda^k \cdot e^{i\omega(x+h)}, \\ u_m^{(k+1)} = \lambda^{k+1} \cdot e^{i\omega x}. \end{cases} \quad (4.30)$$

По формуле (4.23) разностная схема для явного уголка вперед имеет вид

$$\frac{u_m^{(k+1)} - u_m^{(k)}}{\tau} + \frac{u_{m+1}^{(k)} - u_m^{(k)}}{h} = 0.$$

Преобразуем схему (4.23) с учетом соотношений (4.30) к виду

$$\frac{\lambda - 1}{\tau} + \frac{e^{i\omega h} - 1}{h} = 0$$

и с использованием формул (4.29) — к виду

$$\lambda = 1 + [1 - \cos(\omega h)] \frac{\tau}{h} - i \sin(\omega h) \frac{\tau}{h}.$$

Для квадрата модуля λ имеем соотношение

$$|\lambda|^2 = \left[1 + 2 \sin^2\left(\frac{\omega h}{2}\right) \frac{\tau}{h} \right]^2 + \sin^2(\omega h) \frac{\tau^2}{h^2},$$

из которого следует, что $|\lambda| > 1$ при любых τ и h , т. е. разностная схема (4.23) — явный уголок вперед — абсолютно неустойчива. Этот же результат был получен ранее из других соображений.

ЛАКС ПЕТЕР Д. (Lax Peter D.; 1926) — американский математик, чьи основные труды относятся к теории дифференциальных уравнений с частными производными, функциональному анализу и прикладной математике. В вычислительной математике известен метод Лакса — Вендорфа для численного решения задачи одномерного нестационарного истечения идеального газа.

Для явного уголка назад (рис. 4.13, б) разностная схема по формуле (4.25) имеет вид

$$\frac{u_m^{(k+1)} - u_m^{(k)}}{\tau} + \frac{u_m^{(k)} - u_{m-1}^{(k)}}{h} = 0.$$

Выкладки, аналогичные проведенным для явного уголка вперед, дают

$$|\lambda|^2 = 1 + 4\sin^2\left(\frac{\omega h}{2}\right) \frac{\tau(\tau/h - 1)}{h}.$$

Из этого соотношения следует, что разностная схема (4.25) — явный уголок назад, как это было показано и ранее, условно устойчива при $\tau/h \leq 1$.

Схемы (4.23) и (4.25) имеют первый порядок аппроксимации по τ и h , так как отброшены члены порядка малости $O(\tau, h)^*$.

Рассмотрим явную схему второго порядка точности, проиллюстрированную рис. 4.13, в, в которой производная по x аппроксимируется со вторым порядком точности, т. е. отброшены члены малости $O(\tau, h^2)$. Разностная схема имеет вид

$$\frac{u_m^{(k+1)} - u_m^{(k)}}{\tau} + \frac{u_{m+1}^{(k)} - u_{m-1}^{(k)}}{2h} = 0, \quad (4.31)$$

а для квадрата модуля λ может быть получено соотношение

$$|\lambda|^2 = 1 + \sin^2(\omega h) \frac{\tau^2}{h^2},$$

из которого следует, что разностная схема (4.31) абсолютно неустойчива при условии $|\lambda| < 1$. При условии $|\lambda| < 1 + O(\tau)$ схему можно рассматривать как частично устойчивую, если $\tau^2/h^2 < C\tau$. Несмотря на повышение порядка аппроксимации по сравнению со схемой (4.23), при использовании условия $|\lambda| < 1$ схема абсолютно неустойчива.

Для схемы Лакса разностная схема имеет вид

$$\frac{u_m^{(k+1)} - [u_{m+1}^{(k)} + u_{m-1}^{(k)}]/2}{\tau} + \frac{u_{m+1}^{(k)} - u_{m-1}^{(k)}}{2h} = 0. \quad (4.32)$$

Схема Лакса имеет второй порядок точности по h и является условно устойчивой при $\tau/h \leq 1$. Эта схема обычно используется в задачах механики сплошных сред и принадлежит к схемам сквозного счета, так как позволяет без явного выделения проводить расчеты областей с поверхностями разрыва.

* Здесь и далее под записью «...отброшены члены порядка малости $O(x, \mu)$ » или « $O(x, \mu, v)$ » подразумевается, что отброшенные члены стремятся к нулю при $x \rightarrow 0, \mu \rightarrow 0, v \rightarrow 0$.

Для схемы *неявный уголок назад* (рис. 4.13, *δ*) разностная схема имеет вид

$$\frac{u_m^{(k+1)} - u_m^{(k)}}{\tau} + \frac{u_{m+1}^{(k+1)} - u_m^{(k+1)}}{h} = 0. \quad (4.33)$$

В этой схеме значение $|\lambda| < 1$ при любом соотношении шагов, и, следовательно, схема абсолютно устойчива.

4.11. Применение метода гармоник Фурье в задачах исследования устойчивости разностных схем для уравнения теплопроводности

По отношению к одномерному уравнению теплопроводности

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \quad (4.34)$$

рассмотрим две разностные схемы — явную и неявную, шаблоны которых показаны соответственно на рис. 4.14, *a* и 4.14, *б*.

Для *явной схемы* имеем разностную аппроксимацию с порядком малости $O(\tau, h^2)$:

$$\frac{u_m^{(k+1)} - u_m^{(k)}}{\tau} = \frac{u_{m-1}^{(k)} - 2u_m^{(k)} + u_{m+1}^{(k)}}{h^2}. \quad (4.35)$$

Модуль перехода λ для этой схемы составляет

$$\lambda = 1 - \frac{4\tau}{h^2} \sin^2\left(\frac{\omega h}{2}\right).$$

Из этого соотношения следует, что явная схема (4.35) условно устойчива при $\tau/h^2 \leq 1/2$.

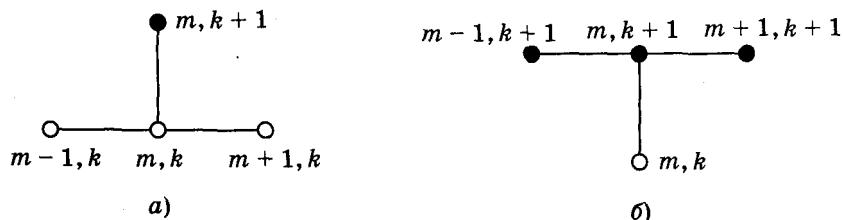


Рис. 4.14

Для неявной схемы имеем разностную аппроксимацию с порядком малости $O(\tau, h^2)$:

$$\frac{u_m^{(k+1)} - u_m^{(k)}}{\tau} = \frac{u_{m-1}^{(k+1)} - 2u_m^{(k+1)} + u_{m+1}^{(k+1)}}{h^2}. \quad (4.36)$$

Можно показать, что для этой схемы $|\lambda| < 1$ при всех τ и h и, следовательно, схема абсолютно устойчива.

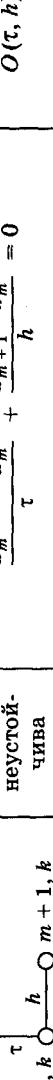
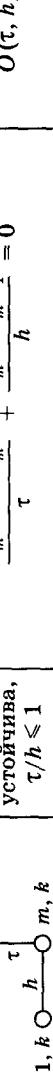
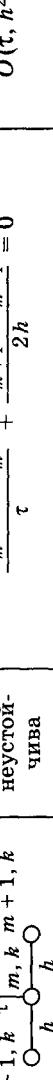
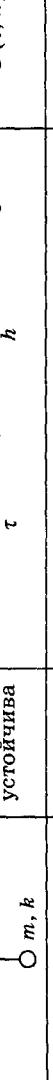
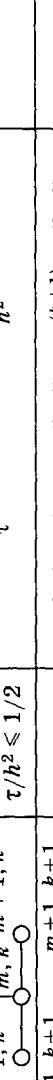
Полученные результаты исследования устойчивости различных разностных схем для уравнения переноса и одномерного уравнения теплопроводности сведены в табл. 4.2, при этом использовано условие устойчивости $|\lambda| < 1$. Из представленных результатов следует, что неявные разностные схемы более устойчивы, чем явные, и, как правило, в неявных схемах не накладываются ограничения на шаги разностной сетки, связанные с устойчивостью. Однако при использовании неявных разностных схем для определения значений неизвестных на $(k+1)$ -м слое необходимо решать систему алгебраических уравнений относительно неизвестных $u_m^{(k+1)}$ ($m = 0, \dots, M$). В то же время для явных схем значения неизвестных $u_m^{(k+1)}$ в каждой точке на $(k+1)$ -м слое явным образом выражаются через известные значения на k -м слое и их вычисление элементарно. Поэтому число операций для нахождения M значений неизвестных на одном слое в явных схемах значительно меньше, чем в неявных схемах. В то же время из-за ограничений, связанных с устойчивостью разностных схем, число шагов по времени для достижения конечного значения T может быть значительно больше в явных схемах, чем в неявных.

Таким образом, предпочтительность той или иной разностной схемы при заданной точности должна определяться с учетом общего числа операций, необходимого для полного цикла вычислений. С этой точки зрения существуют ситуации, когда может быть предпочтительной та или иная схема — явная или неявная. Например, для одномерных по пространству задач часто более выгодными оказываются неявные разностные схемы.

4.12. Метод переменных направлений

На практике приходится иметь дело с многомерными задачами, когда необходимо учитывать изменение искомых функций по координатам x, y, z . Для таких задач выбор разностной схемы

Таблица 4.2

Дифференциальное уравнение	Шаблон	Условие устойчивости схемы	Разностная схема	Порядок аппроксимации
$\frac{\partial u}{\partial t} \pm \frac{\partial u}{\partial x} = 0$		Абсолютно неустойчивая	$\frac{u_m^{(k+1)} - u_m^{(k)}}{\tau} + \frac{u_{m+1}^{(k)} - u_m^{(k)}}{h} = 0$	$O(\tau, h)$
$\frac{\partial u}{\partial t} = 0$		Условно устойчива, $\tau/h \leq 1$	$\frac{u_m^{(k+1)} - u_m^{(k)}}{\tau} + \frac{u_m^{(k)} - u_{m-1}^{(k)}}{h} = 0$	$O(\tau, h)$
$\frac{\partial^2 u}{\partial x^2} = 0$		Условно устойчива, $\tau/h \leq 1/2$	$\frac{u_m^{(k+1)} - [u_{m+1}^{(k)} - u_{m-1}^{(k)}]/2}{\tau} + \frac{u_{m+1}^{(k)} - u_{m-1}^{(k)}}{2h} = 0$	$O(\tau, h^2)$
$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$		Условно устойчива, $\tau/h \leq 1$	$\frac{u_m^{(k+1)} - u_m^{(k)}}{\tau} + \frac{u_{m+1}^{(k+1)} - u_m^{(k+1)}}{h} = 0$	$O\left(\tau, h^2, \frac{h^2}{\tau}\right)$
$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$		Условно устойчива, $\tau/h^2 \leq 1/2$	$\frac{u_m^{(k+1)} - u_m^{(k)}}{\tau} + \frac{u_{m+1}^{(k+1)} - u_{m+1}^{(k)}}{h} = 0$	$O(\tau, h^2)$
$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$		Условно устойчива, $\tau/h^2 \leq 1/2$	$\frac{u_m^{(k+1)} - u_m^{(k)}}{\tau} = \frac{u_{m-1}^{(k)} - 2u_m^{(k)} + u_{m+1}^{(k)}}{h^2}$	$O(\tau, h^2)$

является нетривиальным. Можно показать, что расчет по явной схеме до заданного момента времени T с учетом связанных с устойчивостью ограничений на шаги требует по явной схеме порядка N^{P+2} операций, а по неявной схеме — порядка N^{3P-1} операций. Здесь N — число точек на слое, а P — пространственная размерность задачи. Для рассматриваемого здесь метода переменных направлений требуется N^{P+1} операций. В табл. 4.3 представлено число операций для различных схем и различных пространственных размерностей P задачи. Из таблицы следует, что для задач с пространственной размерностью больше единицы предпочтительным является метод переменных направлений.

Таблица 4.3

P	Явная схема	Неявная схема	Метод переменных направлений
1	N^3	N^2	N^2
2	N^4	N^5	N^3
3	N^5	N^8	N^4

Метод переменных направлений рассмотрим на примере решения двумерного уравнения теплопроводности — уравнения параболического типа с двумя независимыми пространственными переменными

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + f(x, y, t). \quad (4.37)$$

Требуется найти решение этого уравнения внутри параллелепипеда $[0 \leq x \leq a, 0 \leq y \leq b, 0 \leq t \leq T]$ при начальном условии

$$u \Big|_{t=0} = u_0(x, y, 0) \quad (4.38)$$

и граничных условиях на гранях параллелепипеда

$$u \Big|_{\Gamma} = u_0(x, y, t). \quad (4.39)$$

Идея метода переменных направлений состоит в сведении решения трехмерной задачи по переменным t, x и y к решению последовательности двумерных задач по переменным t, x и t, y .

Построим разностную схему. Пусть τ, h_x и h_y — соответствен-
но шаги по t, x, y и $x_m = mh_x, y_n = nh_y, t_k = k\tau$; здесь $m = 0, \dots, M; n = 0, \dots, N; k = 0, \dots, K$. Введем разностные операторы

$$L_x[u_{m,n}^{(k)}] = -\frac{u_{m-1,n}^{(k)} - 2u_{m,n}^{(k)} + u_{m+1,n}^{(k)}}{h_x^2}, \quad m = 1, \dots, M-1,$$

$$L_y[u_{m,n}^{(k)}] = -\frac{u_{m,n-1}^{(k)} - 2u_{m,n}^{(k)} + u_{m,n+1}^{(k)}}{h_y^2}, \quad n = 1, \dots, N-1,$$

$$L[u_{m,n}^{(k)}] = L_x[u_{m,n}^{(k)}] + L_y[u_{m,n}^{(k)}], \quad k = 1, \dots, K-1.$$

Аппроксимируя уравнение (4.37) разностным, можем записать

$$\frac{u_{m,n}^{(k+1)} - u_{m,n}^{(k)}}{\tau} + (L_x + L_y)[\sigma u_{m,n}^{(k+1)} + (1-\sigma)u_{m,n}^{(k)}] = f_{m,n},$$

где σ — параметр разностной схемы, который может принимать значения от 0 до 1. Погрешность аппроксимации такой схемы $O(\tau^\nu, h_x^2, h_y^2)$, где $\nu = 2$ при $\sigma = 1/2$ и $\nu = 1$ при $\sigma \neq 1/2$. Для явной схемы $\sigma = 0$, для чисто неявной схемы $\sigma = 1$.

Рассмотрим шаблоны, изображенные на рис. 4.15. Отметим, что шаблон 4.15, в не переносится на трехмерный случай.

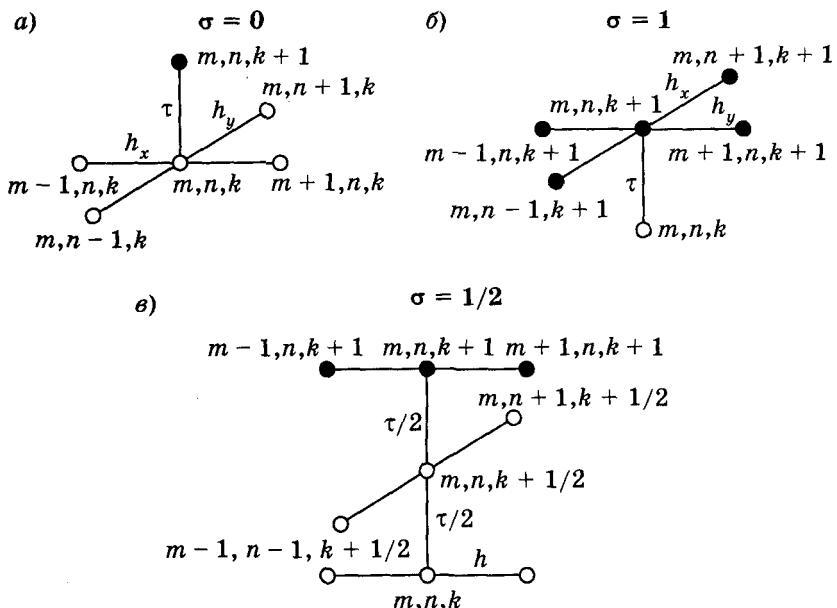


Рис. 4.15

Явная разностная схема (рис. 4.15, а)

$$\frac{u_{m,n}^{(k+1)} - u_{m,n}^{(k)}}{\tau} + L[u_{m,n}^{(k)}] = f_{m,n}^{(k)} \quad (4.40)$$

устойчива при $\tau/h^2 \leq 1/4$ (полагаем $h = h_x = h_y$). Схема позволяет по известным значениям функции на k -м временном слое найти решение на $(k+1)$ -м слое.

Неявная разностная схема (рис. 4.15, б)

$$\frac{u_{m,n}^{(k+1)} - u_{m,n}^{(k)}}{\tau} + L[u_{m,n}^{(k+1)}] = f_{m,n}^{(k)} \quad (4.41)$$

абсолютно устойчива и позволяет определить решение на $(k+1)$ -м слое путем решения системы линейных алгебраических уравнений с ленточной матрицей, имеющей широкую слабо заполненную ленту. Как следует из табл. 4.3, явная и неявная схемы требуют при $P = 2$ порядка N^4 и порядка N^5 операций соответственно, но и в том, и в другом случае необходимое число операций весьма велико. Между тем метод переменных направлений требует лишь N^2 операций.

Рассмотрим разностную схему метода переменных направлений. Для этого введем дополнительный промежуточный слой $(k+1/2)$ и следующую разностную аппроксимацию уравнения (4.37) на этом слое на шаблоне (рис. 4.15, в):

$$\frac{u_{m,n}^{(k+1/2)} - u_{m,n}^{(k)}}{\tau/2} + L_x(u_{m,n}^{(k)}) + L_y(u_{m,n}^{(k+1/2)}) = f_{m,n}^{(k)}. \quad (4.42)$$

Согласно (4.42) вторая производная по x аппроксимируется на k -м слое, а по y — на $(k+1/2)$ -м слое. Поэтому на $(k+1/2)$ -м слое имеем только три неизвестных значения $u_{m,n-1}^{(k+1/2)}$, $u_{m,n}^{(k+1/2)}$, $u_{m,n+1}^{(k+1/2)}$ в каждом уравнении системы, хотя требуется найти $(N-1)^2$ неизвестных (полагаем $N = M$).

Итак, во-первых, на $(k+1/2)$ -м слое имеем систему линейных уравнений с трехдиагональной матрицей, которую можно записать в виде

$$\frac{\tau}{2h_y^2} u_{m,n-1}^{(k+1/2)} - \left(1 + \frac{\tau}{2h_y^2}\right) u_{m,n}^{(k+1/2)} + \frac{\tau}{2h_y^2} u_{m,n+1}^{(k+1/2)} = F_{m,n}^k, \quad (4.43)$$

где $F_{m,n}^k = u_{m,n}^{(k)} - (\tau/2)(L_x[u_{m,n}^{(k)}] - f_{m,n}^{(k)})$ вычисляется по известным значениям функции на k -м слое. Присоединив к системе (4.43) значения $u_{m,0}^{(k+1/2)}$ и $u_{m,N}^{(k+1/2)}$, известные из краевых условий, мо-

жем вычислить значения неизвестных на $(k + 1/2)$ -м слое. Для этого путем решения системы линейных уравнений с трехдиагональной матрицей можно использовать такой эффективный метод решения, как метод прогонки (см. § 1.5).

Во-вторых, система линейных уравнений (4.43) распадается на $M - 1$ независимые трехдиагональные системы порядка $M - 1$ для каждого m . Независимое решение на $(k + 1/2)$ -м слое при фиксированном значении t позволяет существенно уменьшить объем вычислений по сравнению с решением одной системы порядка $(M - 1)^2$.

После того как найдены все неизвестные $u_{m, n}^{(k + 1/2)}$ на промежуточном слое, решение на $(k + 1)$ -м слое определяется разностной аппроксимацией уравнения (4.36):

$$\frac{u_{m, n}^{(k+1)} - u_{m, n}^{(k+1/2)}}{\tau/2} + L_x[u_{m, n}^{(k+1)}] + L_y[u_{m, n}^{(k+1/2)}] = f_{m, n}^{(k+1/2)}. \quad (4.44)$$

Соотношение (4.44) может быть переписано в виде

$$\frac{\tau}{2h_x^2} u_{m-1, n}^{(k+1)} - \left(1 - \frac{\tau}{2h_x^2}\right) u_{m, n}^{(k+1)} + \frac{\tau}{2h_x^2} u_{m+1, n}^{(k+1)} = F_{m, n}^{(k+1/2)}, \quad (4.45)$$

где значение $F_{m, n}^{(k+1/2)} = -(\tau/2)(L_y[u_{m, n}^{(k+1/2)}] - f_{m, n}^{(k)}) + u_{m, n}^{(k+1/2)}$ вычисляется по известным значениям функции на $(k + 1/2)$ -м слое. Аналогично ситуации на предыдущем промежуточном слое система (4.45) с присоединенными граничными значениями распределяется на $(N - 1)$ независимые трехдиагональные линейные системы, отвечающие фиксированным значениям n , каждая из

которых решается методом прогонки. Однако здесь прогонка выполняется по индексу m , т. е. в направлении оси x . Отметим, что коэффициенты при неизвестных в уравнениях (4.43) и (4.45) такие, что прогонка является корректной и устойчивой.

На рис. 4.16 схематически изображен переход с k -го слоя на $(k + 1)$ -й слой с дробным шагом* и указано направление прогонок (прямой и обратный ход).

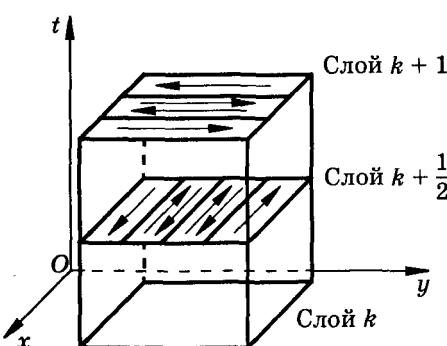


Рис. 4.16

* Метод переменных направлений носит также название метода дробных шагов.

Таким образом, в методе переменных направлений переход с известного k -го слоя на искомый $(k + 1)$ -й слой происходит с использованием дробного шага $\tau/2$ и промежуточного слоя $(k + 1/2)$. Разностная схема (4.42) и (4.44) в совокупности является абсолютно устойчивой, хотя при переходе с k -го слоя на $(k + 1/2)$ -й модуль перехода $|\lambda_x| > 1$, поскольку схема (4.45) локально неустойчива. Однако при переходе со слоя $(k + 1/2)$ на слой k модуль перехода $|\lambda_y| < 1$ и общее значение модуля перехода $|\lambda| = |\lambda_x| \cdot |\lambda_y|$ оказывается меньшим единицы.

Отметим, что существуют так называемые локально-одномерные схемы (или схемы расщепления), в которых на промежуточном $(k + 1/2)$ -м слое отсутствуют и аппроксимация, и устойчивость. Тем не менее при переходе со слоя k на $(k + 1)$ -й слой результирующая схема становится аппроксимирующей и устойчивой. Например, для уравнения (4.37) при $f = 0$ предложена локально одномерная разностная схема: при переходе от k -го слоя к $(k + 1/2)$ -му решение имеет вид

$$\frac{u_{m,n}^{(k+1/2)} - u_{m,n}^{(k)}}{\tau} = L_y[u_{m,n}^{(k+1/2)}] \quad (4.46)$$

с порядком аппроксимации $O(\tau, h^2)$, при переходе с $(k + 1/2)$ -го слоя на $(k + 1)$ -й — соответственно

$$\frac{u_{m,n}^{(k+1)} - u_{m,n}^{(k+1/2)}}{\tau} = L_x[u_{m,n}^{(k+1)}] \quad (4.47)$$

с порядком аппроксимации $O(\tau, h^2)$. На промежуточных слоях ни одна из схем (4.46) и (4.47) не является аппроксимирующей, хотя результирующая схема — аппроксимирующая и устойчивая с порядком аппроксимации $O(\tau, h^2)$.

4.13. Математические и физические основы метода установления

Метод установления обычно используется для решения уравнений эллиптического типа или смешанных эллиптико-гиперболических уравнений. В задачах газовой динамики уравнения эллиптического типа описывают стационарное дозвуковое течение, например, течение за отошедшей ударной волной при сверхзвуковом обтекании затупленных тел или течение в сужающейся

части сопла Лаваля. Уравнение для функции тока и потенциала скорости в плоском несжимаемом течении является уравнением Лапласа. К эллиптическому уравнению приводит задача определения напряжений, возникающих при упругом кручении длинного цилиндрического стержня, задача о стационарных потоках тепла в двумерном теле и многие другие.

Для уравнений эллиптического типа характерной является задача Дирихле, когда на границе Γ области G , в которой ищется решение, задана искомая функция $u(x, y)$ (рис. 4.17, а). Поясним идею метода установления на примере задачи Дирихле для уравнения Лапласа с двумя независимыми переменными x и y :

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad (4.48)$$

$$u(x, y) \Big|_{\Gamma} = u^0(x, y).$$

Задачу (4.48) будем называть СТАЦИОНАРНОЙ, так как ее решение не зависит от времени. Наряду со стационарной задачей рассмотрим эволюционную (нестационарную) задачу для параболического уравнения теплопроводности с теми же граничными условиями и произвольно выбранными начальными данными (рис. 4.17, б):

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}, \\ u(x, y, t) \Big|_{\Gamma} &= u^0(x, y), \\ u(x, y, 0) &= u_0(x, y). \end{aligned} \quad (4.49)$$

В общем случае, если граничные условия зависят от времени, задачу (4.49) называют НЕСТАЦИОНАРНОЙ КРАЕВОЙ ЗАДАЧЕЙ. В представленной формулировке задача (4.49) называется вы-

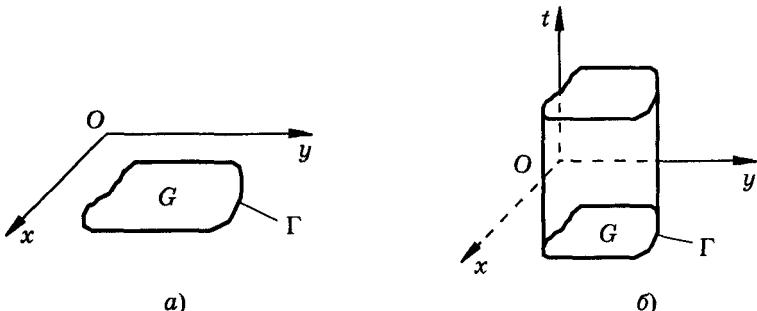


Рис. 4.17

рожденной нестационарной задачей, так как условие на границе области не зависит от времени.

Таким образом, в методе установления вводится новая независимая переменная t и задача формально усложняется — увеличивается размерность задачи. Однако при этом увеличивается набор возможных разностных схем решения задачи, что допускает выбор из них устойчивых и экономичных. Далее будет показано, что кажущееся усложнение задачи не приводит ни к увеличению времени счета, ни к потере точности. Важным достоинством метода установления является то, что краевая задача для эллиптического уравнения сводится к задаче с начальными и краевыми условиями. Из физических соображений достаточно очевидно, что решение эволюционной задачи (4.49) при $t \rightarrow \infty$ будет стремиться к решению стационарной задачи (4.48).

Действительно, решение стационарной задачи (4.48) дает стационарное распределение значений искомой функции (например, температуры) в области G при заданном ее распределении на границе области Γ . Решение эволюционной задачи (4.49) дает распределение значений функции во времени при тех же, что и в стационарной задаче, граничных условиях. Очевидно, что по истечении достаточно большого промежутка времени в области G установится распределение значений функции, не зависящее от начальных условий, а определяемое лишь значениями функции на границе. Приведенное рассуждение дает основание предполагать, что решение эволюционной задачи при $t \rightarrow \infty$ сводится к решению стационарной задачи и решение эволюционной задачи при достаточно большом $t = T$ может быть принято за приближенное решение исходной стационарной задачи.

Приведем некоторые математические рассуждения, подтверждающие сделанное предположение о выходе решения эволюционной задачи (4.49) на стационарный режим. Пусть $u(x, y, t)$ и $u(x, y)$ — решение задач (4.49) и (4.48) соответственно. Рассмотрим разность

$$w(x, y, t) = u(x, y, t) - u(x, y).$$

Функция w удовлетворяет дифференциальному уравнению (4.49). Начальные данные имеют вид

$$w(x, y, 0) = u_0(x, y) - u(x, y),$$

а однородные граничные условия —

$$w(x, y, t) = 0.$$

Используя метод разделения переменных Фурье при получении решения линейной однородной краевой задачи для функции w , покажем, что $w(x, y, t) \rightarrow 0$ при $t \rightarrow 0$. Рассмотрение прове-

дем на примере соответствующей одномерной задачи в области G : $[0 \leq t \leq T, 0 \leq x \leq 1]$. Имеем:

$$\frac{\partial w}{\partial t} = \frac{\partial^2 w}{\partial x^2}, \quad (4.50)$$

$$w(0, t) = w(1, t) = 0, \quad (4.51)$$

$$w(x, 0) = u_0(x) - u(x). \quad (4.52)$$

Будем искать частное решение задачи в виде

$$w(x, t) = X(x)Y(t). \quad (4.53)$$

Подставив (4.53) в (4.50), получаем

$$\frac{X''(x)}{X(x)} = \frac{Y'(t)}{Y(t)} = -\lambda, \quad (4.54)$$

где $\lambda = \text{const}$. Подставляя выражение (4.53) в условие (4.51), получаем

$$w(0, t) = X(0)Y(t) = 0, \quad (4.55)$$

$$w(1, t) = X(1)Y(t) = 0.$$

Из соотношений (4.55) следует, что $X(0) = X(1) = 0$. Выражение (4.54) дает два обыкновенных дифференциальных уравнения для нахождения функций $X(x)$ и $Y(t)$:

$$\frac{dY}{dt} + \lambda Y = 0, \quad (4.56)$$

$$\frac{d^2X}{dx^2} + \lambda X = 0. \quad (4.57)$$

Для уравнения (4.57) решаем краевую задачу с однородными условиями (4.55). Она сводится к простейшей задаче на собственные значения, из которой необходимо определить значения параметра λ , при которых существуют нетривиальные решения задачи (4.57), (4.55).

Рассмотрим три возможных случая.

• $\lambda < 0$; $X(x) = C_1 e^{\sqrt{-\lambda} x} + C_2 e^{-\sqrt{-\lambda} x}$. Из соотношений (4.55) следует, что $C_1 = C_2 = 0$; $X(x) = 0$.

• $\lambda = 0$; $X(x) = C_1 x + C_2$. Из соотношений (4.55) вновь следует, что $C_1 = C_2 = 0$; $X(x) = 0$.

• $\lambda > 0$; $X(x) = C_1 \cos(\sqrt{\lambda} x) + C_2 \sin(\sqrt{\lambda} x)$. Подставляя выражение для $X(x)$ в первую из формул (4.55), получаем: если $X(x) = 0$, то $C_1 = 0$ и $X(1) = C_2 \sin \sqrt{\lambda} = 0$ и если $X(x) \neq 0$, то $C_2 \neq 0$.

Нетривиальные решения задачи (4.57), (4.55) возможны при
 $\lambda = \lambda_j = (\pi j)^2, \quad j = 1, 2, \dots$.

Тогда из уравнения (4.56) следует, что $Y(t) = C_j e^{-\lambda_j t}$. Согласно выражению (4.53) частное решение $w(x, t)$ имеет вид

$$w(x, t) = C_j e^{-\lambda_j t} \sin(\pi j x).$$

Постоянные C_j определяются из разложения начальной функции $w(x, 0)$ в ряд Фурье. Не будем приводить конкретных формул для вычисления этих постоянных. Из последнего соотношения следует, что $w(x, t) \rightarrow 0$ при $t \rightarrow 0$, т. е. решение эволюционной задачи стремится при $t \rightarrow \infty$ к решению стационарной задачи. Таким образом, решение задачи (4.48) можно рассматривать как предел решения нестационарной задачи (4.49) при $t \rightarrow \infty$.

Рассмотрим еще одну эволюционную задачу, описывающую колебания мембранны при наличии вязкого трения:

$$\frac{\partial u}{\partial t} + \frac{\partial^2 u}{\partial t^2} = \Delta u,$$

$$u \Big|_{\Gamma} = u^0(x, y),$$

$$u \Big|_{t=0} = u_0(x, y), \quad \frac{\partial u}{\partial t} \Big|_{t=0} = u_1(x, y).$$

Колебания со временем будут затухать, и решение этой задачи будет стремиться к решению задачи (4.48), которую можно рассматривать как задачу о прогибе мембранны. Если в уравнении отбросить член $\frac{\partial u}{\partial t}$, то оно будет описывать процесс колебания мембранны без затухания амплитуды колебаний и решение устанавливаться не будет.

4.14. Разностные схемы метода установления

Рассмотрим для наглядности следующую прямоугольную область $G: [0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq t \leq T]$. Стационарная задача (4.48) сводится к простейшей разностной аппроксимации

$$\frac{u_{m+1, n} - 2u_{m, n} + u_{m-1, n}}{h_x^2} + \frac{u_{m, n+1} - 2u_{m, n} + u_{m, n-1}}{h_y^2} = 0. \quad (4.58)$$

При решении этой системы линейных уравнений с пятидиагональной матрицей может быть использован один из способов рассмотренных ранее итерационных процессов.

Для аппроксимации задачи (4.49) введем сетку

$$x_m = mh_x, \quad m = 0, \dots, M, \quad h_x = \frac{1}{M},$$

$$y_n = nh_y, \quad n = 0, \dots, N, \quad h_y = \frac{1}{N},$$

$$t_k = k\tau, \quad k = 0, \dots, K$$

и обозначение $u_{m,n}^{(k)} = u(x_m, y_n, t_k)$. Для дифференциального уравнения (4.49) запишем явную и неявную разностные схемы, шаблоны которых показаны соответственно на рис. 4.18, а, б:

$$\begin{aligned} \frac{u_{m,n}^{(k+1)} - u_{m,n}^{(k)}}{\tau} &= \\ &= \frac{u_{m+1,n}^{(k)} - 2u_{m,n}^{(k)} + u_{m-1,n}^{(k)}}{h_x^2} + \frac{u_{m,n+1}^{(k)} - 2u_{m,n}^{(k)} + u_{m,n-1}^{(k)}}{h_y^2}, \end{aligned} \quad (4.59)$$

$$\begin{aligned} \frac{u_{m,n}^{(k+1)} - u_{m,n}^{(k)}}{\tau} &= \\ &= \frac{u_{m+1,n}^{(k+1)} - 2u_{m,n}^{(k+1)} + u_{m-1,n}^{(k+1)}}{h_x^2} + \frac{u_{m,n+1}^{(k+1)} - 2u_{m,n}^{(k+1)} + u_{m,n-1}^{(k+1)}}{h_y^2}. \end{aligned} \quad (4.60)$$

Большинство итерационных методов (в том числе метод простой итерации и метод Зейделя) можно символически записать в виде

$$B_k \frac{u^{(k+1)} - u^{(k)}}{\tau_k} + Au^{(k)} = f,$$

где B_k — произвольная матрица, τ_k — шаг разностной схемы. Если при этом B_k и τ_k не зависят от номера итерации k , процесс на-

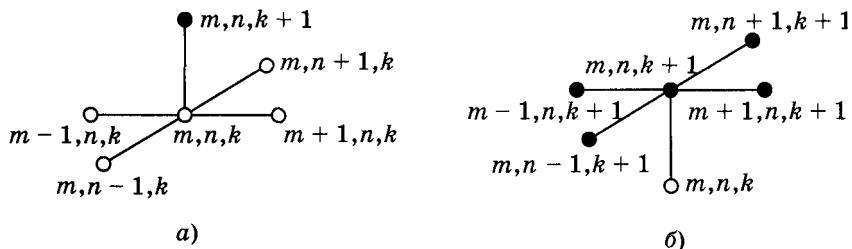


Рис. 4.18

зывается стационарным. В связи с этим разностные схемы (4.59) и (4.60) решения эволюционной задачи можно интерпретировать как итерационный метод решения системы линейных алгебраических уравнений, возникающей при разностной аппроксимации стационарного уравнения (4.48). Тогда для схемы (4.59) матрица $B_k = E$ и $\tau_k = \tau$ (здесь E — единичная матрица). Если для достижения заданной точности в методе установления потребуется меньшее число итераций (т. е. шагов по времени), чем при решении системы уравнений (4.58) каким-либо другим итерационным методом, то применение метода установления целесообразно. Таким образом, при использовании метода установления важно понимать, что итерационный процесс можно рассматривать как разностную схему решения некоторой вырожденной нестационарной (эволюционной) задачи.

Исследование устойчивости новой разностной схемы (4.59) для возмущения специального вида

$$u_{m,n}^{(k)} = \lambda^k e^{i\omega_x m h_x} e^{i\omega_y n h_y} \quad (4.61)$$

(здесь i — мнимая единица) дает для модуля перехода соотношение

$$\lambda = 1 - \frac{4\tau}{h_x^2} \sin^2\left(\frac{\omega_x h_x}{2}\right) - \frac{4\tau}{h_y^2} \sin^2\left(\frac{\omega_y h_y}{2}\right). \quad (4.62)$$

Учитывая, что величины ω_x и ω_y могут принимать произвольные значения, из соотношения (4.62) находим условие устойчивости

$$\tau \left(\frac{1}{h_x^2} + \frac{1}{h_y^2} \right) \leq \frac{1}{2},$$

или в частном случае квадратной сетки, когда $h_x = h_y = h$,

$$\frac{\tau}{h^2} \leq \frac{1}{4}. \quad (4.63)$$

Для неявной разностной схемы (4.60) модуль перехода определяется соотношением

$$\lambda = \left(1 + \frac{4\tau}{h_x^2} \sin^2\left(\frac{\omega_x h_x}{2}\right) + \frac{4\tau}{h_y^2} \sin^2\left(\frac{\omega_y h_y}{2}\right) \right)^{-1}, \quad (4.64)$$

из которого следует, что неявная разностная схема (4.60) абсолютно устойчива, так как $|\lambda| < 1$ при любых шагах τ , h_x и h_y .

Условие устойчивости (4.63) является весьма жестким. Оно, как правило, накладывает более сильные ограничения на шаг, чем естественные требования точности. В случае двух (и более) пространственных переменных применение неявных схем вызывает большие трудности, связанные с решением системы линейных уравнений.

Приведем пример, дающий некоторое представление о существе дела. Рассмотрим квадратную сетку с шагом $h_x = h_y = h$. Введем полуцелые временные слои для аппроксимации уравнения (4.49). Переход с k -го слоя на $(k+1)$ -й разобьем на два полу шага:

$$\frac{u_{m,n}^{(k+1/2)} - u_{m,n}^{(k)}}{\tau} = \frac{u_{m+1,n}^{(k+1/2)} - 2u_{m,n}^{(k+1/2)} + u_{m-1,n}^{(k+1/2)}}{h^2}, \quad (4.65)$$

$$\frac{u_{m,n}^{(k+1)} - u_{m,n}^{(k)}}{\tau} = \frac{u_{m,n+1}^{(k+1)} - 2u_{m,n}^{(k+1)} + u_{m,n-1}^{(k+1)}}{h^2}. \quad (4.66)$$

Отметим, что схемы (4.65) и (4.66) — неявные и содержат неизвестные значения на слое $k+1/2$ в схеме (4.65) и на слое $k+1$ в схеме (4.66). При этом в схеме (4.65) неизвестные значения функции вычисляются прогонкой по слоям $y = \text{const}$, в схеме (4.66) — прогонкой по слоям $x = \text{const}$.

Возникает вопрос о том, в каком смысле пара уравнений (4.65) и (4.66) приближает исходное уравнение (4.49). Заметим, что данное ранее определение аппроксимации здесь явно не годится: сеточные уравнения (4.65) и (4.66) не аппроксимируют уравнение (4.49) на дробных шагах. Однако можно показать, что для перехода с k -го слоя на $(k+1)$ -й результирующая схема является аппроксимирующей. Кроме того, если исходить из физических соображений, то имеются веские причины надеяться, что совокупность сеточных уравнений (4.65) и (4.66) приближенно описывает процесс распространения тепла. Действительно, уравнение (4.65) представляет собой сеточную аппроксимацию прецельно анизотропного процесса теплопередачи, при котором распространение тепла происходит лишь в направлении оси x ; аналогично можно истолковать уравнение (4.66). Можно надеяться, что при соответствующем подборе коэффициентов теплопроводности такое попеременное распределение тепла по двум направлениям будет приближенно описывать реальный (изотропный) процесс теплопроводности.

Покажем, что модуль перехода, соответствующий двум полу шагам, действительно аппроксимирует модуль перехода для уравнения (4.49). Методом гармоник Фурье можно получить выражения для модулей перехода разностных схем (4.65) и (4.66):

$$\lambda' = \left[1 + \frac{4\tau}{h^2} \sin^2 \left(\frac{\omega_1 h}{2} \right) (\omega_1 h / 2) \right]^{-1},$$

$$\lambda'' = \left[1 + \frac{4\tau}{h^2} \sin^2 \left(\frac{\omega_2 h}{2} \right) \right]^{-1}.$$

Для совокупности переходов (4.65) и (4.66) модуль перехода $\lambda = \lambda' \cdot \lambda''$. Из полученного соотношения очевидно, что схема (4.65), (4.66) абсолютно устойчива, т. е. устойчива при любых значениях шагов τ и h .

При $h \rightarrow 0$ можно записать приближенное соотношение для модуля перехода:

$$\lambda \approx 1 - \frac{4\tau}{h^2} \sin^2\left(\frac{\omega_1 h}{2}\right) - \frac{4\tau}{h^2} \sin^2\left(\frac{\omega_2 h}{2}\right) \approx 1 - \tau(\omega_1^2 + \omega_2^2).$$

С другой стороны, для точного решения уравнения теплопроводности

$$u = e^{-(\omega_1^2 + \omega_2^2)t} e^{i\omega_1 x} e^{i\omega_2 y}$$

модуль перехода определяется соотношением

$$\lambda = e^{-(\omega_1^2 + \omega_2^2)\tau} \approx 1 - \tau(\omega_1^2 + \omega_2^2).$$

Таким образом, имеет место аппроксимация по меньшей мере в следующем смысле: при переходе с k -го слоя на $(k+1)$ -й решение специального вида для сеточных уравнений изменяется так же, как и соответствующее решение дифференциального уравнения (4.49), с точностью до малых более высокого порядка, чем τ . Используя преобразование Фурье, можно строго обосновать сходимость приближенного решения к точному.

Для возмущения специального вида (4.61) скорость сходимости определяется модулем перехода λ , который для явной схемы (4.59) имеет вид (4.62). Наиболее медленно убывают низкочастотные возмущения, для которых $\lambda \approx 1 - \tau(\omega_1^2 + \omega_2^2)$. Условие устойчивости (4.63) ограничивает сверху допустимый шаг τ ; положим $\tau = \alpha h^2/4$, $0 < \alpha < 1$. Заметим далее, что минимальные значения частот ω_1 и ω_2 определяются размером области; в нашем случае $\omega_{1,2} \approx \pi$. Таким образом, для минимальных частот имеем $\lambda \approx 1 - \pi^2 \alpha h^2/4$. Из этого соотношения следует, что при малых значениях h сходимость оказывается весьма медленной. Для того чтобы уменьшить в $1/\varepsilon$ раз амплитуду ошибки, требуется $N = N(\varepsilon)$ слоев, где значение N определяется соотношениями

$$|\lambda|^n \approx e^{-1}, \quad N \approx \frac{|\ln \varepsilon|}{|\ln \lambda|} \approx |\ln \varepsilon| \frac{4}{\alpha \pi^2} \frac{1}{h^2}.$$

В частности, для $\varepsilon = 0,01$, $\alpha = 1$ и $h = 0,01$ значение $N = 8000$. Из этого примера видно, что простейшие итерационные методы, основанные на явных аппроксимациях нестационарной краевой задачи, могут оказаться неэффективными.

Рассмотрим теперь неявную аппроксимацию (4.65), (4.66), построенную по методу дробных шагов. Выражение (4.64) для модуля перехода показывает, что скорость затухания возмущений во всем спектре частот ω_1 и ω_2 может быть сделана сколь угодно большой при достаточно большом значении τ . Однако с увеличением значения τ возрастают и ошибки аппроксимации, связанные с представлением оператора перехода от k -го слоя к $(k+1)$ -му в виде произведения операторов, соответствующих полушагам. В предельном случае $\tau = \infty$ получаем два слоя (*целый* и *получелый*), не имеющие ничего общего с искомым решением и не похожие друг на друга. Возникает естественная идея варьирования τ : сначала, когда преобладают возмущения, связанные с ошибками начального слоя, *гасить* побыстрее эти возмущения, а затем, когда начинают все большую роль играть ошибки аппроксимации, постепенно уменьшать τ . На основе этой идеи построены весьма эффективные алгоритмы для решения стационарных сеточных краевых задач.

При использовании метода установления иногда вычисления прекращают, когда удовлетворяется неравенство $|u^{(k+1)} - u^{(k)}| \leq \varepsilon$, где ε — заданная точность. Можно улучшить критерий окончания счета на установление, сделав его по трем последним итерациям (см. гл. 1, § 1.13, метод Эйткена).

4.15. Методы сквозного счета

В задачах механики сплошных сред приходится иметь дело с течениями, в которых возникают поверхности сильного разрыва такие, как ударные волны, центрированные волны разрежения, поверхности контактного разрыва. На этих поверхностях терпят разрыв функции и их производные. При расчете таких течений приходится либо явно выделять эти поверхности, используя законы сохранения, либо строить специальные методы решения, называемые методами сквозного счета, в которых *a priori* неиз-

ГОДУНОВ СЕРГЕЙ КОНСТАНТИНОВИЧ (1929) — математик, чьи основные исследования относятся к теории чисел, теории функций комплексного переменного, теории дифференциальных уравнений, численным методам линейной алгебры, приближенным и численным методам прикладной математики. Г. разработал разностные методы решения нестационарных задач.

вестные поверхности разрыва явно не выделяются и просчитываются *насквозь* вместе с областями непрерывного течения.

Весьма эффективным и красивым методом сквозного счета является МЕТОД ГОДУНОВА. В основе метода Годунова лежат две идеи. Первая состоит в *использовании при построении разностной схемы точных решений с кусочно-постоянными начальными данными*. Для гиперболических уравнений такими точными решениями являются совокупности сравнительно простых и независимых решений задачи о распаде произвольного разрыва.

Вторая идея состоит в *использовании гибких и деформирующихся разностных сеток, связанных с поверхностями разрывов и выделяемых при расчете начальной стадии*. Рассмотрим применение метода Годунова при решении задачи Коши для уравнений акустики:

$$\begin{aligned} \frac{\partial U}{\partial t} + \frac{1}{\rho_0} \frac{\partial P}{\partial x} &= 0, \\ \frac{\partial P}{\partial t} + \rho_0 a_0^2 \frac{\partial U}{\partial x} &= 0, \end{aligned} \tag{4.67}$$

где U и P — соответственно возмущения скорости и давления в акустической волне, a_0 и ρ_0 — скорость звука и плотность в невозмущенном газе. Используя инвариант Римана $I_{\pm} = P \pm a_0 \rho_0 U$ и обозначив через f и g произвольные функции, определяемые из начальных условий, для уравнений (4.67) можно записать общее решение

$$\begin{aligned} U &= \frac{1}{2} [f(x - a_0 t) + g(x + a_0 t)], \\ P &= \frac{\rho_0 a_0}{2} [f(x - a_0 t) - g(x + a_0 t)]. \end{aligned} \tag{4.68}$$

Построим с использованием (4.68) решение простейшей задачи о распаде. Обозначим через U_I , U_{II} , P_I , P_{II} некоторые постоян-

РИМАН ГЕОРГ ФРИДРИХ БЕРНХАРД (Riemann Georg Friedrich Bernhard; 1826—1866) — немецкий математик. Р. дал оригинальное построение теории аналитических функций, положил начало геометрическому направлению в ней, выдвинул ряд основных идей новой науки — топологии. Труды Р. посвящены разработке неевклидовой геометрии, рассматривавшейся им как учение о непрерывных n -мерных многообразиях. Таким образом, им были обобщены идеи Гаусса по внутренней геометрии поверхности, Гаусса — чьи лекции он слушал, будучи студентом Геттингенского университета. Работы Р. оказали глубокое влияние на развитие математики второй половины XIX—XX вв.

ные такие, что $U_I \neq U_{II}$ и $P_I \neq P_{II}$. Пусть в момент $t = 0$ заданы начальные условия

$$\begin{aligned} U_0(x) &= U_I, P_0(x) = P_I && \text{при } x < x^*, \\ U_0(x) &= U_{II}, P_0(x) = P_{II} && \text{при } x > x^*. \end{aligned} \quad (4.69)$$

Начальные условия (4.69) таковы, что имеет место произвольный разрыв функций U и P в точке x^* , который с течением времени каким-то образом трансформируется в пространстве x, t . Такая задача и называется задачей о распаде произвольного разрыва. Построим ее решение, используя свойства инвариантов Римана. Поскольку инвариант Римана I_+ постоянен вдоль характеристики $x - a_0 t = \text{const}$, а инвариант I_- постоянен вдоль характеристики $x + a_0 t = \text{const}$, приходим к структуре решения, изображенной на рис. 4.19. Решение задачи о распаде разрыва имеет вид

$$\begin{aligned} U &= U_I, P = P_I && \text{при } x < x^* - a_0 t \text{ в области I,} \\ U &= U_{II}, P = P_{II} && \text{при } x > x^* + a_0 t \text{ в области II,} \\ U &= U_{III} = \frac{U_I + U_{II}}{2} - \frac{P_{II} - P_I}{2a_0\rho_0}, && (4.70) \\ P &= P_{III} = \frac{P_I + P_{II}}{2} - a_0\rho_0 \frac{U_{II} - U_I}{2} \\ &&& \text{при } x^* - a_0 t < x < x^* + a_0 t \\ &&& \text{в области III.} \end{aligned}$$

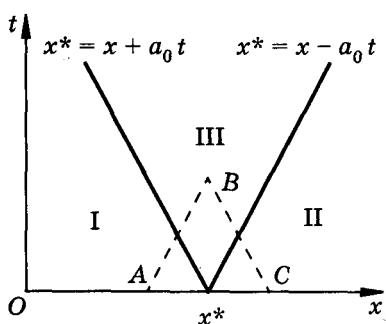


Рис. 4.19

Эти области изображены на рис. 4.19*.

Построенные таким образом функции $U(x, t)$ и $P(x, t)$ имеют разрывы вдоль характеристик $x + a_0 t = x^*$ и $x - a_0 t = x^*$, разделяющих области I, II и III. Поскольку функции U и P не являются непрерывными, то примем, что полученное решение является обобщенным решением задачи о распаде разрыва.

Перейдем к построению разностной схемы метода Годунова.

*Формулы для области III в соотношениях (4.70) получены из рассмотрения инварианта Римана на характеристиках AB и BC . Имеем $P_{III} + a_0\rho_0 U_{III} = P_I + a_0\rho_0 U_I$ на AB , $P_{III} - a_0\rho_0 U_{III} = P_{II} - a_0\rho_0 U_{II}$ на BC . Складывая и вычитая эти соотношения, получим выражения в формулах (4.70).

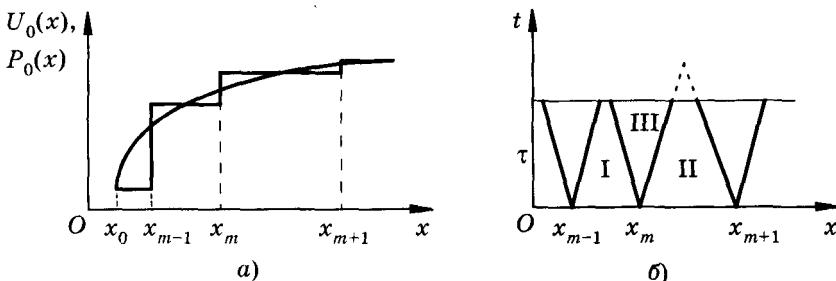


Рис. 4.20

Пусть в момент времени $t = 0$ заданы непрерывные функции $U_0(x)$ и $P_0(x)$. Заменим область непрерывного изменения аргумента дискретным множеством точек — разностной сеткой, узлы которой обозначим через x_m . Расстояние между соседними узлами является шагом разностной сетки. Примем, что между точками x_{m-1} и x_m функции U и P постоянны. Такое предположение эквивалентно тому, что непрерывные функции $U_0(x)$ и $P_0(x)$ заменены некоторыми кусочно-постоянными функциями, сохраняющими постоянные значения между узлами разностной сетки (рис. 4.20, а).

Значения функций $U(x)$ и $P(x)$ между узлами разностной сетки обозначим через $U_{m-1/2}$ и $P_{m-1/2}$, присвоив соответствующему слою полуцелый индекс ($m - 1/2$).

На границе между слоями возникает распад разрыва. В результате в каждом узле сетки образуются звуковые волны, распространяющиеся вправо и влево со скоростью звука a_0 . Через некоторый отрезок времени τ структура решения примет вид, представленный на рис. 4.20, б. Согласно формулам (4.70) решение в окрестности точки x_m будет

$$\begin{aligned}
 U &= U_{m-1/2}, P = P_{m-1/2} \\
 &\text{в области I: } x_{m-1} + a_0 t < x < x_m - a_0 t, \\
 U &= U_{m+1/2}, P = P_{m+1/2} \\
 &\text{в области II: } x_m + a_0 t < x < x_{m+1} - a_0 t, \\
 U &= U_m = \frac{U_{m-1/2} + U_{m+1/2}}{2} - \frac{P_{m+1/2} - P_{m-1/2}}{2a_0\rho_0}, \\
 P &= P_m = \frac{P_{m-1/2} - P_{m+1/2}}{2} - a_0\rho_0 \frac{U_{m+1/2} - U_{m-1/2}}{2} \\
 &\text{в области III: } x_m - a_0 t < x < x_m + a_0 t.
 \end{aligned} \tag{4.71}$$

Здесь через U_m и P_m обозначены значения, возникающие при распаде разрыва в узле m .

Изображенная на рис. 4.20 структура решения сохранится до тех пор, пока звуковые волны, вышедшие из соседних узлов, не встретятся между собой. После этого ее нужно перестроить заново, и даже в случае равномерной сетки эта процедура оказывается чрезвычайно громоздкой.

В связи с этим предлагается следующий способ построения решения в момент времени $t = \tau$. Заменим решение (4.71) другим решением таким, которое имеет ту же структуру, что и в начальный момент $t = 0$. Приближенно примем, что решение при $t = \tau$ также кусочно-постоянно внутри интервалов, ограниченных прежними узлами x_m . Новые средние значения этих функций при $t = \tau$ между узлами x_{m-1} и x_m обозначим через $U^{m-1/2}$ и $P^{m-1/2}$.

Эти величины можно определить по формулам

$$\begin{aligned} U^{m-1/2} &= U_{m-1/2} - \frac{\tau}{h}(P_m - P_{m-1}), \\ P^{m-1/2} &= P_{m-1/2} - \frac{\tau}{h}\rho_0 a_0^2(U_m - U_{m-1}), \end{aligned} \quad (4.72)$$

в которых значения U и P в правой части определяются по формулам (4.61) для точек m и $m-1$ соответственно.

Действительно, рассмотрим законы сохранения в интегральной форме для первого и второго уравнений (4.66), которые следуют из формулы Грина

$$\iint \left(\frac{\partial \Phi_1}{\partial t} + \frac{\partial \Phi_2}{\partial x} \right) dx dt = \oint (\Phi_1 dx - \Phi_2 dt),$$

в которой $\Phi_1 = U$ или $\Phi_1 = P$ и $\Phi_2 = P/\rho_0$ или $\Phi_2 = \rho_0 a_0^2 U$ для первого и второго уравнений соответственно. Имеем

$$\oint (U dx - (P/\rho_0) dt) = 0,$$

$$\oint (P dx - \rho_0 a_0^2 U dt) = 0.$$

Запишем эти соотношения для прямоугольной ячейки в плоскости (x, t) , ограниченной прямыми $x = x_{m-1}$, $x = x_m$, $t = 0$, $t = \tau$. Имеем

ГРИН ДЖОРДЖ (Green George; 1793—1841) — английский математик. Г. нашел соотношение между интегралами по объему и поверхности, ограничивающей объем. Он ввел термин *потенциал* и развил теорию электричества и магнетизма.

$$\begin{aligned}
 & \int_{x_{m-1}}^{x_m} U(x, t) dx = \\
 &= \int_{x_{m-1}}^{x_m} U(x, 0) dx - \frac{1}{\rho_0} \int_0^\tau \left[P(x_m, t) - P(x_{m-1}, t) \right] dt, \\
 & \int_{x_{m-1}}^{x_m} P(x, t) dx = \\
 &= \int_{x_{m-1}}^{x_m} P(x, 0) dx - \rho_0 a_0^2 \int_0^\tau \left[U(x_m, t) - U(x_{m-1}, t) \right] dt.
 \end{aligned} \tag{4.73}$$

Так как $U(x, 0)$ и $P(x, 0)$ постоянны на интервале от x_{m-1} до x_m и равны соответственно $U_{m-1/2}$ и $P_{m-1/2}$ и, аналогично, $U(x, \tau)$ и $P(x, \tau)$ равны $U^{m-1/2}$ и $P^{m-1/2}$, а на левой и правой сторонах ячейки $U(x, t)$ и $P(x, t)$ принимают постоянные значения U_{m-1} , U_m и P_{m-1} , P_m соответственно, которые вырабатываются при распаде разрыва в узлах m и $(m-1)$, то из формул (4.73) следуют соотношения (4.72).

Формулы (4.73) справедливы до тех пор, пока не происходит пересечения звуковых волн, приходящих из соседних ячеек. Это имеет место при $\tau < h/a_0$, что, как видно, является условием Куранта — Фридрихса — Леви. Описанная разностная схема перехода с одного временного слоя на другой имеет первый порядок точности по τ и h .

4.16. Метод прямых

Основная идея метода прямых и рассматриваемого далее метода характеристик состоит в сведении уравнений в частных производных к решению системы обыкновенных дифференциальных уравнений. В этом состоит их отличие от метода сеток, который непосредственно сводит решение систем уравнений в частных производных к решению систем алгебраических уравнений. Метод прямых может быть использован для решения уравнений в частных производных любого типа, но используется в основном для решения эллиптических и параболических уравнений.

В методе прямых уравнения в частных производных рассматриваются на линиях, параллельных одной из координатных осей. Это приводит к исчезновению зависимости от одной из независимых переменных. Сущность метода состоит в следующем. Пусть в прямоугольной области G : $[\alpha < x < \beta, y_0 < y < y_0 + H]$ необходимо найти решение дифференциального уравнения

$$\begin{aligned} a(x, y) \frac{\partial^2 u}{\partial x^2} + b(x, y) \frac{\partial^2 u}{\partial y^2} + c(x, y) \frac{\partial u}{\partial x} + d(x, y) \frac{\partial u}{\partial y} + e(x, y)u = \\ = f(x, y), \end{aligned} \quad (4.74)$$

где параметры a, b, c, d — положительные числа. Обозначим через $\varphi_j(x)$, $\psi_j(y)$ ($j = 0, 1$) заданные функции (рис. 4.21). Пусть уравнение (4.74) удовлетворяет граничным условиям

$$\begin{aligned} u(x, y_0) = \varphi_0(x), \quad u(x, y_0 + H) = \varphi_1(x) \quad \text{при } \alpha \leq x \leq \beta, \\ u(\alpha, y) = \psi_0(y), \quad u(\beta, y) = \psi_1(y) \quad \text{при } y_0 < y < y_0 + H. \end{aligned} \quad (4.75)$$

На отрезке $[y_0, y_0 + H]$ через точки $y_n = y_0 + nh$ ($n = 1, \dots, N$), $h = H/(1 + N)$ проведем прямые $y = y_n = \text{const}$ (рис. 4.21). Предполагая существование достаточно гладкого решения $u(x, y)$ задачи (4.74) — (4.75), положим в уравнении (4.74) $y = y_n$ и заменим производные по y разностными отношениями, приняв $u_n(x) = u(x, y_n)$ и воспользовавшись, например, соотношениями

$$\left. \frac{\partial u}{\partial y} \right|_{y=y_n} = \frac{1}{2h} [u_{n+1}(x) - u_{n-1}(x)] + o(h^2),$$

$$\left. \frac{\partial^2 u}{\partial y^2} \right|_{y=y_n} = \frac{1}{h^2} [u_{n+1}(x) - 2u_n(x) + u_{n-1}(x)] + o(h^2).$$

Проводя такую замену на всех линиях $y_n = \text{const}$ ($n = 1, \dots, N$) и пренебрегая малыми порядка $O(h^2)$, получаем систему N обык-

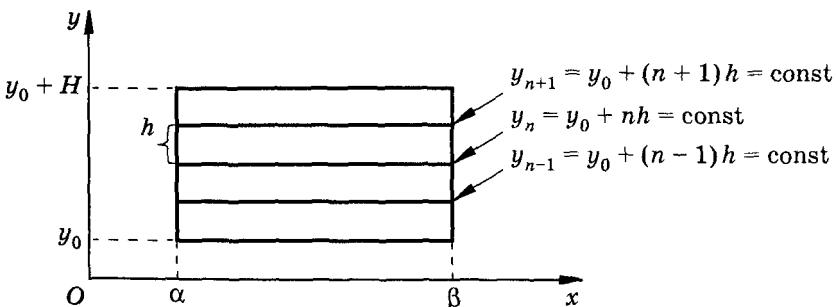


Рис. 4.21

новенных линейных дифференциальных уравнений второго порядка:

$$a_n(x)u_n''(x) + \frac{b_n(x)}{h^2} [u_{n+1}(x) - 2u_n(x) + u_{n-1}(x)] + \\ + c_n(x)u_n'(x) + \frac{d_n(x)}{2h} [u_{n+1}(x) - u_{n-1}(x)] + \\ + e_n(x)u_n(x) = f_n(x), \quad n = 1, \dots, N, \quad (4.76)$$

где через $u_n(x)$ обозначено приближенное значение решения $u(x, y)$ на прямой $y = y_n$; а через $a_n(x), b_n(x), c_n(x), d_n(x), e_n(x)$ и $f_n(x)$ — значения коэффициентов a, b, c, d, e, f в уравнении (4.74) на линии $y = y_n = \text{const}$.

Используя граничные условия на границе области, получим

$$\begin{aligned} u_0(x) &= \varphi_0(x), & \alpha \leq x \leq \beta, \\ \dots &\dots & \\ u_N(x) &= \varphi_1(x), & \alpha \leq x \leq \beta, \\ u_n(\alpha) &= \psi_0(y_n), \quad u_n(\beta) = \psi_1(y_n), & n = 1, \dots, N. \end{aligned} \tag{4.77}$$

На прямой $y = y_n$ каждое из уравнений (4.76) содержит три неизвестных значения функции $u_{n+1}(x)$, $u_n(x)$, $u_{n-1}(x)$. Всего с учетом граничных значений число неизвестных равно числу линий $y_n = \text{const}$ и равно числу уравнений в системе.

Система (4.76) обыкновенных дифференциальных уравнений с граничными условиями (4.77) аппроксимирует уравнение (4.74) с граничными условиями (4.75) и называется СИСТЕМОЙ УРАВНЕНИЙ МЕТОДА ПРЯМЫХ.

Эта система уравнений вместе с граничными условиями решается численно (см. § 3.7—3.9) либо аналитически.

В зависимости от способа замены производных по координате y разностными соотношениями будем иметь различные системы метода прямых, с различной точностью аппроксимирующие дифференциальное уравнение (4.74). Метод прямых удобнее всего применять, если коэффициенты в уравнении (4.74) не зависят от x . В этом случае система (4.76) будет системой обыкновенных дифференциальных уравнений с постоянными коэффициентами.

Метод прямых можно рассматривать как предельный случай метода сеток, если шаг сетки по оси x устремить к нулю. Отметим важное свойство метода прямых. Если изменение искомой функции внутри области незначительно, то можно ограничиться разбиением области на одну или две полосы, что в ряде случаев существенно упрощает задачу, а иногда позволяет получить анали-

тическое решение. В механике сплошных сред широко применяется метод интегральных соотношений, аналогичный методу прямых.

4.17. Метод характеристик

Метод характеристик применяется для решения уравнений гиперболического типа в задачах газовой динамики и теории упругости. Метод характеристик позволяет уменьшать на единицу размерность задачи и в случае уравнений в частных производных с двумя независимыми переменными сводит их к системе обыкновенных дифференциальных уравнений. Метод характеристик тесно связан с решением задачи Коши. Характеристические поверхности (линии) определяются как поверхности (линии), на которых решение задачи Коши либо не существует, либо не единствено.

Рассмотрим квазилинейную систему порядка p вида

$$A \frac{\partial u}{\partial x} + B \frac{\partial u}{\partial y} = f, \quad (4.78)$$

в которой матрицы A и B и вектор f с элементами a_{ij} , b_{ij} , f_i являются функциями независимых переменных x , y и компонент u_1, \dots, u_p искомой вектор-функции. Оказывается, что в плоскости (x, y) существуют линии, называемые характеристиками, на которых решение задачи либо не существует, либо не единствено. Уравнение этих линий можно получить, приравняв к нулю определитель

$$\left| B - A \frac{\partial y}{\partial x} \right| = 0. \quad (4.79)$$

Из соотношения (4.79), которое является многочленом степени p , можно найти производные dy/dx , определяющие характеристические направления. Таким образом, получаем p дифференциальных уравнений

$$\frac{dy}{dx} = \lambda_i(x, y, u), \quad i = 1, \dots, p, \quad (4.80)$$

которые называют уравнениями направлений или уравнениями характеристик. Отметим, что p действительных корней уравнения (4.79) существуют только для гиперболических систем.

Через каждую точку плоскости (x, y) проходит p характеристик. На характеристиках выполняются соотношения, называе-

мые уравнениями совместности. Эти соотношения связывают дифференциалы искомых функций с дифференциалами независимых переменных вдоль характеристик и получаются, если приравнять к нулю определитель порядка p расширенной матрицы ($Bdx - Ady$, $fdx - Adu$). Таким образом, можно видеть, что система дифференциальных уравнений в частных производных (4.78) сводится к эквивалентной системе обыкновенных дифференциальных уравнений.

Поясним сказанное на примере системы уравнений в частных производных с двумя неизвестными функциями u_1 и u_2 . Имеем систему уравнений

$$A \frac{\partial u}{\partial x} + B \frac{\partial u}{\partial y} = f, \quad (4.81)$$

где A и B — матрицы второго порядка, $u = (u_1, u_2)^T$ — векторы-столбцы, $f = (f_1, f_2)^T$ — векторы-столбцы правых частей. Перешифтуем систему (4.81) в виде

$$\begin{aligned} a_{11} \frac{\partial u_1}{\partial x} + a_{12} \frac{\partial u_2}{\partial x} + b_{11} \frac{\partial u_1}{\partial y} + b_{12} \frac{\partial u_2}{\partial y} &= f_1, \\ a_{21} \frac{\partial u_1}{\partial x} + a_{22} \frac{\partial u_2}{\partial x} + b_{21} \frac{\partial u_1}{\partial y} + b_{22} \frac{\partial u_2}{\partial y} &= f_2. \end{aligned} \quad (4.82)$$

Введем обозначения: $a = |A|$, $b = |B|$, $c = \begin{vmatrix} b_{12} & b_{21} \\ a_{12} & a_{22} \end{vmatrix} - \begin{vmatrix} a_{11} & b_{11} \\ a_{12} & b_{22} \end{vmatrix}$.

Уравнение (4.79) для системы (4.82) имеет вид $a\lambda^2 + c\lambda + b = 0$ или

$$\begin{vmatrix} b_{11} - \lambda a_{11} & b_{12} - \lambda a_{12} \\ b_{21} - \lambda a_{21} & b_{22} - \lambda a_{22} \end{vmatrix} = 0.$$

Решив это уравнение, найдем два уравнения для направлений двух характеристик, проходящих через каждую точку плоскости (x, y) (рис. 4.22, а):

$$\frac{dy}{dx} = \lambda^{(1)}(x, y, u_1, u_2) = \frac{-c + \sqrt{c^2 - 4ab}}{2a}, \quad (4.83)$$

$$\frac{dy}{dx} = \lambda^{(2)}(x, y, u_1, u_2) = \frac{-c - \sqrt{c^2 - 4ab}}{2a}. \quad (4.84)$$

Уравнения совместности получим, если приравняем к нулю определители первого и третьего или второго и третьего столбцов расширенной матрицы

$$\begin{pmatrix} b_{11} dx - a_{11} dy & b_{12} dx - a_{12} dy & f_1 dx - a_{11} du_1 - a_{12} du_2 \\ b_{21} dx - a_{21} dy & b_{22} dx - a_{22} dy & f_2 dx - a_{21} du_1 - a_{22} du_2 \end{pmatrix}.$$

Если ввести обозначения $\bar{c} = a_{22}b_{12} - a_{12}b_{22}$, $h = b_{22}f_1 - b_{12}f_2$, $e = a_{12}f_2 - a_{22}f_1$, $g = h + e\lambda^{(k)}$, то уравнения совместности примут вид

$$\begin{aligned}[a\lambda^{(k)} + b]du_1 + \bar{c}du_2 + hdx +edy &= 0, \quad k = 1, 2, \\ hdx + edy &= gdx.\end{aligned}\tag{4.85}$$

Уравнения (4.83)–(4.85) образуют систему четырех обыкновенных дифференциальных уравнений для нахождения значений x , y , u_1 , u_2 вдоль характеристик.

Перейдем к изложению существа численного метода характеристик. Рассмотрим в плоскости (x, y) две достаточно близкие точки 1 и 2 (рис. 4.22, б), не лежащие на одной характеристике, и зададимся целью определить значения u_1 и u_2 в точке 3 пересечения характеристик, проходящих через эти точки. Координаты точки l будем обозначать через (x_l, y_l) , а значение любой функции q в точке l — через q_l .

Переходя в уравнениях (4.83), (4.84) от дифференциалов к разностям, получим для определения координат точки 3 (x_3, y_3) систему двух линейных уравнений

$$\begin{cases} y_3 - y_1 = \lambda_1^{(1)}(x_3 - x_1), \\ y_3 - y_2 = \lambda_2^{(2)}(x_3 - x_2). \end{cases}$$

Аналогично из соотношений (4.85) получаем

$$(\lambda^{(1)}a + b)_1(u_{13} - u_{11}) + c_1(u_{23} - u_{21}) + g_1(x_3 - x_1) = 0,$$

$$(\lambda^{(2)}a + b)_2(u_{13} - u_{12}) + c_2(u_{23} - u_{22}) + g_2(x_3 - x_2) = 0,$$

где $u_{ij} = u_i(x_j, y_j)$. Решая эту систему относительно x_3 , y_3 , u_{13} и u_{23} , находим приближенные значения x_3 , y_3 и u_1 , u_2 в точке 3. Это приближение может оказаться недостаточно точным, так как характеристики, выходящие из точек 1 и 2, заменены отрезками прямых. В действительности точка 3 должна быть точкой пересечения, вообще говоря, криволинейных характеристик.

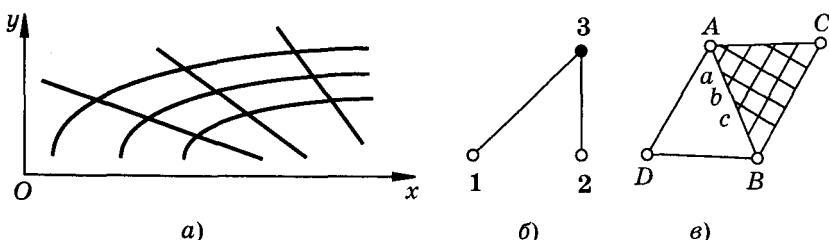


Рис. 4.22

Кроме того, дифференциалы всюду заменим разностями. Ошибка в определении переменных x_3 , y_3 , u_{13} и u_{23} будет иметь порядок квадрата расстояния от точки 3 до точек 1 и 2. Решение можно уточнить, если при вычислении коэффициентов при дифференциалах в (4.77)–(4.79) использовать значения искомых величин в точке 3. Этого можно достичь, используя итерационный процесс:

$$y_3^{(k)} - y_1 = (\lambda^{(1)})_1^{(k)} (x_3^{(k)} - x_1),$$

$$y_3^{(k)} - y_2 = (\lambda^{(2)})_2^{(k)} (x_3^{(k)} - x_2),$$

$$(\lambda^{(1)}a + b)_1^{(k)} (u_{1,3}^{(k)} - u_{1,1}) + c_1^{(k)} (u_{2,3}^{(k)} - u_{2,1}) + g_1^{(k)} (x_3^{(k)} - x_1) = 0,$$

$$(\lambda^{(2)}a + b)_2^{(k)} (u_{1,3}^{(k)} - u_{1,2}) + c_2^{(k)} (u_{2,3}^{(k)} - u_{2,2}) + g_2^{(k)} (x_3^{(k)} - x_2) = 0,$$

где $k = 1, 2, 3, \dots$ — номер итерации, $q_l^{(k)} = (q_3^{(k-1)} + q_l)/2$ — общая формула для определения значений коэффициентов уравнений при переходе с итерации на итерацию, при этом первая итерация соответствует методу Эйлера, а вторая — методу Эйлера с пересчетом.

Пусть теперь u_1 и u_2 заданы на дуге AB некоторой кривой, которая ни в одной точке не имеет характеристического направления. Требуется определить решение в окрестности отрезка AB , что является задачей Коши. Выберем на отрезке AB ряд точек A, a, b, \dots, B (рис. 4.22, в) и проведем через каждую из них характеристики обоих семейств (4.77) и (4.78). В точках их пересечения описанным выше способом можно вычислить искомые функции. Зная решение в этих точках, можно вычислить решение еще на одном слое, и так до тех пор, пока не будет найдено решение в точке C . Таким образом последовательно определяется решение и одновременно выстраивается характеристическая сетка. Аналогично определяется решение в характеристическом треугольнике ABD (рис. 4.22, в).

4.18. Метод конечных элементов

В последние годы для решения краевых задач механики сплошной среды интенсивно развивается метод конечных элементов. Этот метод начал использоваться для задач строительной меха-

ники, а затем превратился в общий численный метод решения систем обыкновенных дифференциальных уравнений и уравнений в частных производных.

Метод конечных элементов содержит в себе основные концепции *метода сеток*, связанные с дискретизацией областей непрерывного изменения аргумента и искомой функции, и *метода Галёркина*, использующего вариационные принципы для нахождения искомых функций в узлах рассматриваемой области. В конечном итоге решение краевой задачи сводится к решению системы линейных алгебраических уравнений, обладающих ленточной матрицей. Порядок системы равен числу узловых точек.

Основная идея МЕТОДА КОНЕЧНЫХ ЭЛЕМЕНТОВ состоит в том, что любую непрерывную функцию — такую, как зависимость температуры, давления, перемещения от координат — можно аппроксимировать дискретной моделью, которая строится на множестве кусочно-непрерывных функций, определенных на конечном числе подобластей. Кусочно-непрерывные функции определяются с помощью значений непрерывной функции в конечном числе точек рассматриваемой области.

В общем случае значения непрерывной функции заранее неизвестны и нужно определить их в некоторых внутренних точках области. Дискретную модель, однако, легко построить, если сначала предположить, что числовые значения искомой функции в каждой внутренней точке известны.

Построение дискретной модели состоит из нескольких этапов.

- В рассматриваемой области фиксируется конечное число точек. Эти точки называются узловыми точками или просто узлами.
- Значения непрерывной функции в каждой узловой точке считаются неизвестными, которые должны быть определены.

ГАЛЁРКИН БОРИС ГРИГОРЬЕВИЧ (1871—1943) — инженер и ученый, труды которого посвящены строительной механике и теории упругости. Г. разработал эффективные методы точного и приближенного решения уравнений теории упругости, предложил метод приближенного интегрирования краевых задач (метод Бубнова — Галёркина), нашедший широкое применение при решении задач теории упругости, вариационного исчисления, математической физики.

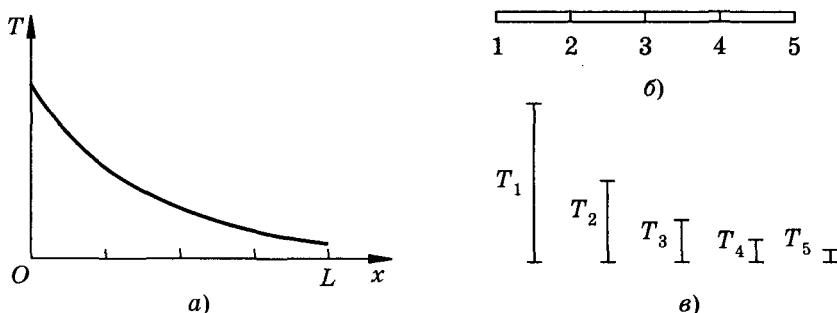


Рис. 4.23

- Область определения разбивается на конечное число подобластей, называемых **элементами**. Эти элементы имеют общие узловые точки и в совокупности аппроксимируют форму области.
- Непрерывная функция аппроксимируется на каждом элементе полиномом, который определяется с помощью узловых значений функции. Для каждого элемента определяется свой полином, но полиномы подбираются таким образом, чтобы сохранилась непрерывность функции вдоль границы.

Основные концепции метода конечных элементов могут быть наглядно проиллюстрированы на примере заданного одномерного распределения температуры в стержне (рис. 4.23, а). Рассматривается непрерывная функция $T(x)$, областью определения которой является отрезок OL вдоль оси x . Фиксированы и пронумерованы пять точек на оси x (рис. 4.23, б). Эти узловые точки не обязательно располагать на равном расстоянии друг от друга. Очевидно, можно ввести в рассмотрение и более пяти точек, но этих пяти вполне достаточно, чтобы проиллюстрировать основную идею метода. Значения функции T в данном случае известны в каждой узловой точке. Эти фиксированные значения представлены графически на рис. 4.23, в и обозначены в соответствии с номерами узловых точек.

Разбиение области на элементы может быть проведено различными способами. Можно, например, ограничить каждый элемент двумя соседними узловыми точками, образовав четыре элемента (рис. 4.24, а), или разбить на два элемента, каждый из которых содержит три узла (рис. 4.24, б).

Соответствующий элементу полином определяется по значению функции T в узловых точках элемента. В случае разбиения области на четыре элемента, когда на каждый элемент приходится по два узла, функция элемента будет линейна по T (две точки

однозначно определяют прямую линию): окончательная аппроксимация функции T будет состоять из четырех кусочно-линейных функций, каждая из которых определена на отдельном элементе. В случае разбиения области на два элемента с тремя узловыми точками в каждом функция элемента представляет собой полином второй степени; окончательная аппроксимация функции T будет состоять из двух кусочно-непрерывных квадратичных функций. Отметим, что и в том, и в другом случае приближение будет именно кусочно-непрерывным, так как углы наклона аппроксимирующих полиномов в граничных угловых точках, принадлежащих двум элементам, могут иметь различные значения.

В общем случае распределение температуры неизвестно и необходимо определить ее значение в некоторых фиксированных точках. Методика построения дискретной модели остается точно такой же, как описано в предыдущем случае, но с добавлением одного дополнительного шага. Снова находится множество узлов и значения температуры в этих узлах, но теперь эти значения являются неизвестными. Область решения разбивается на элементы, на каждом из которых определяется соответствующая функция элемента. Узловые значения функции T должны быть теперь определены таким образом, чтобы обеспечивалось в каком-то смысле наилучшее приближение к истинному распределению температуры. Это наилучшее приближение находится из условия минимизации некоторой величины, связанной с физической сущностью задачи. Если рассматривается задача о распределении тепла, то минимизируется функционал, связанный с соответствующим дифференциальным уравнением. Процесс минимизации сводится к решению системы линейных алгебраических уравнений относительно узловых значений T .

При построении дискретной модели непрерывной величины, определенной в двух- или трехмерной области, основная концепция метода конечных элементов используется аналогично. В двумерном случае элементы описываются функциями от пере-

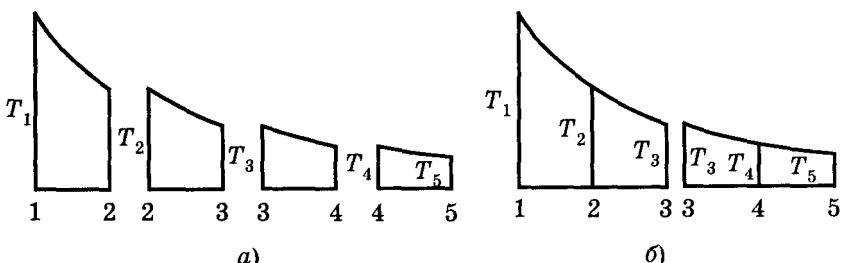


Рис. 4.24

менных x и y . При этом чаще всего рассматриваются элементы в форме треугольника или четырехугольника. Функции элемента представляются здесь плоскими или криволинейными поверхностями: функция элемента представляется плоскостью, если для данного элемента взято минимальное число точек. Для треугольного элемента оно равно трем, а для четырехугольного — четырем.

Важным аспектом метода конечных элементов является возможность выделить из набора элементов типичный элемент при нахождении функции элемента. Это позволяет определить функцию элемента независимо от относительного положения элемента в общей модели и от других функций элементов. Задание функции элемента через произвольное множество узловых значений и координат дает возможность использовать функции элемента для аппроксимации искомой функции на выбранных узлах.

Простейшим из элементов является одномерный элемент, схематически изображенный в виде отрезка. Простейший одномерный элемент имеет два узла, по одному на каждом конце элемента. Одномерные элементы более высокого порядка, трехузловые (квадратичные) и четырехузловые (кубичные), содержат по три и четыре узла соответственно. Порядок элемента, как будет показано, определяется порядком интерполяционного многочлена, с помощью которого аппроксимируется искомая функция. Так, на рис. 4.24, *a* представлены одномерные линейные элементы, а на рис. 4.24, *б* — квадратичные. В качестве двумерных элементов используются треугольники и четырехугольники. При этом количество узлов, которые содержит элемент, определяет его порядок и порядок его интерполяционного многочлена.

В соответствии с этими рассуждениями искомую функцию φ записывают на элементе в виде интерполяционного многочлена. В случае линейного одномерного элемента, содержащего два узла (так называемого одномерного симплекс-элемента), интерполяционный многочлен имеет вид

$$\varphi = \alpha_1 + \alpha_2 x,$$

а в случае двумерного линейного элемента, содержащего три узла (так называемого двумерного симплекс-элемента), интерполяционный многочлен имеет вид

$$\varphi = \alpha_1 + \alpha_2 x + \alpha_3 y.$$

В случае квадратичного одномерного элемента (так называемого одномерного комплекс-элемента) интерполяционный многочлен имеет вид

$$\varphi = \alpha_1 + \alpha_2 x + \alpha_3 x^2,$$

а в случае квадратичного двумерного элемента (так называемого двумерного комплекс-элемента) интерполяционный многочлен имеет вид

$$\varphi = \alpha_1 + \alpha_2 x + \alpha_3 y + \alpha_4 x^2 + \alpha_5 xy + \alpha_6 y^2.$$

Входящие в интерполяционные формулы коэффициенты α_i выражаются через значения функции в узлах. В частности, если ввести так называемые функции формы N_i и N_j , определяемые соотношениями

$$N_i = \frac{(X_j - x)}{L}, \quad N_j = \frac{(x - X_i)}{L},$$

где X_i и X_j — координаты узлов элемента, L — длина элемента, Φ_i и Φ_j — значения искомых функций в узлах i и j , $[N] = [N_i, N_j]$ — матричная строка, а $\{\Phi\} = \{\Phi_i \Phi_j\}$ — вектор-столбец искомых функций, то для одномерного симплекс-элемента имеем

$$\varphi = N_i \Phi_i + N_j \Phi_j = [N] \{\Phi\}.$$

Если ввести обозначения

$$N_i = \frac{1}{2A} (a_i + b_i x + c_i y), \quad \begin{aligned} a_i &= X_j Y_k - X_k Y_j, \\ b_i &= Y_j - Y_k, \\ c_i &= X_k - X_j, \end{aligned}$$

$$N_j = \frac{1}{2A} (a_j + b_j x + c_j y), \quad \begin{aligned} a_j &= X_k Y_i - X_i Y_k, \\ b_j &= Y_k - Y_i, \\ c_j &= X_i - X_k, \end{aligned}$$

$$N_k = \frac{1}{2A} (a_k + b_k x + c_k y), \quad \begin{aligned} a_k &= X_i Y_j - X_j Y_i, \\ b_k &= Y_i - Y_j, \\ c_k &= X_j - X_i, \end{aligned}$$

$$A = \begin{vmatrix} 1 & X_i & Y_i \\ 1 & X_j & Y_j \\ 1 & X_k & Y_k \end{vmatrix},$$

$$[N] = [N_i, N_j, N_k],$$

$$\{\Phi\} = \{\Phi_i, \Phi_j, \Phi_k\},$$

то для двумерного комплекс-элемента имеем

$$\varphi = N_i \Phi_i + N_j \Phi_j + N_k \Phi_k = [N] \{\Phi\}. \quad (4.86)$$

Функция формы N , равная единице в p -м узле, равна нулю во всех остальных узлах в соответствии с построением интерполяционных многочленов.

Следующим этапом метода конечных элементов является получение системы уравнений для нахождения неизвестных значений функций в узлах. С этой целью данному дифференциальному уравнению с заданными граничными условиями ставится в соответствие некоторый функционал, минимум которого достигается в том случае, при котором удовлетворяется исходное дифференциальное уравнение. Иными словами, вариационным уравнением Эйлера для данного функционала является наше исходное уравнение. Например, задача Дирихле для уравнения Лапласа с целью отыскания потенциала скорости

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0$$

эквивалентна отысканию минимума функционала

$$\chi = \frac{1}{2} \int_G \left[\left(\frac{\partial \phi}{\partial x} \right)^2 + \left(\frac{\partial \phi}{\partial y} \right)^2 \right] ds.$$

Минимизация этого функционала должна быть осуществлена на множестве узловых значений $\{\Phi\}$. Перед процедурой минимизации проведем некоторые преобразования. Представим функционал в виде

$$\chi = \frac{1}{2} \int_G \{g\}^T \{g\} ds, \quad (4.87)$$

где $\{g\}^T = \left\{ \frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial y} \right\}$. Рассмотрим область G , в которой ищется решение на M элементах, и заменим интеграл (4.87) суммой интегралов по этим элементам. Имеем

$$\chi = \sum_{m=1}^M \frac{1}{2} \int_G \{g^{(m)}\}^T \{g^{(m)}\} ds = \chi^{(1)} + \chi^{(2)} + \dots + \chi^{(M)}.$$

Потребуем теперь, чтобы функционал χ был минимальным на множестве узловых значений $\{\Phi\}$. Для этого необходимо, чтобы

$$\frac{\partial \chi}{\partial \{\Phi\}} = \frac{\partial}{\partial \{\Phi\}} \sum_{m=1}^M \chi^{(m)} = \sum_{m=1}^M \frac{\partial \chi^{(m)}}{\partial \{\Phi\}} = 0.$$

Поскольку согласно (4.80) имеем $\varphi^{(m)} = [N^{(m)}]\{\Phi\}$, то

$$\{g^{(m)}\} = \begin{pmatrix} \frac{\partial \varphi^{(m)}}{\partial x} \\ \frac{\partial \varphi^{(m)}}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial N_1^{(m)}}{\partial x} & \frac{\partial N_2^{(m)}}{\partial x} & \dots & \frac{\partial N_p^{(m)}}{\partial x} \\ \frac{\partial N_1^{(m)}}{\partial y} & \frac{\partial N_2^{(m)}}{\partial y} & \dots & \frac{\partial N_p^{(m)}}{\partial y} \end{pmatrix} \begin{Bmatrix} \Phi_1 \\ \Phi_2 \\ \vdots \\ \Phi_p \end{Bmatrix} \quad (4.88)$$

или

$$\{g^{(m)}\} = [B^{(m)}]\{\Phi\}, \quad (4.89)$$

где $B^{(m)}$ — матрица, содержащая информацию, связанную с функцией формы. Учитывая соотношения (4.88), (4.89), функционал χ можно переписать в виде

$$\chi^{(m)} = \frac{1}{2} \int_G \{\Phi\}^T [B^{(m)}]^T [B^{(m)}] \{\Phi\} ds.$$

Окончательно условие минимума функционала принимает вид

$$\frac{\partial \chi}{\partial \{\Phi\}} = \sum_{m=1}^M [K^{(m)}] \{\Phi\} = 0, \quad (4.90)$$

где $K^{(m)} = \int_{G^{(m)}} [B^{(m)}]^T [B^{(m)}] ds$. Из (4.86) следует, что матрицы

$B^{(m)}$ и $K^{(m)}$ определяются следующими соотношениями для двумерного комплекс-элемента:

$$B^{(m)} = \frac{1}{2A} \begin{bmatrix} b_i & b_j & b_k \\ c_i & c_j & c_k \end{bmatrix},$$

$$K^{(m)} = \frac{1}{4A} \begin{bmatrix} b_i^2 & b_i b_j & b_i b_k \\ b_j b_i & b_j^2 & b_j b_k \\ b_k b_i & b_k b_j & b_k^2 \end{bmatrix} + \frac{1}{4A} \begin{bmatrix} c_i^2 & c_i c_j & c_i c_k \\ c_j c_i & c_j^2 & c_j c_k \\ c_k c_i & c_k c_j & c_k^2 \end{bmatrix}.$$

Уравнения (4.90) представляют собой систему линейных алгебраических уравнений относительно значений искомой функции в узлах $\{\Phi_1, \Phi_2, \dots, \Phi_p\}$. Таким образом, решение краевой задачи для уравнения в частных производных сведено к решению системы линейных алгебраических уравнений. Матрица этой системы является матрицей ленточного типа, содержащей большое число нулевых элементов. Матрица системы будет трехдиагональной, если используются двумерные комплекс-элементы. При использовании квадратичных элементов и элементов более

высокого порядка матрица системы будет содержать большое число *диагоналей*. Матрица системы является хорошо обусловленной.

Важным свойством матрицы является ее симметрия и положительная определенность. Симметрия матрицы позволяет не хранить в памяти почти половину ее элементов. Положительная определенность матрицы означает, что элементы, стоящие на главной диагонали, всегда положительны и обычно много больше по величине, чем любой коэффициент соответствующей строки или столбца. Для симметричных положительно определенных матриц удается значительно сократить время вычислений и уменьшить влияние ошибок округления. При решении систем линейных уравнений такого типа эффективным является метод Гаусса и его различные варианты, в том числе метод прогонки.

Литература

1. Бахвалов Н. С. Численные методы. М.: Наука, 1973.
2. Вержбицкий В. М. Численные методы. Линейная алгебра и нелинейные уравнения. М.: Высшая школа, 2000.
3. Вержбицкий В. М. Численные методы. Математический анализ и обыкновенные дифференциальные уравнения. М.: Высшая школа, 2001.
4. Волков Е. А. Численные методы. М.: Наука, 1987.
5. Годунов С. К. и др. Численное решение многомерных задач газовой динамики / Под ред. С. К. Годунова. — М.: Наука, 1976.
6. Демидович Б. П., Марон И. А. Основы вычислительной математики. М.: Наука, 1970.
7. Калиткин Н. Н. Численные методы. М.: Наука, 1976.
8. Киреев В. И., Пантелеев А. В. Численные методы в примерах и задачах. М.: Изд-во МАИ, 2000.
9. Пиругов У. Г., Росляков Г. С. Численные методы газовой динамики. М.: Высшая школа, 1987.
10. Роуч П. Вычислительная гидродинамика. М.: Мир, 1980.
11. Самарский А. А., Попов Ю. П. Разностные схемы газовой динамики. М.: Наука, 1975.
12. Самарский А. А. Введение в численные методы. М.: Наука, 1982.
13. Самарский А. А. Теория разностных схем. М.: Наука, 1977.
14. Самарский А. А., Николаев Е. С. Методы решения сеточных уравнений. М.: Наука, 1978.
15. Сегерлинд Л. Применение метода конечных элементов. М.: Мир, 1979.

Именной указатель

- Адамар (*Hadamard J.S.*) 9
Адамс (*Adams J.C.*) 139
- Бернулли Якоб старший
(*Bernoulli J.*) 108
Больцман (*Boltzmann L.*) 107
Бюффон (*Buffon L.L.*) 106
- Вандермонд (*Vandermonde A.Th.*) 74
Вольтерра (*Volterra V.*) 149
- Галёркин Б.Г. 208
Гаусс (*Gauß C.F.*) 15
Гильберт (*Hilbert D.*) 89
Годунов С.К. 196
Грин (*Green G.*) 200
Гурса (*Goursat E.J.B.*) 154
Гюа де (*De Gua de Malves J.*) 33
- Декарт (*Descartes R.*) 32
Дирихле (*Dirichlet P.G.L.*) 120
- Евклид (*Ευκλείδης*) 58
Зейдель (*Seidel Ph.L.*) 28
- Картезий см. Декарт
Коши (*Cauchy A.L.*) 127
Крамер (*Cramer G.*) 13
Кристоффель (*Christoffel E.B.*) 150
Кронекер (*Kronecker L.*) 25
Крылов А.Н. 49
Курант (*Courant R.*) 173
Кутта (*Kutta M.W.*) 126
- Лангранж (*Lagrange J.L.*) 75
Лакс (*Lax P.D.*) 178
Ланцош (*Lanczos C.*) 49
- Лаплас (*Laplace P.S.*) 120
Лаппо-Данилевский И.А. 49
Леверье (*Le Verrier U.J.J.*) 49
Леви (*Lewy H.*) 173
Липшиц (*Lipschitz P.O.S.*) 63
- Мюллер см. Региомонтанус
- Нейман фон (*Neumann J. von*) 112
Нейман (*Neumann K.G.*) 157
Ньютона (*Newton I.*) 38
- Пуассон (*Poisson S.D.*) 156
- Региомонтанус (*Regiomontanus*) 42
Риман (*Riemann G.F.B.*) 197
Ричардсон (*Richardson L.F.*) 103
Ролль (*Rolle M.*) 81
Рунге (*Runge C.D.T.*) 82
- Симпсон (*Simpson Th.*) 102
- Тейлор (*Taylor B.*) 39
- Уlam (*Ulam S.M.*) 106
- Фредгольм (*Fredholm E.I.*) 148
Фридрихс (*Friedrichs K.O.*) 173
Фробениус (*Frobenius F.G.*) 54
Фурье (*Fourier J.B.J.*) 176
- Хильберт см. Гильберт
- Чебышёв П.Л. 83
- Эвклид см. Евклид
Эйлер (*Euler L.*) 133
Эйткен (*Aitken A.C.*) 47
- Якоби (*Jacobi C.G.J.*) 45

СБОРНИК ЗАДАЧ

Предисловие

Настоящее учебное пособие представляет собой сборник задач по основным разделам курса «Численные методы и алгоритмы». Излагаемый материал соответствует учебному пособию У. Г. Пирумова «Численные методы» [7].

Книга содержит пять основных разделов. Первый раздел посвящен численным методам линейной алгебры. Здесь рассмотрены базовые алгоритмы решения систем линейных алгебраических уравнений и решения проблемы собственных значений и собственных векторов матриц. Во втором разделе изучаются численные методы решения нелинейных уравнений и систем. Третий раздел посвящен вопросам приближения функций, численному дифференцированию и интегрированию. В четвертом разделе рассматриваются численные алгоритмы решения начальных и краевых задач для обыкновенных дифференциальных уравнений (ОДУ) и систем ОДУ, а в пятом — конечно-разностные алгоритмы решения задач для уравнений с частными производными.

Пособие предназначено для освоения читателями навыков применения численных методов при решении конкретных задач. С этой целью изложение материала построено по единой схеме, включающей постановку задачи, описание алгоритма решения, детально разобранные типовые примеры, демонстрирующие работу изучаемого алгоритма. В предлагаемых задачах требуется довести результат до конкретных численных значений. Во многих задачах в качестве ответа представлено точное (аналитическое) решение. Это позволит читателю самостоятельно провести анализ поведения погрешности численного решения в зависимости от параметров метода.

Материал, изложенный в книге, соответствует программам практических занятий по дисциплине «Численные методы и алгоритмы», преподаваемой в технических вузах. Для освоения представленного материала читатель должен обладать знанием основ линейной алгебры, математического анализа, теории обыкновенных дифференциальных уравнений и уравнений в частных

производных. Авторы рекомендуют читателям при решении задач активно использовать современные компьютерные технологии. Это касается как автоматизации расчетов на основе известных программных продуктов, так и разработки собственных программ. Применение компьютерных технологий сделает работу с книгой более эффективной и интересной.

Книга предназначена для студентов технических вузов, аспирантов и преподавателей, а также инженеров и научных работников, использующих в практической деятельности численные методы.

Глава 1

ЧИСЛЕННЫЕ МЕТОДЫ ЛИНЕЙНОЙ АЛГЕБРЫ

Рассматриваются численные методы решения систем линейных алгебраических уравнений (СЛАУ) и задач на собственные значения и собственные векторы матриц.

Среди численных методов алгебры существуют *прямые*, в которых решение получается после фиксированного числа операций, и *итерационные*, в которых результат достигается в процессе последовательных приближений.

1.1. Численные методы решения систем линейных алгебраических уравнений

В качестве прямых методов решения СЛАУ рассмотрим методы Гаусса и прогонки.

1.1.1. Метод Гаусса

При решении методом Гаусса матрица СЛАУ с помощью равносильных преобразований превращается в верхнюю треугольную матрицу в результате прямого хода. В обратном ходе определяются неизвестные.

Пусть дана СЛАУ

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n. \end{cases}$$

Выпишем расширенную матрицу системы:

→ *Ведущая строка*

↑ *Ведущий столбец*

	x_1	x_2	x_3	\dots	x_n	b
a_{11}	a_{12}	a_{13}	\dots	a_{n1}	b_1	
a_{21}	a_{22}	a_{23}	\dots	a_{n2}	b_2	
a_{31}	a_{32}	a_{33}	\dots	a_{n3}	b_3	
\dots						
a_{n1}	a_{n2}	a_{n3}	\dots	a_{nn}	b_n	

→ $\left(-\frac{a_{21}}{a_{11}} \right); \left(-\frac{a_{31}}{a_{11}} \right); \dots; \left(\frac{a_{n1}}{a_{11}} \right)$.

→ *1-й шаг*

На первом шаге алгоритма Гаусса выберем диагональный элемент $a_{11} \neq 0$ (если он равен нулю, то первую строку переставляем с какой-либо нижележащей строкой) и объявляем его *ведущим*, а соответствующую строку и столбец, на пересечении которых он стоит, — *ведущими*. Обнулим элементы a_{21}, \dots, a_{n1} ведущего столбца. Для этого сформируем числа $(-a_{21}/a_{11}), (-a_{31}/a_{11}), \dots, (-a_{n1}/a_{11})$. Умножая ведущую строку на число $(-a_{21}/a_{11})$, складывая со второй и ставя результат на место второй строки, получим вместо элемента a_{21} нуль, а вместо элементов $a_{2j}, j = 2, \dots, n$, b_2^1 — соответственно элементы

$$a_{2j}^1 = a_{2j} + a_{1j}(-a_{21}/a_{11}), \quad b_2^1 = b_2 + b_1(-a_{21}/a_{11})$$

и т.д.

Умножая ведущую строку на число $(-a_{n1}/a_{11})$, складывая с n -й строкой и ставя результат на место n -й строки, получим вместо элемента a_{n1} нуль, а остальные элементы этой строки будут иметь вид

$$a_{nj}^1 = a_{nj} + a_{1j}(-a_{n1}/a_{11}), \quad b_n^1 = b_n + b_1(-a_{n1}/a_{11}).$$

Сохраняя ведущую строку неизменной, получим в результате первого шага алгоритма Гаусса следующую матрицу:

	x_1	x_2	x_3	...	x_n	b
<i>Ведущая строка</i>	a_{11}	a_{12}	a_{13}	...	a_{1n}	b_1
	0	a_{22}^1	a_{23}^1	...	a_{2n}^1	b_2^1
	0	a_{32}^1	a_{33}^1	...	a_{3n}^1	b_3^1
			
<i>Ведущий столбец</i>	0	a_{n2}^1	a_{n3}^1	...	a_{nn}^1	b_n^1

$\left(-\frac{a_{31}}{a_{22}} \right); \dots; \left(\frac{a_{n2}^1}{a_{22}^1} \right)$.

2-й шаг

На втором шаге алгоритма Гаусса в качестве ведущего элемента выбирается элемент $a_{22}^1 \uparrow 0$ (если он равен нулю, то вторую строку взаимно меняем на нижележащую строку). Формируются следующие числа: $(a_{32}^1/a_{22}^1); \dots; (a_{n2}^1/a_{22}^1)$. Умножая ведущую строку на число (a_{32}^1/a_{22}^1) и складывая результат с третьей

строкой, получим вместо элемента a_{32}^1 нуль, а вместо элементов $a_{3j}^1 \quad j=3, 4, \dots, n$, b_3^1 — элементы $a_{3j}^2 \quad a_{3j}^1 + a_{2j}^1 \left(a_{32}^1 / a_{22}^1 \right)$, $b_3^2 \quad b_3^1 + b_2^1 \left(a_{32}^1 / a_{22}^1 \right)$ и т.д. Умножая ведущую строку на число $\left(a_{n2}^1 / a_{22}^1 \right)$, складывая результат с n -й строкой и ставя полученную сумму на место n -й строки, получим вместо элемента a_{n2}^1 нуль, а вместо элементов a_{nj}^1, b_n^1 — соответственно элементы $a_{nj}^2 \quad a_{nj}^1 + a_{2j}^1 \left(a_{n2}^1 / a_{22}^1 \right)$, $b_n^2 \quad b_n^1 + b_2^1 \left(a_{n2}^1 / a_{22}^1 \right), j = 3, 4, \dots, n$.

Сохраняя первую и вторую строки матрицы неизменными, получим в результате второго шага алгоритма Гаусса следующую матрицу:

	x_1	x_2	x_3	...	x_n	b
<i>Ведущая строка</i>	a_{11}	a_{12}	a_{13}	...	a_{1n}	b_1
	0	a_{22}^1	a_{23}^1	...	a_{2n}^1	b_2^1
	0	0	a_{33}^2	...	a_{3n}^2	b_3^2
			
<i>Ведущий столбец</i>	0	0	a_{n3}^2	...	a_{nn}^2	b_n^2

↗ 3-й шаг ... ↗ (n - 1)-й шаг

После $(n - 1)$ -го шага алгоритма Гаусса получаем следующую расширенную матрицу, содержащую верхнюю треугольную матрицу СЛАУ:

$$\left[\begin{array}{cccccc} x_1 & x_2 & x_3 & \dots & x_n & b \\ \hline a_{11} & a_{12} & a_{13} & \dots & a_{1n} & b_1 \\ 0 & a_{22}^1 & a_{23}^1 & \dots & a_{2n}^1 & b_2^1 \\ 0 & 0 & a_{33}^2 & \dots & a_{3n}^2 & b_3^2 \\ \dots & & & & & \\ 0 & 0 & 0 & \dots & a_{nn}^{n-1} & b_n^{n-1} \end{array} \right].$$

Прямой ход алгоритма Гаусса завершен.

В обратном ходе алгоритма Гаусса из последнего уравнения сразу определяется x_n , из предпоследнего — x_{n-1} и т.д. Из первого уравнения определяется x_1 :

$$\begin{cases} a_{nn}^{n-1}x_n = b_n^{n-1} \Rightarrow x_n, \\ a_{n-1,n-1}^{n-2}x_{n-1} + a_{n-1,n}^{n-2}x_n = b_{n-1}^{n-2} \Rightarrow x_{n-1}, \\ \dots \\ a_{11}x_1 + \dots + a_{1n}x_n = b_1 \Rightarrow x_1. \end{cases}$$

Замечание 1. Если элементы какой-либо строки матрицы системы в результате преобразований стали равными нулю, а правая часть не равна нулю, то СЛАУ несовместна, поскольку не выполняются условия теоремы Кронекера — Капелли.

Замечание 2. Если элементы какой-либо строки матрицы системы и правая часть в результате преобразований стали равными нулю, то СЛАУ совместна, но имеет бесконечное множество решений, получающееся с помощью метода Гаусса для СЛАУ порядка r , где r — ранг матрицы исходной СЛАУ.

Замечание 3. В результате прямого хода метода Гаусса можно вычислить определитель матрицы A исходной СЛАУ:

$$\det A = (-1)^p a_{11} \cdot a_{22}^1 \cdot a_{33}^2 \cdots \cdot a_{nn}^{n-1}.$$

При этом с помощью множителя $(-1)^p$, где p — число перестановок строк в процессе прямого хода, учитываются соответствующие переменны знаков вследствие перестановок строк.

Замечание 4. Метод Гаусса можно применить для обращения невырожденной ($\det A \neq 0$) матрицы.

Действительно, пусть требуется обратить невырожденную матрицу $A = [a_{ij}]$, $i, j = 1, 2, \dots, n$. Тогда, обозначив $A^{-1} = X$, $X = [x_{ij}]$, можно выписать матричное уравнение $AX = E$, где

$$E = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots \\ 0 & 0 & \dots & 1 \end{bmatrix} -$$

единичная матрица, и на его основе записать цепочку СЛАУ

$$A \cdot \begin{pmatrix} x_{11} \\ x_{21} \\ \dots \\ x_{n1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}, \quad A \cdot \begin{pmatrix} x_{12} \\ x_{22} \\ \dots \\ x_{n2} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \end{pmatrix}, \quad \dots, \quad A \cdot \begin{pmatrix} x_{1n} \\ x_{2n} \\ \dots \\ x_{nn} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix},$$

каждую из которых можно решить методом Гаусса. При этом, поскольку верхняя треугольная матрица для всех СЛАУ будет одной и той же, прямой ход метода Гаусса применяется лишь один раз. Строится следующая расширенная матрица:

$$\left| \begin{array}{cccc|c} x_{1n} & x_{2n} & \dots & x_{nn} & \\ \cdots & \cdots & & \cdots & \\ x_{12} & x_{22} & \dots & x_{n2} & \\ \hline x_{11} & x_{21} & \dots & x_{n1} & b_1 & b_2 & \dots & b_n \\ a_{11} & a_{12} & \dots & a_{1n} & 1 & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & a_{2n} & 0 & 1 & \dots & 0 \\ \cdots & \cdots & & \cdots & \cdots & \cdots & & \cdots \\ a_{n1} & a_{n2} & \dots & a_{nn} & 0 & 0 & \dots & 1 \end{array} \right|.$$

В результате применения $(n - 1)$ -го шага метода Гаусса получаем:

$$\left| \begin{array}{cccc|c} x_{1n} & x_{2n} & \dots & x_{nn} & \\ \cdots & \cdots & & \cdots & \\ x_{12} & x_{22} & \dots & x_{n2} & \\ \hline x_{11} & x_{21} & \dots & x_{n1} & b_1 & b_2 & \dots & b_n \\ a_{11} & a_{12} & \dots & a_{1n} & b_{11} & b_{12} & \dots & b_{1n} \\ 0 & a_{22}^1 & \dots & a_{2n}^1 & b_{21}^1 & b_{22}^1 & \dots & b_{2n}^1 \\ \cdots & \cdots & & \cdots & \cdots & \cdots & & \cdots \\ 0 & 0 & \dots & a_{nn}^{n-1} & b_{n1}^{n-1} & b_{n2}^{n-1} & \dots & b_{nn}^{n-1} \end{array} \right|.$$

При этом первый столбец $(x_{11} \ x_{21} \ \dots \ x_{n1})^T$ обратной матрицы определяется в обратном ходе метода Гаусса с правой частью b_1 , второй столбец $(x_{12} \ x_{22} \ \dots \ x_{n2})^T$ — с правой частью b_2 и т.д. Столбец $(x_{1n} \ x_{2n} \ \dots \ x_{nn})^T$ определяется с правой частью b_n .

Пример 1.1.

Методом Гаусса решить СЛАУ

$$\begin{cases} 10x_1 + x_2 + x_3 = 12, \\ 2x_1 + 10x_2 + x_3 = 13, \\ 2x_1 + 2x_2 + 10x_3 = 14. \end{cases}$$

Решение. Прямой ход:

$$\left(\begin{array}{ccc|c} x_1 & x_2 & x_3 & b \\ \hline 10 & 1 & 1 & 12 \\ 2 & 10 & 1 & 13 \\ 2 & 2 & 10 & 14 \end{array} \right) \xrightarrow{\begin{matrix} (-2/10); & (-2/10) \\ 1\text{-й шаг} \end{matrix}}$$

$$\left(\begin{array}{ccc|c} x_1 & x_2 & x_3 & b \\ 10 & 1 & 1 & 12 \\ 0 & 9.8 & 0.8 & 10.6 \\ 0 & 1.8 & 9.8 & 11.6 \end{array} \right) \xrightarrow{(-1.8/9.8)} \left(\begin{array}{ccc|c} x_1 & x_2 & x_3 & b \\ 10 & 1 & 1 & 12 \\ 0 & 9.8 & 0.8 & 10.6 \\ 0 & 0 & 9.653 & 9.653 \end{array} \right).$$

Обратный ход:

$$9.653x_3 = 9.653 \Rightarrow x_3 = 1,$$

$$9.8x_2 + 0.8x_3 = 10.6 \Rightarrow x_2 = 1,$$

$$10x_1 + x_2 + x_3 = 12 \Rightarrow x_1 = 1.$$

Ответ: $x_1 = x_2 = x_3 = 1$.

Пример 1.2.

Методом Гаусса вычислить определитель матрицы и обратить матрицу СЛАУ из примера 1.1.

Решение. Имеем

$$A = \begin{pmatrix} 10 & 1 & 1 \\ 2 & 10 & 1 \\ 2 & 2 & 10 \end{pmatrix};$$

$\det A \approx 10 \cdot 9.8 \cdot 9.65 = 945.994$ (точное значение — 946).

Прямой ход:

$$\left(\begin{array}{ccc|ccc} x_{13} & x_{23} & x_{33} & & & & \\ x_{12} & x_{22} & x_{23} & & & & \\ x_{11} & x_{21} & x_{31} & b_1 & b_2 & b_3 & \\ \hline 10 & 1 & 1 & 1 & 0 & 0 \\ 2 & 10 & 1 & 0 & 1 & 0 \\ 2 & 2 & 10 & 0 & 0 & 1 \end{array} \right) \xrightarrow{(-2/10); (-2/10)} \left(\begin{array}{ccc|ccc} x_{13} & x_{23} & x_{33} & & & & \\ x_{12} & x_{22} & x_{23} & & & & \\ x_{11} & x_{21} & x_{31} & b_1 & b_2 & b_3 & \\ \hline 10 & 1 & 1 & 1 & 0 & 0 \\ 0 & 9.8 & 0.8 & -0.2 & 1 & 0 \\ 0 & 1.8 & 9.8 & -0.2 & 0 & 1 \end{array} \right) \xrightarrow{1-\text{й шаг}}$$

$$\left(\begin{array}{ccc|ccc} x_{13} & x_{23} & x_{33} & & & & \\ x_{12} & x_{22} & x_{23} & & & & \\ x_{11} & x_{21} & x_{31} & b_1 & b_2 & b_3 & \\ \hline 10 & 1 & 1 & 1 & 0 & 0 \\ 0 & 9.8 & 0.8 & -0.2 & 1 & 0 \\ 0 & 0 & 9.653 & -0.163 & 0.184 & 1 \end{array} \right) \xrightarrow{2-\text{й шаг}}$$

$$\left(\begin{array}{ccc|ccc} x_{13} & x_{23} & x_{33} & & & & \\ x_{12} & x_{22} & x_{23} & & & & \\ x_{11} & x_{21} & x_{31} & b_1 & b_2 & b_3 & \\ \hline 10 & 1 & 1 & 1 & 0 & 0 \\ 0 & 9.8 & 0.8 & -0.2 & 1 & 0 \\ 0 & 0 & 9.653 & -0.163 & 0.184 & 1 \end{array} \right).$$

Обратный ход:

$$\begin{cases} 9.653x_{31} = -0.163, \\ 9.8x_{21} + 0.8x_{31} = -0.2, \\ 10x_{11} + x_{21} + x_{31} = 1. \end{cases} \quad \begin{cases} 9.653x_{32} = -0.814, \\ 9.8x_{22} + 0.8x_{32} = 1, \\ 10x_{12} + x_{22} + x_{32} = 0. \end{cases}$$

$$\begin{cases} 9.653x_{33} = 1, \\ 9.8x_{23} + 0.8x_{33} = 0, \\ 10x_{13} + x_{23} + x_{33} = 0. \end{cases}$$

Отсюда следует:

$$A^{-1} = \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} = \begin{pmatrix} 0.104 & -0.0085 & -0.0095 \\ -0.019 & 0.104 & -0.0085 \\ -0.0169 & -0.019 & 0.104 \end{pmatrix}.$$

Проверка:

$$\begin{aligned} A \cdot A^{-1} &= \begin{pmatrix} 10 & 1 & 1 \\ 2 & 10 & 1 \\ 2 & 2 & 10 \end{pmatrix} \cdot \begin{pmatrix} 0.104 & -0.0085 & -0.0095 \\ -0.019 & 0.104 & -0.0085 \\ -0.0169 & -0.019 & 0.104 \end{pmatrix} = \\ &= \begin{pmatrix} 0.104 & 0 & -0.0005 \\ 0.001 & 0.104 & -0.0085 \\ 0.001 & 0.001 & 1.004 \end{pmatrix}, \end{aligned}$$

т.е. с точностью до ошибки округления получена единичная матрица.

Замечание 5. Компьютерная реализация метода Гаусса часто осуществляется с использованием *LU-разложения матриц*.

LU-разложение матрицы A представляет собой разложение матрицы A в произведение нижней и верхней треугольных матриц, т.е.

$$A = LU,$$

где L — нижняя треугольная матрица (матрица, у которой все элементы, находящиеся выше главной диагонали, равны нулю, $l_{ij} = 0$ при $i < j$), U — верхняя треугольная матрица (матрица, у которой все элементы, находящиеся ниже главной диагонали, равны нулю, $u_{ij} = 0$ при $i > j$).

LU-разложение может быть построено с использованием описанного выше метода Гаусса. Рассмотрим k -й шаг метода Гаусса, на котором осуществляется присваивание нулевых значений (обнуление) поддиагональным элементам k -го столбца матрицы $A^{(k-1)}$. Как было описано, с этой целью используется следующая операция:

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - \mu_i^{(k)} a_{kj}^{(k-1)}, \quad \mu_i^{(k)} = \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}, \quad i = k+1, \dots, n, \quad j = k, \dots, n.$$

В терминах матричных операций такая операция эквивалентна умножению $A(k) = M_k A^{(k-1)}$, где элементы матрицы M_k определяются следующим образом:

$$m_{ij}^k = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \quad j \neq k, \\ -\mu_{k+1}^{(k)}, & i \neq j, \quad j = k, \end{cases}$$

т.е. матрица M_k имеет вид

$$\left(\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -\mu_{k+1}^{(k)} & 1 & 0 & 0 \\ \dots & & & & & \\ 0 & 0 & -\mu_n^{(k)} & 0 & 0 & 1 \end{array} \right).$$

Выражение для обратной операции запишется в виде $A^{(k-1)} = M_k^{-1} A^{(k)}$, здесь

$$M_k^{-1} = \left(\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \mu_{k+1}^{(k)} & 1 & 0 & 0 \\ \dots & & & & & \\ 0 & 0 & \mu_n^{(k)} & 0 & 0 & 1 \end{array} \right).$$

В результате прямого хода Гаусса получим $A^{(n-1)} = U$,

$$A = A^{(0)} = M_1^{-1} A^{(1)} = M_1^{-1} M_2^{-1} A^{(2)} = M_1^{-1} M_2^{-1} \dots M_n^{-1} A^{(n-1)},$$

где $A^{(n-1)} = U$ — верхняя треугольная матрица, а

$$L = M_1^{-1} M_2^{-1} \dots M_n^{-1}$$

— нижняя треугольная матрица, имеющая вид

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \mu_2^{(1)} & 1 & 0 & 0 & 0 & 0 \\ \mu_3^{(1)} & \mu_3^{(2)} & 1 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \mu_{k+1} & 1 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \mu_n^{(1)} & \mu_n^{(2)} & \mu_n^{(k)} & \mu_n^{(k+1)} & \mu_n^{(n-1)} & 1 \end{pmatrix}.$$

Таким образом, получено искомое разложение $A = LU$.

В частности, для рассматриваемого выше примера 1.1 LU -разложение матрицы A имеет вид

$$A = \begin{pmatrix} 10 & 1 & 1 \\ 2 & 10 & 1 \\ 2 & 2 & 10 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0.2 & 1 & 0 \\ 0.2 & 0.18 & 1 \end{pmatrix} \cdot \begin{pmatrix} 10 & 1 & 1 \\ 0 & 9.8 & 0.8 \\ 0 & 0 & 9.65 \end{pmatrix} = LU.$$

В дальнейшем LU -разложение может быть эффективно использовано при решении систем линейных алгебраических уравнений вида $A\mathbf{x} = \mathbf{b}$. Действительно, подставляя LU -разложение в СЛАУ, получим $L\mathbf{Ux} = \mathbf{b}$, или $L\mathbf{z} = \mathbf{b}$. Далее процесс решения СЛАУ сводится к двум простым этапам.

На первом этапе решается СЛАУ $L\mathbf{z} = \mathbf{b}$. Поскольку матрица системы — нижняя треугольная, решение можно записать в явном виде:

$$z_1 = b_1, \quad z_i = b_i - \sum_{j=1}^{i-1} l_{ij} z_j, \quad i = 2, 3, \dots, n.$$

На втором этапе решается СЛАУ $\mathbf{Ux} = \mathbf{z}$ с верхней треугольной матрицей. Здесь, как и на предыдущем этапе, решение представляется в явном виде:

$$x_n = \frac{z_n}{u_{nn}}, \quad x_i = \frac{1}{u_{ii}} \left(z_i - \sum_{j=i+1}^n u_{ij} x_j \right), \quad i = n-1, \dots, 1.$$

Отметим, что второй этап эквивалентен обратному ходу метода Гаусса, тогда как первый соответствует преобразованию правой части СЛАУ в процессе прямого хода.

Замечание 6. В программных продуктах для научных расчетов обще-принятой считается реализация метода Гаусса с помощью LU -разложения с выбором ведущего элемента. Выбор ведущего (главного) элемента необходим, чтобы избежать деления на малые или нулевые величины при обнулении поддиагональных элементов. Строго говоря, в этой модификации метода LU -разложение выполняется над матрицей PA , где P – матрица перестановок.

Пусть на k -м шаге метода известна матрица $A^{(k-1)}$

$$A^{(k-1)} = \begin{pmatrix} a_{11}^{k-1} & a_{12} & \dots & a_{1k-1} & a_{1k} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2k-1} & a_{2k} & \dots & a_{2n} \\ 0 & 0 & \dots & a_{3k-1} & a_{3k} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{k-1k-1} & a_{k-1k} & \dots & a_{k-1n} \\ 0 & 0 & \dots & 0 & a_{kk} & \dots & a_{kn} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{nk} & \dots & a_{nn} \end{pmatrix}.$$

Для матрицы $A^{(k-1)}$ производится выбор главного элемента a_{kk} . Главным элементом будем считать максимальный по абсолютному значению элемент из подмножества элементов k -го столбца. Пусть

$a_{mk}^{(k-1)} = \max |a_{ik}^{(k-1)}|, i \geq k$, тогда матрицу перестановок $P^{(k)}$ получим перестановкой в единичной матрице строк m и k . Матрица $\hat{A}^{(k-1)}$ с выбранным главным элементом $\hat{a}_{kk}^{(k-1)} = a_{mk}^{(k-1)}$ будет равна $\hat{A}^{(k-1)} = P^{(k-1)} A^{(k-1)}$.

Далее будем производить замещение поддиагональных элементов k -го столбца матрицы $\hat{A}^{(k-1)}$ согласно следующим правилам:

$$a_{ij}^{(k)} = \hat{a}_{ij}^{(k-1)} - \mu_i^{(k)} \hat{a}_{kj}^{(k-1)}, \mu_i^{(k)} = \frac{\hat{a}_{ik}^{(k-1)}}{\hat{a}_{kk}^{(k-1)}}, i = k+1, \dots, n, j = k, \dots, n.$$

Это выражение эквивалентно произведению матриц

$$A^{(k)} = M_k \hat{A}^{(k-1)} = M_k P^{(k)} A^{(k-1)},$$

где

$$M_k = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & -\mu_{k+1}^{(k)} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & -\mu_n^{(k)} & \dots & 1 \end{pmatrix}.$$

Нижнетреугольная матрица на k -м шаге метода определяется выражением

$$L^{(k)} = P^{(k)} L^{(k-1)} P^{(k)} M_k^{-1}.$$

На первом шаге $L^{(1)} = M_1^{-1}$.

Алгоритм LU-разложения завершается после $(n-1)$ -го шага:

$$L = L^{(n-1)}, U = A^{(n-1)}, P = P^{(n-2)} \times \dots \times P^{(1)}.$$

Окончательно LU-разложение имеет вид:

$$PA = LU.$$

Отметим, что $\det P = (-1)^P$. Этим объясняется наличие множителя $(-1)^P$ в формуле вычисления определителя.

Решение СЛАУ вида $Ax = b$ осуществляется в два этапа. На первом этапе решается СЛАУ с нижнетреугольной матрицей $Lz = Pb$. На втором этапе решается СЛАУ с верхнеугольной матрицей $Ux = z$.

1.1.2. Метод прогонки

Метод прогонки является частным случаем метода Гаусса. Он применяется для решения СЛАУ с трехдиагональными матрицами. Такие системы часто возникают при конечно-разностной аппроксимации задач для обыкновенных дифференциальных уравнений (ОДУ) и уравнений в частных производных второго порядка. Рассмотрим СЛАУ

$$\left\{ \begin{array}{l} b_1x_1 + c_1x_2 = d_1, \\ a_2x_1 + b_2x_2 + c_2x_3 = d_2, \\ a_3x_2 + b_3x_3 + c_3x_4 = d_3, \\ \dots \\ a_{n-1}x_{n-2} + b_{n-1}x_{n-1} + c_{n-1}x_n = d_{n-1}, \\ a_nx_{n-1} + b_nx_n = d_n, \end{array} \right. \quad (1.1)$$

при этом будем полагать, что

$$\begin{aligned} a_1 &= 0, \\ c_n &= 0. \end{aligned}$$

Решение (1.1) будем искать в виде

$$x_i = P_i x_{i+1} + Q_i, \quad i = 1, 2, \dots, n, \quad (1.2)$$

где P_i, Q_i — прогоночные коэффициенты, подлежащие определению. Для этого выразим x_1 из первого уравнения системы (1.1) через x_2 и получим

$$x_1 = \frac{-c_1}{b_1}x_2 + \frac{d_1}{b_1} = P_1 x_2 + Q_1, \quad (1.3)$$

откуда следует

$$P_1 = \frac{-c_1}{b_1}, \quad Q_1 = \frac{d_1}{b_1}.$$

Из второго уравнения системы (1.1) с помощью (1.3) выразим x_2 через x_3 , получим

$$x_2 = \frac{-c_2}{b_2 + a_2 P_1} x_3 + \frac{d_2 - a_2 Q_1}{b_2 + a_2 P_1} = P_2 x_3 + Q_2,$$

откуда следует

$$P_2 = \frac{-c_2}{b_2 + a_2 P_1}, \quad Q_2 = \frac{d_2 - a_2 Q_1}{b_2 + a_2 P_1}.$$

Продолжая этот процесс, получим из i -го уравнения системы (1.1)

$$x_i = \frac{-c_i}{b_i + a_i P_{i-1}} x_{i+1} + \frac{d_i - a_i Q_{i-1}}{b_i + a_i P_{i-1}},$$

следовательно,

$$P_i = \frac{-c_i}{b_i + a_i P_{i-1}}, \quad Q_i = \frac{d_i - a_i Q_{i-1}}{b_i + a_i P_{i-1}}.$$

Из последнего уравнения системы (1.1) имеем

$$x_n = \frac{-c_n}{b_n + a_n P_{n-1}} x_{n+1} + \frac{d_n - a_n Q_{n-1}}{b_n + a_n P_{n-1}} = 0 \cdot x_{n+1} + Q_n,$$

т.е., так как $c_n = 0$,

$$P_n = 0, \quad Q_n = \frac{d_n - a_n Q_{n-1}}{b_n + a_n P_{n-1}} = x_n.$$

Таким образом, прямой ход определения прогоночных коэффициентов $P_i, Q_i, i = 1, 2, \dots, n$, завершен. В результате эти коэффициенты вычисляются по следующим формулам:

$$P_i = \frac{-c_i}{b_i + a_i P_{i-1}}, \quad Q_i = \frac{d_i - a_i Q_{i-1}}{b_i + a_i P_{i-1}}, \quad i = 2, 3, \dots, n-1, \quad (1.4)$$

$$P_1 = \frac{-c_1}{b_1}, \quad Q_1 = \frac{d_1}{b_1}, \quad (1.5)$$

$$P_n = 0, \quad Q_n = \frac{d_n - a_n Q_{n-1}}{b_n + a_n P_{n-1}} = x_n. \quad (1.6)$$

Обратный ход метода прогонки осуществляется в соответствии с выражением (1.2)

$$\begin{cases} x_n = P_n x_{n+1} + Q_n = 0 \cdot x_{n+1} + Q_n = Q_n, \\ x_{n-1} = P_{n-1} x_n + Q_{n-1}, \\ x_{n-2} = P_{n-2} x_{n-1} + Q_{n-2}, \\ \dots \\ x_1 = P_1 x_2 + Q_1. \end{cases} \quad (1.7)$$

Формулы (1.4)–(1.7) являются формулами *правой прогонки*. Аналогично, начиная с последнего уравнения СЛАУ (1.1), можно вывести формулы *левой прогонки*.

Общее число операций в методе прогонки равно $8n + 1$, т.е. пропорционально числу уравнений. Такие методы решения СЛАУ называют *экономичными*. Для сравнения число операций в методе Гаусса пропорционально n^3 [1].

Для устойчивости метода прогонки (1.4)–(1.7) достаточно выполнения следующих условий:

$$\begin{aligned} a_i \neq 0, \quad c_i \neq 0, \quad & i = 2, 3, \dots, n-1, \\ |b_i| \geq |a_i| + |c_i|, \quad & i = 1, 2, \dots, n, \end{aligned} \quad (1.8)$$

причем строгое неравенство имеет место хотя бы при одном i . Здесь устойчивость понимается в смысле ненакопления погрешности решения в ходе вычислительного процесса при малых погрешностях входных данных (правых частей и элементов матрицы СЛАУ).

Пример 1.3. Методом прогонки решить СЛАУ

$$\begin{cases} 8x_1 - 2x_2 = 6, \\ -x_1 + 6x_2 - 2x_3 = 3, \\ 2x_2 + 10x_3 - 4x_4 = 8, \\ -x_3 + 6x_4 = 5. \end{cases}$$

Решение. Имеем

$$\begin{aligned} P_1 &= \frac{-c_1}{b_1} = \frac{2}{8} = 0.25; & Q_1 &= \frac{d_1}{b_1} = 0.75; \\ P_2 &= \frac{-c_2}{b_2 + a_2 P_1} = \frac{2}{6 - 1 \cdot 0.25} = & Q_2 &= \frac{d_2 - a_2 Q_1}{b_2 + a_2 P_1} = \frac{(3 + 1 \cdot 0.75)}{5.75} = \\ &= 0.3478; & &= 0.6522; \\ P_3 &= \frac{-c_3}{b_3 + a_3 P_2} = 0.374; & Q_3 &= \frac{d_3 - a_3 Q_2}{b_3 + a_3 P_2} = 0.626; \\ P_4 &= 0 (c_4 = 0); & Q_4 &= \frac{d_4 - a_4 Q_3}{b_4 + a_4 P_3} = 1.0; \\ x_4 &= P_4 x_5 + Q_4 = 1.0; & x_2 &= P_2 x_3 + Q_2 = 1.0; \\ x_3 &= P_3 x_4 + Q_3 = 1.0; & x_1 &= P_1 x_2 + Q_1 = 1.0. \end{aligned}$$

1.1.3. Нормы векторов и матриц

Для исследования сходимости численных методов решения задач линейной алгебры вводятся понятия нормы векторов и матриц.

Нормой вектора $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ (обозначают $\|\mathbf{x}\|$) в n -мерном вещественном пространстве векторов $\mathbf{x} \in R^n$ называют неотри-

цательное число, вычисляемое с помощью компонент вектора и обладающее следующими свойствами:

- $\|\mathbf{x}\| \geq 0$ $\|\mathbf{x}\| = 0$ тогда и только тогда, когда \mathbf{x} – нулевой вектор; для нулевого вектора введем обозначение \emptyset ;
- $\|\alpha \cdot \mathbf{x}\| = |\alpha| \cdot \|\mathbf{x}\|$ для любых действительных чисел α ;
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Нормой матрицы $A_{n \times n}$ (обозначается $\|A\|$) с вещественными элементами в пространстве матриц называют неотрицательное число, вычисляемое с помощью элементов матрицы и обладающее следующими свойствами:

- $\|A\| > 0$ $\|A\| = 0$ тогда и только тогда, когда A – нулевая матрица; для нулевой матрицы введем обозначение Θ ;
- $\|\alpha \cdot A\| = |\alpha| \cdot \|A\|$ для любых действительных чисел α ;
- $\|A + B\| \leq \|A\| + \|B\|$ для всех $n \times n$ матриц A и B рассматриваемого пространства;
- $\|A \cdot B\| \leq \|A\| \cdot \|B\|$ для всех $n \times n$ матриц A и соответствующих матриц B .

Как видно из последнего свойства (если в качестве матрицы B использовать вектор \mathbf{x}), норма матриц должна быть согласована с нормой векторов. Это согласование осуществляется связью

$$\|Ax\|'' \|A\| \|\mathbf{x}\|. \quad (1.9)$$

Наиболее употребительными являются следующие нормы векторов:

$$\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|, \quad (1.10)$$

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{(\mathbf{x}, \mathbf{x})}, \quad (1.11)$$

$$\|\mathbf{x}\|_c = \max_i |x_i|. \quad (1.12)$$

Наиболее распространенными согласованными с ними с помощью связи (1.9) нормами матриц будут соответственно:

$$\|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|, \quad (1.13)$$

$$\|A\|_2 = \sqrt{\sum_{i,j=1}^n a_{ij}^2}, \quad (1.14)$$

$$\|A\|_c = \max_i \sum_{j=1}^n |a_{ij}|. \quad (1.15)$$

Отметим, что норма (1.15) согласована со всеми приведенными выше нормами векторов.

Для исследования погрешностей, возникающих при решении СЛАУ, вводят понятие *числа обусловленности матрицы*:

$$\chi(A) = \|A\| \cdot \|A^{-1}\|.$$

Число обусловленности характеризует степень зависимости относительной погрешности решения СЛАУ от погрешности входных данных (правые части, элементы матрицы). Можно показать, что для ненулевых векторов \mathbf{x} справедливы следующие неравенства:

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \chi(A) \frac{\|\Delta \mathbf{b}\|}{\|\mathbf{b}\|}, \quad \frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \chi(A) \frac{\|\Delta A\|}{\|A + \Delta A\|}.$$

Таким образом, чем больше число обусловленности, тем сильнее влияние погрешности входных данных на конечный результат. Матрица считается плохо обусловленной, если $\chi(A) >> 1$.

Если в качестве нормы матрицы принять ее спектральный радиус $\max_i |\lambda_i|$ (см. раздел 1.2), то

$$\chi(A) = \max_i |\lambda_i| \frac{1}{\min_i |\lambda_i|} \geq 1,$$

поскольку спектральный радиус обратной матрицы A^{-1} равен обратной величине модуля минимального собственного значения исходной матрицы.

Пример 1.4. Для матрицы A и вектора \mathbf{b} вычислить различные нормы $\|A\|_1, \|A\|_2, \|A\|_c$. Проверить выполнение условия согласованности норм

$\|Ax\|'' \|A\|\|x\|$ для различных комбинаций норм. Вычислить число обусловленности матрицы A , если

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

Решение. Вычислим соответствующие нормы:

$$\|b\|_1 = |3| + |-4| = 7, \quad \|b\|_2 = \left(3^2 + (-4)^2\right)^{1/2} = 5, \quad \|b\|_c = \max(|3|, |-4|) = 4;$$

$$\|A\|_1 = \max(|-1| + |3|, |2| + |-5|) = 7,$$

$$\|A\|_2 = \left((-1)^2 + 3^2 + 2^2 + (-5)^2\right)^{1/2} = \sqrt{39};$$

$$\|A\|_c = \max(|-1| + |2|, |3| + |-5|) = 8.$$

Для проверки условия согласованности вычислим различные нормы вектора

$$c = Ab = \begin{pmatrix} -11 \\ 29 \end{pmatrix}:$$

$$\|c\|_1 = |-11| + |29| = 40, \quad \|c\|_2 = \left((-11)^2 + 29^2\right)^{1/2} = \sqrt{962},$$

$$\|c\|_c = \max(|-11|, |29|) = 29.$$

Легко убедиться в том, что условие согласованности выполняется для следующих комбинаций норм:

$$\|c\|_1 = 40 \leq \|A\|_1 \|b\|_1 = 7 \cdot 7 = 49, \quad \|c\|_2 = \sqrt{962} \leq \|A\|_2 \|b\|_2 = \sqrt{39} \cdot 5 = \sqrt{975},$$

$$\|c\|_c = 29 \leq \|A\|_c \|b\|_c = 8 \cdot 4 = 32.$$

Кроме того, известно, что матричная норма $\|A\|_c$ согласована со всеми введенными выше нормами векторов. В данном примере это подтверждается выполнением неравенств

$$\|c\|_1 = 40 \leq \|A\|_1 \|b\|_1 = 8 \cdot 7 = 56, \quad \|c\|_2 = \sqrt{962} \leq \|A\|_2 \|b\|_2 = 8 \cdot 5 = 40.$$

В то же время использование ряда других комбинаций норм матрицы и вектора приводит в данном случае к нарушению условия согласованности

$$\|c\|_c = 29 > \|A\|_1 \|b\|_c = 7 \cdot 4 = 28, \quad \|c\|_c = 29 > \|A\|_2 \|b\|_c = \sqrt{39} \cdot 4.$$

Рассмотренный пример наглядно иллюстрирует важность использования согласованных норм матрицы и вектора.

Вычислим число обусловленности матрицы A , взяв в качестве нормы матрицы $\|\cdot\|_c$. Для этого найдем сначала обратную матрицу

$$A^{-1} = \begin{pmatrix} 5 & 2 \\ 3 & 1 \end{pmatrix}$$

и вычислим ее норму

$$\|A^{-1}\|_c = \max(|5| + |2|, |3| + |1|) = 7.$$

В результате получим

$$\chi(A) = \|A\|_c \|A^{-1}\|_c = 8 \cdot 7 = 56.$$

1.1.4. Итерационные методы решения СЛАУ

Метод простых итераций. При большом числе уравнений прямые методы решения СЛАУ (за исключением метода прогонки) становятся труднореализуемыми на ЭВМ, прежде всего из-за сложности хранения и обработки матриц большой размерности. В то же время характерной особенностью ряда часто встречающихся в прикладных задачах СЛАУ является разреженность матриц. Число ненулевых элементов таких матриц мало по сравнению с их размерностью. Для решения СЛАУ с разреженными матрицами предпочтительнее использовать итерационные методы.

Методы последовательных приближений, в которых при вычислении последующего приближения используются предыдущие, уже известные приближенные решения, называются *итерационными*.

Рассмотрим СЛАУ с невырожденной матрицей ($\det A \neq 0$)

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n. \end{cases} \quad (1.16)$$

Приведем СЛАУ к эквивалентному виду

$$\begin{cases} x_1 = \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1n}x_n + \beta_1, \\ x_2 = \alpha_{21}x_1 + \alpha_{22}x_2 + \dots + \alpha_{2n}x_n + \beta_2, \\ \dots \\ x_n = \alpha_{n1}x_1 + \alpha_{n2}x_2 + \dots + \alpha_{nn}x_n + \beta_n. \end{cases} \quad (1.17)$$

или, в векторно-матричной форме,

$$\mathbf{x} = \alpha \mathbf{x} + \beta, \\ \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}, \quad \alpha = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \dots & \dots & \dots \\ \alpha_{n1} & \dots & \alpha_{nn} \end{pmatrix}. \quad (1.18)$$

Такое приведение может быть выполнено различными способами. Одним из наиболее распространенных является следующий.

Разрешим систему (1.16) относительно неизвестных при ненулевых диагональных элементах $a_{ii} \neq 0$, $i = 1, 2, \dots, n$; если какой-либо коэффициент на главной диагонали равен нулю, достаточно соответствующее уравнение поменять местами с любым другим уравнением. Получим следующие выражения для компонентов вектора β и матрицы α эквивалентной системы:

$$\beta_i = \frac{b_i}{a_{ii}}; \quad \alpha_{ij} = -\frac{a_{ij}}{a_{ii}}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, n, \quad i \neq j; \\ \alpha_{ii} = 0, \quad i = 1, 2, \dots, n. \quad (1.18)$$

При таком способе приведения исходной СЛАУ к эквивалентному виду метод простых итераций носит название *метода Якоби*.

В качестве нулевого приближения $\mathbf{x}^{(0)}$ вектора неизвестных примем вектор правых частей $\mathbf{x}^{(0)} = \beta$ или $\left(x_1^{(0)} x_2^{(0)} \dots x_n^{(0)} \right)^T = (\beta_1 \beta_2 \dots \beta_n)^T$. Тогда *метод простых итераций* примет вид

$$\begin{cases} \mathbf{x}^{(0)} = \beta, \\ \mathbf{x}^{(1)} = \alpha \mathbf{x}^{(0)} + \beta, \\ \mathbf{x}^{(2)} = \alpha \mathbf{x}^{(1)} + \beta, \\ \dots \\ \mathbf{x}^{(k)} = \alpha \mathbf{x}^{(k-1)} + \beta. \end{cases} \quad (1.19)$$

Из (1.19) видно преимущество итерационных методов по сравнению, например, с рассмотренным методом Гаусса. В вычислительном процессе участвуют только произведения матрицы

на вектор. Это позволяет работать только с ненулевыми элементами матрицы, что значительно упрощает процесс хранения и обработки матриц.

Имеет место следующее *достаточное* условие сходимости метода простых итераций.

- Метод простых итераций (1.19) сходится к единственному решению СЛАУ (1.17) (а следовательно, и к решению исходной СЛАУ (1.16)) при любом начальном приближении $\mathbf{x}^{(0)}$, если какая-либо норма матрицы α эквивалентной системы меньше единицы $\|\alpha\| < 1$.

Если используется метод Якоби (выражения (1.18) для эквивалентной СЛАУ), то достаточным условием сходимости является *диагональное преобладание матрицы A*, т.е. $|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}| \forall i$

(для каждой строки матрицы A модули элементов, стоящих на главной диагонали, больше суммы модулей недиагональных элементов). Очевидно, что в этом случае $\|\alpha\|_c$ меньше единицы и, следовательно, итерационный процесс (1.19) сходится.

Приведем также *необходимое и достаточное* условие сходимости метода простых итераций.

- Для сходимости итерационного процесса (1.19) необходимо и достаточно, чтобы спектр матрицы α эквивалентной системы лежал внутри круга с радиусом, равным единице.

При выполнении достаточного условия сходимости оценка погрешности решения на k -й итерации дается выражением

$$\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \leq \varepsilon^{(k)} = \frac{\|\alpha\|}{1 - \|\alpha\|} \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|, \quad (1.20)$$

где \mathbf{x}^* — точное решение СЛАУ. Процесс итераций останавливается про выполнении условия $\varepsilon^{(k)} \leq \varepsilon$, где ε — задаваемая вычислителем погрешность.

Принимая во внимание, что из (1.20) следует неравенство

$$\|\mathbf{x}^{(k)} - \mathbf{x}^*\| \leq \frac{\|\alpha\|^k}{1 - \|\alpha\|} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|,$$

можно получить априорную оценку необходимого для достижения заданной точности числа итераций. При использовании

в качестве начального приближения вектора β такая оценка определяется неравенством

$$\frac{\|\alpha\|^{k+1}}{1-\|\alpha\|} \|\beta\| \leq \varepsilon,$$

из которого получаем априорную оценку числа итераций k при $\|\alpha\| < 1$:

$$k + 1 \geq \frac{\lg \varepsilon - \lg \|\beta\| + \lg(\|\alpha\|)}{\lg \|\alpha\|}.$$

Следует подчеркнуть, что это неравенство дает завышенное число итераций k , поэтому редко используется на практике.

Замечание. Поскольку $\|\alpha\| < 1$ является только достаточным (не необходимым) условием сходимости метода простых итераций, итерационный процесс может сходиться и в случае, если оно не выполнено. Тогда критерием окончания итераций может служить неравенство $\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\| < \varepsilon$. Отметим, однако, что при этом не гарантируется достижение заданной точности.

Пример 1.5. Методом простых итераций с погрешностью $\varepsilon = 0.01$ решить СЛАУ

$$\begin{aligned} 10x_1 + x_2 + x_3 &= 12, \\ 2x_1 + 10x_2 + x_3 &= 13, \\ 2x_1 + 2x_2 + 10x_3 &= 14. \end{aligned}$$

Решение. Приведем СЛАУ к эквивалентному виду:

$$\begin{cases} x_1 = 1.2 - 0.1x_2 - 0.1x_3, \\ x_2 = 1.3 - 0.2x_1 - 0.1x_3, \\ x_3 = 1.4 - 0.2x_1 - 0.2x_2 \end{cases}$$

или $\mathbf{x} = \alpha\mathbf{x} + \beta$, где

$$\alpha = \begin{pmatrix} 0 & -0.1 & -0.1 \\ -0.2 & 0 & -0.1 \\ -0.2 & -0.2 & 0 \end{pmatrix}; \quad \beta = (1.2 \quad 1.3 \quad 1.4)^T.$$

Здесь $\|\alpha\|_c = 0.4 < 1$, следовательно, достаточное условие сходимости метода простых итераций выполнено.

Итерационный процесс выглядит следующим образом:

$$\mathbf{x}^{(0)} = \beta;$$

$$\mathbf{x}^{(1)} = \alpha\beta + \beta = (0.93 \quad 0.92 \quad 0.9)^T; \quad \varepsilon^{(1)} = 0.333 > \varepsilon;$$

$$\mathbf{x}^{(2)} = \alpha\mathbf{x}^{(1)} + \beta = (1.018 \quad 1.024 \quad 1.03)^T; \quad \varepsilon^{(2)} = 0.0867 > \varepsilon;$$

$$\mathbf{x}^{(3)} = \alpha\mathbf{x}^{(2)} + \beta = (0.9946 \quad 0.9934 \quad 0.9916)^T; \quad \varepsilon^{(3)} = 0.0256 > \varepsilon;$$

$$\mathbf{x}^{(4)} = \alpha\mathbf{x}^{(3)} + \beta = (1.0015 \quad 1.00192 \quad 1.0024)^T; \quad \varepsilon^{(4)} = 0.0072 > \varepsilon.$$

Таким образом, вычислительный процесс завершен за четыре итерации. Отметим, что точное решение исходной СЛАУ в данном случае известно: $\mathbf{x}^* = (1 \ 1 \ 1)^T$. Отсюда следует, что заданной погрешности $\varepsilon = 0.01$ удовлетворяло решение, полученное уже на третьей итерации. Но поскольку для вычисления погрешности используется оценочное выражение (1.20) (в данном случае $\|\mathbf{x}(3) - \mathbf{x}^*\| \leq \varepsilon^{(3)}$, при этом $\varepsilon^{(3)} > \varepsilon$, хотя $\|\mathbf{x}(3) - \mathbf{x}^*\| \leq \varepsilon$), процесс останавливается только на четвертой итерации.

Отметим также, что априорная оценка необходимого количества итераций в данной задаче дает: $(k+1) \geq (-2 + \lg 0.6 - \lg 1.4)/\lg 0.4 = 5.95$, т.е. для достижения погрешности $\varepsilon = 0.01$, согласно априорной оценке, необходимо сделать не менее пяти итераций. Это иллюстрирует характерную для априорной оценки тенденцию к завышению числа итераций.

Метод Зейделя решения СЛАУ. Метод простых итераций сходится довольно медленно. Для его ускорения существует *метод Зейделя*, заключающийся в том, что при вычислении компоненты $x_i^{(k+1)}$ вектора неизвестных на $(k+1)$ -й итерации используются компоненты $x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{i-1}^{(k+1)}$, уже вычисленные на $(k+1)$ -й итерации. Значения остальных компонент $x_{i+1}^{(k)}, x_{i+2}^{(k)}, \dots, x_n^{(k)}$ берутся из предыдущей итерации. Так же, как и в методе простых итераций, строится эквивалентная СЛАУ (1.17) и за начальное приближение принимается вектор правых частей $\mathbf{x}_0 = (\beta_1, \beta_2, \dots, \beta_n)^T$. Тогда метод Зейделя для известного вектора $(x_1^k, x_2^k, \dots, x_n^k)^T$ на k -й итерации имеет вид

$$\mathbf{x} = \alpha\mathbf{x} + \beta;$$

$$\begin{cases} x_1^{(k+1)} = \alpha_{11}x_1^{(k)} + \alpha_{12}x_2^{(k)} + \dots + \alpha_{1n}x_n^{(k)} + \beta_1, \\ x_2^{(k+1)} = \alpha_{21}x_1^{(k+1)} + \alpha_{22}x_2^{(k)} + \dots + \alpha_{2n}x_n^{(k)} + \beta_2, \\ x_3^{(k+1)} = \alpha_{31}x_1^{(k+1)} + \alpha_{32}x_2^{(k+1)} + \alpha_{33}x_3^{(k)} + \dots + \alpha_{3n}x_n^{(k)} + \beta_3, \\ \dots \\ x_n^{(k+1)} = \alpha_{n1}x_1^{(k+1)} + \alpha_{n2}x_2^{(k+1)} + \dots + \alpha_{nn-1}x_{n-1}^{(k+1)} + \alpha_{nn}x_n^{(k)} + \beta_n. \end{cases}$$

Из этой системы видно, что $\mathbf{x} = B\mathbf{x}^{(k+1)} + C\mathbf{x}^{(k)} + \beta$, где B — нижняя треугольная матрица с диагональными элементами, равными нулю, а C — верхняя треугольная матрица с диагональными элементами, отличными от нуля, $\alpha = B + C$. Следовательно,

$$(E - B)\mathbf{x}^{(k+1)} = C\mathbf{x}^{(k)} + \beta,$$

откуда получаем

$$\mathbf{x}^{(k+1)} = (E - B)^{-1}C\mathbf{x}^{(k)} + (E - B)^{-1}\beta.$$

Таким образом, метод Зейделя является методом простых итераций с матрицей правых частей $\alpha = (E - B)^{-1}C$ и вектором правых частей $(E - B)^{-1}\beta$. Следовательно, сходимость и погрешность метода Зейделя можно исследовать с помощью формул, выведенных для метода простых итераций, в которых вместо матрицы α подставлена матрица $(E - B)^{-1}C$, а вместо вектора правых частей — вектор $(E - B)^{-1}\beta$. Для практических вычислений важно, что в качестве достаточных условий сходимости метода Зейделя могут быть использованы условия, приведенные выше для метода простых итераций ($\|\alpha\| < 1$ или, если используется эквивалентная СЛАУ в форме (1.18), — диагональное преобладание матрицы A). При выполнении этих условий для оценки погрешности на k -й итерации можно использовать выражение

$$\varepsilon^{(k)} = \frac{\|C\|}{1 - \|\alpha\|} \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|.$$

Отметим, что, как и метод простых итераций, метод Зейделя может сходиться и при нарушении условия $\|\alpha\| < 1$. В этом случае $\varepsilon^{(k)} = \|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|$.

Пример 1.6. Методом Зейделя решить СЛАУ из примера 1.5 с той же погрешностью $\varepsilon = 0.01$.

Решение. Приведение СЛАУ к эквивалентному виду аналогично примеру (1.5). Диагональное преобразование элементов исходной матрицы СЛАУ гарантирует сходимость метода Зейделя.

Итерационный процесс выглядит следующим образом:

$$\begin{aligned} \mathbf{x}^{(0)} &= (1.2 \quad 1.3 \quad 1.4)^T; \\ \left\{ \begin{array}{l} x_1^{(1)} = -0.1 \cdot 1.3 - 0.1 \cdot 1.4 + 1.2 = 0.93; \\ x_2^{(1)} = -0.2 \cdot 0.93 - 0.1 \cdot 1.4 + 1.3 = 0.974; \\ x_3^{(1)} = -0.2 \cdot 0.93 - 0.2 \cdot 0.974 + 1.4 = 1.0192; \end{array} \right. \\ \left\{ \begin{array}{l} x_1^{(2)} = -0.1 \cdot 0.974 - 0.1 \cdot 1.0192 + 1.2 = 1.0007; \\ x_2^{(2)} = -0.2 \cdot 1.0007 - 0.1 \cdot 1.0192 + 1.3 = 0.998; \\ x_3^{(2)} = -0.2 \cdot 1.0007 - 0.2 \cdot 0.998 + 1.4 = 1.0003. \end{array} \right. \end{aligned}$$

Таким образом, уже на второй итерации погрешность $\|\mathbf{x}^{(2)} - \mathbf{x}^*\| < \varepsilon$, т.е. метод Зейделя в данном случае сходится быстрее метода простых итераций.

1.2. Численные методы решения задач на собственные значения и собственные векторы матриц

1.2.1. Основные определения и спектральные свойства матриц

Рассмотрим матрицу $A_{n \times n}$ в n -мерном вещественном пространстве R^n векторов $\mathbf{x} = (x_1 \ x_2 \ \dots \ x_n)^T$.

Собственным вектором \mathbf{x} матрицы A называется ненулевой вектор ($\mathbf{x} \neq \vartheta$), удовлетворяющий равенству

$$A\mathbf{x} = \lambda\mathbf{x}, \quad (1.21)$$

где λ — собственное значение матрицы A , соответствующее рассматриваемому собственному вектору.

Собственные значения матрицы A с действительными элементами могут быть вещественными различными, вещественными кратными, комплексными попарно сопряженными, комплексными кратными.

Классический способ нахождения собственных значений и собственных векторов заключается в следующем. Для однородной СЛАУ, полученной из (1.21),

$$(A - \lambda E)\mathbf{x} = \vartheta, \quad \vartheta = (0 \ 0 \ \dots \ 0)^T,$$

ненулевые решения $\mathbf{x} \neq \vartheta$ (а именно такие решения и находятся) имеют место при

$$\det(A - \lambda E) = 0, \quad (1.22)$$

причем уравнение (1.22) называют *характеристическим уравнением*, а выражение в левой части — *характеристическим многочленом*. Сначала каким-либо способом находят решения $\lambda_1, \lambda_2, \dots, \lambda_n$ алгебраического уравнения (1.22) n -й степени (предположим, что они вещественны и различны). Затем, решая однородную СЛАУ (1.22) для различных собственных значений $\lambda_j, j = 1, 2, \dots, n$,

$$(A - \lambda_j E) \mathbf{x}^j = \vartheta,$$

получают линейно независимые собственные векторы \mathbf{x}^j , соответствующие собственным значениям λ_j .

Попарно различным собственным значениям соответствуют линейно независимые собственные векторы; k -кратному корню характеристического уравнения (1.22), построенного для произвольной матрицы $A_{n \times n}$, соответствуют не более k линейно независимых собственных векторов. Если количество линейно независимых собственных векторов матрицы $A_{n \times n}$ совпадает с размерностью пространства R_n , то их можно принять за новый базис, в котором матрица $A_{n \times n}$ примет диагональный вид

$$\Lambda = U^{-1} \cdot A \cdot U. \quad (1.23)$$

Здесь на главной диагонали находятся собственные значения, а столбцы матрицы преобразования U являются собственными векторами матрицы A . Матрицы Λ и A , удовлетворяющие равенству (1.23), называются подобными. Собственные значения подобных матриц Λ и A совпадают.

Симметрическая матрица A , $A = A^T$ имеет полный спектр $\lambda_j, j = 1, 2, \dots, n$, вещественных собственных значений; положительно определенная симметрическая матрица $A = A^T, (Ax, x) > 0$, имеет полный спектр вещественных положительных собственных значений; k -кратному корню характеристического уравнения (1.22) симметрической матрицы соответствуют ровно k линейно независимых собственных векторов.

Симметрическая матрица имеет ровно n ортогональных собственных векторов. Приняв их за новый базис (т.е. построив матрицу преобразования U , в качестве столбцов которой взяты

координатные столбцы собственных векторов), можно преобразовать симметрическую матрицу A к *диагональному* виду с помощью преобразования (1.23). Для симметрической матрицы A матрица преобразования U в (1.23) является ортогональной, $U^{-1} = U^T$, следовательно, преобразование (1.23) имеет вид

$$\Lambda = U^T \cdot A \cdot U. \quad (1.24)$$

1.2.2. Метод вращений Якоби численного решения задач на собственные значения и собственные векторы матриц

Метод вращений Якоби применим только для симметрических матриц $A_{n \times n}$ ($A = A^T$) и решает полную проблему собственных значений и собственных векторов таких матриц. Он основан на отыскании с помощью итерационных процедур матрицы U в преобразовании подобия $\Lambda = U^{-1}AU$, а поскольку для симметрических матриц A матрица преобразования подобия U является ортогональной $U^{-1} = U^T$, имеем $\Lambda = U^T A U$, здесь Λ — диагональная матрица с собственными значениями на главной диагонали

$$\Lambda = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \dots & \ddots & \dots \\ 0 & \dots & \lambda_n \end{pmatrix}.$$

Пусть дана симметрическая матрица A . Требуется вычислить для нее с погрешностью ε все собственные значения и соответствующие им собственные векторы. Приведем алгоритм метода вращения. Пусть известна матрица $A^{(k)}$ на k -й итерации, при этом $k = 0$ для $A^{(0)} = A$.

Выбирается максимальный по модулю недиагональный элемент $a_{ij}^{(k)}$ матрицы $A^{(k)}$, $|a_{ij}^{(k)}| = \max |a_{lm}^{(k)}|$, $l < m$.

Ставится задача найти ортогональную матрицу $U^{(k)}$ такую, чтобы в результате преобразования подобия $A^{(k+1)} = U^{(k)} T A^{(k)} U^{(k)}$ произошло обнуление элемента $a_{ij}^{(k+1)}$ матрицы $A^{(k+1)}$. В качестве ортогональной матрицы выбирается матрица вращения, имеющая следующий вид:

$$U^k = \begin{pmatrix} & i & & j \\ 1 & \vdots & \vdots & \\ \ddots & \vdots & \vdots & 0 \\ & 1 & \vdots & \\ & \dots \dots \cos \varphi^{(k)} \dots -\sin \varphi^{(k)} \dots \dots & & i \\ & \vdots & 1 & \vdots \\ & \vdots & \ddots & \vdots \\ & \vdots & 1 & \vdots \\ & \dots \dots \sin \varphi^{(k)} \dots \cos \varphi^{(k)} \dots \dots & & j \\ & \vdots & \vdots & 1 \\ 0 & \vdots & \vdots & \ddots \\ & \vdots & \vdots & 1 \end{pmatrix}.$$

В матрице вращения на пересечении i -й строки и j -го столбца находится элемент $u_{ij}^{(k)} = \sin \varphi^{(k)}$, где $\varphi^{(k)}$ — угол вращения, подлежащий определению. Симметрично относительно главной диагонали (j -я строка, i -й столбец) расположен элемент $u_{ji}^{(k)} = \sin \varphi^{(k)}$.

Диагональные элементы $u_{ii}^{(k)}$ и $u_{jj}^{(k)}$ равны соответственно $u_{ii}^{(k)} = 1$, $u_{jj}^{(k)} = \cos \varphi^{(k)}$; другие диагональные элементы $u_{mm}^{(k)} = 1$, $m = 1, 2, \dots, n$, $m \neq i, m \neq j$. Остальные элементы в матрице вращения $U^{(k)}$ равны нулю.

Угол вращения $\varphi^{(k)}$ определяется из условия $a_{ij}^{(k+1)} = 0$:

$$\varphi^{(k)} = \frac{1}{2} \operatorname{arctg} \frac{2a_{ij}^{(k)}}{a_{ii}^{(k)} - a_{jj}^{(k)}},$$

причем если $a_{ii}^{(k)} = a_{jj}^{(k)}$, то $\varphi^{(k)} = \pi/4$.

- Строится матрица $A^{(k+1)}$

$$A^{(k+1)} = U^{(k)T} A^{(k)} U^{(k)},$$

в которой элемент $a_{ij}^{(k+1)} \approx 0$.

В качестве критерия окончания итерационного процесса используется условие малости суммы квадратов внедиагональных элементов:

$$t(A^{(k+1)}) = \left(\sum_{l,m; l < m} \left(a_{lm}^{(k+1)} \right)^2 \right)^{1/2}.$$

Если $t(A^{(k+1)}) > \varepsilon$, то итерационный процесс

$$A^{(k+1)} = U^{(k)T} A^{(k)} U^{(k)} = U^{(k)T} U^{(k-1)T} \dots U^{(0)T} A^{(0)} U^{(0)} U^{(1)} \dots U^{(k)}$$

продолжается. Если $t(A^{(k+1)}) < \varepsilon$, то итерационный процесс останавливается, и в качестве искомых собственных значений принимаются

$$\lambda_1 \approx a_{11}^{(k+1)}, \lambda_2 \approx a_{22}^{(k+1)}, \dots, \lambda_n \approx a_{nn}^{(k+1)}.$$

Координатными столбцами собственных векторов матрицы A в единичном базисе будут столбцы матрицы $U = U^{(0)T} A^{(0)} U^{(0)} U^{(1)} \dots U^{(k)}$, т.е.

$$\begin{aligned} (x^1)^T &= (u_{11}, u_{21}, \dots, u_{n1}), \quad (x^2)^T = (u_{12}, u_{22}, \dots, u_{n2}), \\ (x^n)^T &= (u_{1n}, u_{2n}, \dots, u_{nn}), \end{aligned}$$

причем эти собственные векторы будут ортогональны между собой, т.е.

$$(x^l, x^m) \approx 0, \quad l \neq m.$$

Пример 1.7. С погрешностью $\varepsilon = 0.3$ вычислить собственные значения и собственные векторы матрицы

$$A = \begin{bmatrix} 4 & 2 & 1 \\ 2 & 5 & 3 \\ 1 & 3 & 6 \end{bmatrix} \equiv A^{(0)}.$$

Решение. Используем приведенный алгоритм.

- Выбираем максимальный по модулю внедиагональный элемент матрицы $A^{(0)}$, т.е. находим $a_{ij}^{(0)}$ такой, что $|a_{ij}^{(0)}| = \max_{l,m} |a_{lm}^{(0)}|$, $l < m$. Им является элемент $a_{23}^{(0)} = 3$.

- Находим соответствующую этому элементу матрицу вращения:

$$U^{(0)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi^{(0)} & -\sin\varphi^{(0)} \\ 0 & \sin\varphi^{(0)} & \cos\varphi^{(0)} \end{bmatrix};$$

$$\varphi^{(0)} = \frac{1}{2} \operatorname{arctg} \frac{2 \cdot 3}{5 - 6} = 0.7033; \quad \sin\varphi^{(0)} = -0.65; \quad \cos\varphi^{(0)} = 0.76;$$

$$U^{(0)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.76 & 0.65 \\ 0 & -0.65 & 0.76 \end{bmatrix}.$$

- Вычисляем матрицу $A^{(1)}$:

$$A^{(1)} = U^{(0)T} A^{(0)} U^{(0)} = \begin{bmatrix} 4 & 0.87 & \mathbf{2.06} \\ 0.87 & 2.46 & -0.03 \\ 2.06 & -0.03 & 8.54 \end{bmatrix}.$$

В полученной матрице с точностью до ошибок округления элемент $a_{23}^{(0)} = 0$. Так как

$$t(A^{(1)}) = \left(\sum_{l,m;l < m} \left(a_{lm}^{(1)} \right)^2 \right)^{1/2} = \left(0.87^2 + 2.06^2 + (-0.03)^2 \right)^{1/2} > \epsilon,$$

итерационный процесс необходимо продолжить.

Переходим к следующей итерации ($k = 1$):

$$a_{13}^{(1)} = 2.06; \quad \left(\left| a_{13}^{(1)} \right| = \max_{l,m;l < m} \left| a_{lm}^{(1)} \right| \right);$$

$$U^{(1)} = \begin{bmatrix} \cos\varphi^{(1)} & 0 & -\sin\varphi^{(1)} \\ 0 & 1 & 0 \\ \sin\varphi^{(1)} & 0 & \cos\varphi^{(1)} \end{bmatrix};$$

$$\varphi^{(1)} = \frac{1}{2} \operatorname{arctg} \frac{2 \cdot 2.06}{4 - 8.54} = -0.3693; \quad \sin\varphi^{(1)} = -0.361; \quad \cos\varphi^{(1)} = 0.933;$$

$$U^{(1)} = \begin{bmatrix} 0.933 & 0 & 0.361 \\ 0 & 1 & 0 \\ 0 & 0 & 0.933 \end{bmatrix};$$

$$A^{(2)} = U^{(1)T} A^{(1)} U^{(1)} = \begin{bmatrix} 3.19 & \mathbf{0.819} & 0.005 \\ 0.819 & 2.46 & 0.28 \\ 0.005 & 0.28 & 9.38 \end{bmatrix};$$

$$t(A^{(2)}) = \left(\sum_{l,m;l < m} (a_{lm}^{(2)})^2 \right)^{1/2} = (0.819^2 + 0.28^2 + 0.005^2)^{1/2} > \varepsilon.$$

Переходим к следующей итерации ($k = 2$):

$$a_{12}^{(2)} = 0.819; \quad \left(\left| a_{12}^{(2)} \right| = \max_{l,m;l < m} \left| a_{lm}^{(2)} \right| \right);$$

$$\varphi^{(2)} = \frac{1}{2} \operatorname{arctg} \frac{2 \cdot 0.819}{3.19 - 2.46} = 0.5758; \quad \sin \varphi^{(2)} = 0.5445; \quad \cos \varphi^{(2)} = 0.8388;$$

$$U^{(2)} = \begin{bmatrix} 0.8388 & -0.5445 & 0 \\ 0.5445 & 0.8388 & 0 \\ 0 & 0 & 1 \end{bmatrix};$$

$$A^{(3)} = U^{(2)T} A^{(2)} U^{(2)} = \begin{bmatrix} 3.706 & 0.0003 & 0.1565 \\ 0.0003 & 1.929 & 0.232 \\ 0.1565 & 0.232 & 9.38 \end{bmatrix};$$

$$t(A^{(3)}) = (0.0003^2 + 0.1565^2 + 0.232^2)^{1/2} = 0.07839^{1/2} < \varepsilon.$$

Таким образом, в качестве искомых собственных значений могут быть приняты диагональные элементы матрицы $A^{(3)}$:

$$\lambda_1 \approx 3.706; \quad \lambda_2 \approx 1.929; \quad \lambda_3 \approx 9.38.$$

Собственные векторы определяются из произведения

$$U^{(0)} U^{(1)} U^{(2)} = \begin{bmatrix} 0.78 & -0.5064 & 0.361 \\ 0.2209 & 0.7625 & 0.6 \\ -0.58 & -0.398 & 0.7 \end{bmatrix};$$

$$\mathbf{x}^1 = \begin{bmatrix} 0.78 \\ 0.2209 \\ -0.58 \end{bmatrix}; \quad \mathbf{x}^2 = \begin{bmatrix} -0.5064 \\ 0.7625 \\ -0.398 \end{bmatrix}; \quad \mathbf{x}^3 = \begin{bmatrix} 0.361 \\ 0.6 \\ 0.7 \end{bmatrix}.$$

Полученные собственные векторы ортогональны в пределах заданной точности, т.е. $(\mathbf{x}^1, \mathbf{x}^2) = -0.00384; (\mathbf{x}^1, \mathbf{x}^3) = 0.0081; (\mathbf{x}^2, \mathbf{x}^3) = -0.0039$.

1.2.3. Частичная проблема собственных значений и собственных векторов матрицы.

Степенной метод

Рассмотренный метод вращения решает полную проблему собственных значений и собственных векторов матриц (симметрических) в том смысле, что определяются все собственные значения и собственные векторы.

Зачастую не нужно находить все собственные значения (спектр), а необходимо найти лишь максимальное по модулю из них (спектральный радиус). Существует степенной метод по определению спектрального радиуса матрицы, т.е. максимального собственного значения матрицы и соответствующего ему собственного вектора.

Пусть дана матрица A , и пусть ее собственные значения упорядочены по абсолютным величинам:

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|. \quad (1.25)$$

Тогда, выбрав некоторый вектор $\mathbf{y}^{(0)}$, например вектор, компоненты которого равны единице $\mathbf{y}^{(0)} = (1, 1, \dots, 1)^T$, можно для определения λ_1 построить следующий итерационный процесс:

$$\begin{aligned} \mathbf{y}^{(1)} &= A\mathbf{y}^{(0)}, \lambda_1^{(1)} = \frac{y_j^{(1)}}{y_j^{(0)}}, \\ \mathbf{y}^{(2)} &= A\mathbf{y}^{(1)}, \lambda_1^{(2)} = \frac{y_j^{(2)}}{y_j^{(1)}}, \end{aligned} \quad (1.26)$$

.....

$$\mathbf{y}^{(k)} = A\mathbf{y}^{(k-1)}, \lambda_1^{(k)} = \frac{y_j^{(k)}}{y_j^{(k-1)}},$$

где $y_j^{(k-1)}$, $y_j^{(k)}$ — соответствующие компоненты векторов $\mathbf{y}^{(k-1)}$, $\mathbf{y}^{(k)}$.

При этом в качестве номера j может использоваться любое число из диапазона $j = 1, 2, \dots, n$.

В связи с тем, что вектор $\mathbf{y}^{(k)}$ на k -й итерации может быть представлен в виде $\mathbf{y}^{(k)} = A\mathbf{y}^{(k-1)} = A\mathbf{y}^{(0)}$, рассматриваемый итера-

ционный процесс носит название *степенного метода*. При выполнении условий (1.25) итерационный процесс сходится к искомому собственному значению λ_1 и соответствующему собственному вектору, причем скорость сходимости определяется отношением $|\lambda_2|/\lambda_1|$ (чем оно меньше, тем выше скорость сходимости). В качестве критерия завершения вычислений используется условие $\varepsilon^{(k)} \left| \lambda_1^{(k)} - \lambda_1^{(k-1)} \right| \leq \varepsilon$, где ε — задаваемая вычислителем погрешность расчета.

Пример 1.8. Вычислить спектральный радиус матрицы

$$A = \begin{bmatrix} 5 & 1 & 2 \\ 1 & 4 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

с погрешностью $\varepsilon = 0.1$.

Решение. В качестве начального приближения собственного вектора возьмем $\mathbf{y}^{(0)} = (1, 1, 1)^T$. Реализуем итерационный процесс (1.26), полагая $j=1$:

$$\mathbf{y}^{(1)} = A\mathbf{y}^{(0)} = (8 \ 6 \ 6)^T,$$

$$\lambda_1^{(1)} = \frac{y_1^{(1)}}{y_1^{(0)}} = \frac{8}{1} = 8;$$

$$\mathbf{y}^{(2)} = A\mathbf{y}^{(1)} = (58 \ 38 \ 40)^T,$$

$$\lambda_1^{(2)} = \frac{y_1^{(2)}}{y_1^{(1)}} = \frac{58}{8} = 7.25;$$

$$\varepsilon^{(2)} = \left| \lambda_1^{(2)} - \lambda_1^{(1)} \right| = 0.75 > \varepsilon;$$

$$\mathbf{y}^{(3)} = A\mathbf{y}^{(2)} = (480 \ 250 \ 274)^T,$$

$$\lambda_1^{(3)} = \frac{y_1^{(3)}}{y_1^{(2)}} = \frac{408}{58} = 7.034;$$

$$\varepsilon^{(3)} = \left| \lambda_1^{(3)} - \lambda_1^{(2)} \right| = 0.216 > \varepsilon;$$

$$\mathbf{y}^{(4)} = A\mathbf{y}^{(3)} = (2838 \ 1682 \ 1888)^T,$$

$$\lambda_1^{(4)} = \frac{y_1^{(4)}}{y_1^{(3)}} = \frac{2838}{408} = 6.9559;$$

$$\varepsilon^{(4)} = \left| \lambda_1^{(4)} - \lambda_1^{(3)} \right| = 0.078 > \varepsilon.$$

Таким образом, полученное на четвертой итерации значение $\lambda_1^{(4)} = 6.9559$ удовлетворяет заданной погрешности и может быть взято в ка-

честве приближенного значения λ_1 . Искомое значение спектрального радиуса $\rho(A) = \max_i |\lambda_i| = |\lambda_1| = 6.9559$.

Рассмотренный пример наглядно иллюстрирует существенный недостаток алгоритма (1.26), связанный с возрастанием компонент итерируемого вектора $\mathbf{y}^{(k)}$ в ходе итерационного процесса.

Видно, что $\left| y_j^{(k)} / y_j^{(k-1)} \right| \approx |\lambda_1|$. Во избежание неограниченного возрастания (при $|\lambda_1| > 1$) или убывания (при $|\lambda_1| < 1$) компонент $\mathbf{y}^{(k)}$ по мере увеличения числа итераций k обычно при проведении компьютерных расчетов применяется степенной метод с нормировкой итерируемого вектора. С этой целью алгоритм (1.26) модифицируется следующим образом:

$$\mathbf{z}^{(k)} = A\mathbf{y}^{(k-1)}, \quad \lambda_1^{(k)} = \frac{z_j^{(k)}}{z_j^{(k-1)}}, \quad \mathbf{y}^{(k)} = \frac{\mathbf{z}^{(k)}}{\|\mathbf{z}^{(k)}\|}. \quad (1.27)$$

При этом в качестве начального приближения $\mathbf{y}^{(0)}$ берется вектор с единичной нормой.

Широко распространена также версия *степенного метода*, использующая скалярные произведения:

$$\mathbf{z}^{(k)} = A\mathbf{y}^{(k-1)}, \quad \mathbf{y}^{(k)} = \frac{\mathbf{z}^{(k)}}{\|\mathbf{z}^{(k)}\|}, \quad \lambda_1^{(k)} = (\mathbf{y}^{(k)}, A\mathbf{y}^{(k)}). \quad (1.28)$$

1.2.4. QR-алгоритм нахождения собственных значений матриц

При решении полной проблемы собственных значений для несимметричных матриц эффективным является подход, основанный на приведении матриц к подобным, имеющим треугольный или квазитреугольный вид. Одним из наиболее распространенных методов этого класса является *QR-алгоритм*, позволяющий находить как вещественные, так и комплексные собственные значения.

В основе *QR*-алгоритма лежит представление матрицы в виде $A = QR$, где Q — ортогональная матрица, $Q^{-1} = Q^T$, а R — верхняя треугольная.

Такое разложение существует для любой квадратной матрицы. Одним из возможных подходов к построению *QR*-разложения является использование *преобразования Хаусхолдера*, позволяю-

щего обратить в нуль группу поддиагональных элементов столбца матрицы.

Преобразование Хаусхолдера осуществляется с использованием *матрицы Хаусхолдера*, имеющей вид

$$H = E - \frac{2}{v^T v} vv^T, \quad (1.29)$$

где v — произвольный ненулевой вектор-столбец, E — единичная матрица, vv^T — квадратная матрица того же размера. Легко убедиться, что любая матрица такого вида является симметричной и ортогональной. При этом произвол в выборе вектора v дает возможность построить матрицу, отвечающую некоторым дополнительным требованиям.

Рассмотрим случай, когда необходимо обратить в нуль все элементы какого-либо вектора, кроме первого, т.е. построить матрицу Хаусхолдера такую, что

$$\tilde{\mathbf{b}} = H\mathbf{b}, \quad \mathbf{b} = (b_1, b_2, \dots, b_n)^T, \quad \tilde{\mathbf{b}} = (\tilde{b}_1, 0, \dots, 0)^T.$$

Тогда вектор v определится следующим образом:

$$v = \mathbf{b} + \text{sign}(b_1) \|\mathbf{b}\|_2 \mathbf{e}_1. \quad (1.30)$$

Здесь $\|\mathbf{b}\|_2 \left(\sum_i b_i^2 \right)^{1/2}$ — евклидова норма вектора, $\mathbf{e}_1 = (1, 0, \dots, 0)^T$.

Применяя описанную процедуру с целью обнуления поддиагональных элементов каждого из столбцов исходной матрицы, можно за фиксированное число шагов получить ее *QR*-разложение. Рассмотрим подробнее реализацию этого процесса.

Положим $A_0 = A$ и построим преобразование Хаусхолдера H_1 , переводящее матрицу A_0 в матрицу A_1 с нулевыми элементами первого столбца под главной диагональю $A_1 = H_1 A_0$:

$$A_0 = \begin{bmatrix} a_{11}^0 & a_{12}^0 & \cdots & a_{1n}^0 \\ a_{21}^0 & a_{22}^0 & \cdots & a_{2n}^0 \\ \dots & \dots & \dots & \dots \\ a_{n1}^0 & a_{n2}^0 & \cdots & a_{nn}^0 \end{bmatrix} \xrightarrow{H_1} A_1 = \begin{bmatrix} a_{11}^1 & a_{12}^1 & \cdots & a_{1n}^1 \\ 0 & a_{22}^1 & \cdots & a_{2n}^1 \\ \dots & \dots & \dots & \dots \\ 0 & a_{n2}^1 & \cdots & a_{nn}^1 \end{bmatrix}.$$

Ясно, что матрица Хаусхолдера H_1 должна определяться по первому столбцу матрицы A_0 , т.е. в качестве вектора \mathbf{b} в выра-

жении (1.30) берется вектор $(a_{11}^0, a_{21}^0, \dots, a_{n1}^0)^T$. Тогда компоненты вектора v вычисляются следующим образом:

$$v_1^1 = a_{11}^0 + \operatorname{sign}(a_{11}^0) \left(\sum_{j=1}^n (a_{j1}^0)^2 \right)^{1/2},$$

$$v_i^1 = a_{i1}^0, \quad i = 2, 3, \dots, n.$$

Матрица Хаусхолдера H_1 вычисляется согласно (1.29):

$$H_1 = E - 2 \frac{v^1 v^{1T}}{v^{1T} v^1}.$$

На следующем, втором шаге рассматриваемого процесса строится преобразование Хаусхолдера H_2 , обнуляющее элементы второго столбца матрицы A_1 , расположенные ниже главной диагонали: $A_2 = H_2 A_1$. Взяв в качестве вектора b вектор $(a_{22}^1, a_{32}^1, \dots, a_{n2}^1)^T$ размерности $(n - 1)$, получим следующие выражения для компонент вектора v :

$$v_1^2 = 0,$$

$$v_2^2 = a_{22}^1 + \operatorname{sign}(a_{22}^1) \left(\sum_{j=2}^n (a_{j2}^1)^2 \right)^{1/2},$$

$$v_i^2 = a_{i2}^1, \quad i = 3, 4, \dots, n.$$

Повторяя процесс $(n - 1)$ раз, получим искомое разложение $A = QR$, где

$$Q = (H_{n-1} H_{n-2} \dots H_1)^T = H_1 H_2 \dots H_{n-1}, \quad R = A_{n-1}.$$

Следует отметить определенное сходство рассматриваемого процесса с алгоритмом Гаусса. Отличие заключается в том, что здесь обнуление поддиагональных элементов соответствующего столбца осуществляется с использованием ортогонального преобразования.

Процедура QR -разложения многократно используется в QR -алгоритме вычисления собственных значений. Строится следующий итерационный процесс:

$$A^{(0)} = A,$$

$A^{(0)} = Q^{(0)}R^{(0)}$ — производится QR -разложение,

$A^{(1)} = R^{(0)}Q^{(0)}$ — производится перемножение матриц,

.....

$A^{(k)} = Q^{(k)}R^{(k)}$ — разложение,

$A^{(k+1)} = R^{(k+1)}Q^{(k)}$ — перемножение.

Таким образом, каждая итерация реализуется в два этапа. На первом этапе осуществляется разложение матрицы $A^{(k)}$ в произведение ортогональной $Q^{(k)}$ и верхней треугольной $R^{(k)}$ матриц, а на втором — полученные матрицы перемножаются в обратном порядке.

Нетрудно показать подобие матриц $A^{(k+1)}$ и $A^{(k)}$. Действительно, учитывая ортогональность $Q^{(k)}$, $Q^{(k)T}Q^{(k)} = E$, можно записать:

$$A^{(k+1)} = R^{(k)}Q^{(k)} = Q^{(k)T}Q^{(k)}R^{(k)}Q^{(k)} = Q^{(k)T}A^{(k)}Q^{(k)}.$$

Аналогично можно показать, что любая из матриц $A^{(k)}$ ортогонально подобна матрице A . При отсутствии у матрицы кратных собственных значений последовательность $A^{(k)}$ сходится к верхней треугольной матрице (в случае, когда все собственные значения вещественны) или к верхней квазитреугольной матрице (если имеются комплексно-сопряженные пары собственных значений).

Таким образом, каждому вещественному собственному значению будет соответствовать столбец со стремящимися к нулю поддиагональными элементами и в качестве критерия сходимости итерационного процесса для таких собственных значений можно

использовать неравенство $\left(\sum_{l=m+1}^n \left(a_{lm}^{(k)} \right)^2 \right)^{1/2} \leq \epsilon$. При этом соответ-

ствующее собственное значение принимается равным диагональному элементу данного столбца.

Каждой комплексно-сопряженной паре соответствует диагональный блок размерностью 2×2 , т.е. матрица $A^{(k)}$ имеет блочно-диагональную структуру. Принципиально то, что элементы этих блоков изменяются от итерации к итерации без видимой закономерности, в то время как комплексно-сопряженные собственные значения, определяемые каждым блоком, имеют тенденцию к сходимости. Это обстоятельство необходимо учитывать при формировании критерия выхода из итерационного процесса. Если в ходе итераций прослеживается комплексно-сопряженная пара собственных значений, соответствующая блоку, образуемому элементами j -го и $(j+1)$ -го столбцов $a_{jj}^{(k)}, a_{jj+1}^{(k)}, a_{j+1j}^{(k)}, a_{j+1j+1}^{(k)}$

, то, несмотря на значительное изменение в ходе итераций самих этих элементов, собственные значения, соответствующие данному блоку и определяемые из решения квадратного уравнения

$$\left(a_{jj}^{(k)} - \lambda^{(k)} \right) \left(a_{j+1,j+1}^{(k)} + \lambda^{(k)} \right) = a_{jj+1}^{(k)} + a_{j+1,j}^{(k)},$$

начиная с некоторого k отличаются незначительно. В качестве критерия окончания итераций для таких блоков может быть использовано условие $|\lambda^{(k)} - \lambda^{(k-1)}| \leq \epsilon$.

Замечание. Существенным недостатком рассмотренного алгоритма является большое число операций (оно пропорционально n^3 , где n – размерность матрицы), необходимое для QR-факторизации матрицы на каждой итерации. Эффективность QR-алгоритма может быть повышенна, если предварительно с помощью преобразования подобия привести матрицу к верхней Хессенберговой форме, в которой равны нулю все элементы, находящиеся ниже главной диагонали, за исключением элементов первой поддиагонали. Иными словами, предварительно производится операция

$$A^{(0)} = H^T A H,$$

где $A^{(0)}$ – матрица Хессенberга, имеющая следующую структуру (знак x обозначает ненулевые элементы):

$$\begin{pmatrix} x & x & x & \dots & x & x \\ x & x & x & \dots & x & x \\ 0 & x & x & \dots & x & x \\ 0 & 0 & x & \dots & x & x \\ \cdots & & & & x & x \\ 0 & 0 & 0 & 0 & x & x \end{pmatrix}.$$

Здесь принципиально то, что в дальнейшем, в ходе QR-итераций, матрицы $A^{(k)}$ сохраняют верхнюю Хессенбергову форму, что позволяет более экономно проводить их QR-разложение. Подробное изложение данного вопроса можно найти, например, в [2].

Пример 1.9. Используя преобразование Хаусхолдера, построить QR-разложение матрицы

$$A = \begin{bmatrix} 1 & 3 & 1 \\ 1 & 1 & 4 \\ 4 & 3 & 1 \end{bmatrix}.$$

Решение. Будем следовать предложенному алгоритму.

• Положим $A_0 = A$ и найдем ортогональную матрицу Хаусхолдера H_1 такую, что в матрице $A_1 = H_1 A_0$ все поддиагональные элементы первого столбца равны нулю. С этой целью компоненты вектора v определим, используя элементы первого столбца матрицы A_0 :

$$v_1^1 = a_{11}^0 + \operatorname{sign}(a_{11}^0) \left(\sum_{j=1}^3 (a_{j1}^0)^2 \right)^{1/2} = 1 + (1+1+4^2)^{1/2} = 5.24,$$

$$v_2^1 = a_{21}^0 = 1,$$

$$v_3^1 = a_{31}^0 = 4.$$

В результате получен вектор $v^1 = (5.24 \ 1 \ 4)^T$. Найдем соответствующую этому вектору матрицу Хаусхолдера:

$$H_1 = E - 2 \frac{v^1 v^{1T}}{v^{1T} v^1} = \begin{pmatrix} -0.24 & -0.24 & -0.94 \\ -0.24 & 0.96 & -0.18 \\ -0.94 & -0.18 & 0.28 \end{pmatrix}.$$

В заключение первого шага вычислим матрицу A_1 :

$$A_1 = H_1 A_0 = \begin{pmatrix} -4.24 & -3.77 & -2.12 \\ 0 & -0.29 & 3.40 \\ 0 & -2.17 & -1.38 \end{pmatrix}.$$

Таким образом, после первого шага получена матрица с нулевыми поддиагональными элементами в первом столбце.

• На втором шаге проделаем аналогичную процедуру, обнуляя поддиагональный элемент второго столбца.

$$v_1^2 = 0,$$

$$v_2^2 = a_{22}^1 + \operatorname{sign}(a_{22}^1) \left(\sum_{j=2}^3 (a_{j2}^1)^2 \right)^{1/2} = -0.29 - (0.29^2 + 2.17^2)^{1/2} = -2.48,$$

$$v_3^2 = a_{32}^1 = -2.17,$$

т.е. искомый вектор $v^2 = (0, -2.48, -2.17)^T$.

Далее найдем соответствующую ему матрицу Хаусхолдера:

$$H_2 = E - 2 \frac{v^2 v^{2T}}{v^{2T} v^2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -0.13 & -0.99 \\ 0 & -0.99 & 0.13 \end{pmatrix}$$

и вычислим матрицу A_2 :

$$A_2 = H_2 A_1 = \begin{pmatrix} -4.24 & -3.77 & -2.12 \\ 0 & 2.19 & 0.91 \\ 0 & 0 & -3.56 \end{pmatrix}.$$

Таким образом, исходная матрица A приведена к верхнему треугольному виду, т.е. получена матрица $R = A_2$ искомого разложения.

Результирующая ортогональная матрица преобразования Q получается в результате перемножения матриц H_i , $i = 1, 2$:

$$Q = H_1 H_2 = \begin{pmatrix} -0.24 & 0.97 & 0.11 \\ -0.24 & 0.05 & -0.97 \\ -0.94 & -0.25 & 0.22 \end{pmatrix}.$$

В заключение выпишем окончательный результат $A = QR$ в явном виде:

$$A = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 1 & 4 \\ 4 & 3 & 1 \end{pmatrix} = \begin{pmatrix} -0.24 & 0.97 & 0.11 \\ -0.24 & 0.05 & -0.97 \\ -0.94 & -0.25 & 0.22 \end{pmatrix} \cdot \begin{pmatrix} -4.24 & -3.77 & -2.12 \\ 0 & 2.19 & 0.91 \\ 0 & 0 & -3.56 \end{pmatrix}.$$

Пример 1.10. С помощью QR -алгоритма вычислить собственные значения матрицы A из примера 1.9 с погрешностью $\varepsilon = 0.01$.

Решение. Приступим к процедуре решения.

• Положим $A^{(0)} = A$ и найдем QR -разложение этой матрицы $A^{(0)} = Q^{(0)}R^{(0)}$. Эта процедура подробно рассмотрена в предыдущем примере. Получены следующие матрицы $Q^{(0)}, R^{(0)}$:

$$Q^{(0)} = \begin{pmatrix} -0.24 & 0.97 & 0.11 \\ -0.24 & 0.05 & -0.97 \\ -0.94 & -0.25 & 0.22 \end{pmatrix}, \quad R^{(0)} = \begin{pmatrix} -4.24 & -3.77 & -2.12 \\ 0 & 2.19 & 0.91 \\ 0 & 0 & -3.56 \end{pmatrix}.$$

Матрицу $A^{(1)}$ определим перемножением полученных в результате QR -разложения матриц в обратном порядке $A^{(1)} = R^{(0)} Q^{(0)}$:

$$A^{(1)} = \begin{pmatrix} 3.89 & -3.75 & 2.74 \\ -1.38 & -0.12 & -1.92 \\ 3.35 & 0.9 & = 0.77 \end{pmatrix}.$$

Первая итерация завершена. Поддиагональные элементы матрицы $A^{(1)}$ достаточно велики, поэтому итерационный процесс необходимо продолжить.

- Находим QR -разложение $A^{(1)} = Q^{(1)}R^{(1)}$, используя преобразование Хаусхолдера аналогично примеру 1.9:

$$Q^{(1)} = \begin{pmatrix} -0.73 & 0.68 & 0.05 \\ 0.26 & 0.21 & 0.94 \\ -0.63 & -0.7 & 0.33 \end{pmatrix}, \quad R^{(1)} = \begin{pmatrix} -5.32 & 2.14 & -2.02 \\ 0 & 3.21 & -2.0 \\ 0 & 0 & -1.93 \end{pmatrix}.$$

Перемножая полученные выше матрицы в обратном порядке, находим матрицу $A^{(2)}$:

$$A^{(2)} = \begin{pmatrix} 5.72 & 1.75 & 1.09 \\ 2.09 & -2.08 & 2.37 \\ 1.22 & -1.36 & -0.64 \end{pmatrix}.$$

Продолжая итерационный процесс, получим соответственно на 6-й и 7-й итерациях следующие матрицы:

$$A^{(6)} = \begin{pmatrix} 6.34 & 0.94 & -0.73 \\ 0.034 & -2.53 & 1.69 \\ 0.023 & -1.86 & -0.81 \end{pmatrix}, \quad A^{(7)} = \begin{pmatrix} 6.34 & 0.27 & 1.13 \\ -0.0014 & -2.01 & -2.58 \\ 0.0006 & 0.98 & -1.33 \end{pmatrix}.$$

Видно, что поддиагональные элементы первого столбца становятся достаточно малыми, и, следовательно, диагональный элемент $a_1^{(7)}$ может быть принят в качестве собственного значения. В то же время отчетливо прослеживается комплексно-сопряженная пара собственных значений, соответствующая блоку, образуемому элементами второго и третьего столбцов $a_{22}^{(k)}, a_{23}^{(k)}, a_{32}^{(k)}, a_{33}^{(k)}$. Несмотря на значительное изменение в ходе итераций самих этих элементов, собственные значения, соответствующие данному блоку и определяемые из решения квадратного уравнения $(a_{22}^{(k)} - \lambda^{(k)}) \left(a_{33}^{(k)} - \lambda^{(k)} \right) = a_{23}^{(k)} a_{32}^{(k)}$, меняются незначительно — в пределах допустимой погрешности. Таким образом, окончательное решение задачи можно записать в виде:

$$\lambda_1 = \lambda_1^{(7)} = 6.34, \quad \lambda_2 = \lambda_2^{(7)} = -1.67 + 1.55i, \quad \lambda_3 = \lambda_3^{(7)} = -1.67 - 1.55i.$$

Задачи

- 1.1.** Методом Гаусса решить системы линейных алгебраических уравнений (СЛАУ). Для матрицы СЛАУ вычислить определитель и обратную матрицу.

1.
$$\begin{cases} x_1 + 2x_2 - 2x_3 + 6x_4 = 24, \\ -3x_1 - 5x_2 + 14x_3 + 13x_4 = 41, \\ x_1 + 2x_2 - 2x_3 - 2x_4 = 0, \\ -2x_1 - 4x_2 + 5x_3 + 10x_4 = 20. \end{cases}$$
2.
$$\begin{cases} 2x_1 + 7x_2 - 8x_3 + 6x_4 = -39, \\ 4x_1 + 4x_2 - 7x_4 = 41, \\ -x_1 - 3x_2 + 6x_3 + 3x_4 = 4, \\ 9x_1 - 7x_2 - 2x_3 - 8x_4 = 113. \end{cases}$$
3.
$$\begin{cases} 9x_1 - 5x_2 - 6x_3 + 3x_4 = -8, \\ x_1 - 7x_2 + x_3 = 38, \\ 3x_1 - 4x_2 + 9x_3 = 47, \\ 6x_1 - x_2 + 9x_3 + 8x_4 = -8. \end{cases}$$
4.
$$\begin{cases} -x_1 - 7x_2 - 3x_3 - 2x_4 = -12, \\ -8x_1 + x_2 - 9x_3 = -60, \\ 8x_1 + 2x_2 - 5x_3 - 3x_4 = -91, \\ -5x_1 + 3x_2 + 5x_3 - 9x_4 = -43. \end{cases}$$
5.
$$\begin{cases} 3x_1 - 8x_2 + x_3 - 7x_4 = 96, \\ 6x_1 + 4x_2 + 8x_3 + 5x_4 = -13, \\ -x_1 + x_2 - 9x_3 - 3x_4 = -54, \\ -6x_1 + 6x_2 + 9x_3 - 4x_4 = 82. \end{cases}$$
6.
$$\begin{cases} x_1 + 2x_2 - x_3 - 7x_4 = -23, \\ 8x_1 - 9x_3 - 3x_4 = 39, \\ 2x_1 - 3x_2 + 7x_3 + x_4 = -7, \\ x_1 - 5x_2 - 6x_3 + 8x_4 = 30. \end{cases}$$
7.
$$\begin{cases} x_1 - 5x_2 - 7x_3 + x_4 = -75, \\ x_1 - 3x_2 - 9x_3 - 4x_4 = -41, \\ -2x_1 + 4x_2 + 2x_3 + x_4 = 18, \\ -9x_1 + 9x_2 + 5x_3 + 3x_4 = 29. \end{cases}$$
8.
$$\begin{cases} -4x_1 - 9x_2 + 4x_3 + 3x_4 = -51, \\ 2x_1 + 7x_2 + 9x_3 + 8x_4 = 76, \\ 4x_1 - 4x_2 + 2x_4 = 26, \\ -8x_1 + 5x_2 + 2x_3 + 9x_4 = -73. \end{cases}$$
9.
$$\begin{cases} -7x_1 - 9x_2 + x_3 - 9x_4 = 29, \\ -6x_1 - 8x_2 - 5x_3 + 2x_4 = 42, \\ -3x_1 + 6x_2 + 5x_3 - 9x_4 = 11, \\ -2x_1 - 5x_3 - 9x_4 = 75. \end{cases}$$
10.
$$\begin{cases} -7x_1 + 3x_2 - 4x_3 + 7x_4 = -126, \\ 8x_1 - x_2 - 7x_3 + 6x_4 = 29, \\ 9x_1 + 9x_2 + 3x_3 - 6x_4 = 27, \\ -7x_1 - 9x_2 - 8x_3 - 5x_4 = 34. \end{cases}$$
11.
$$\begin{cases} 2x_1 - 8x_2 + 5x_4 = -40, \\ -9x_1 + 9x_2 - 7x_3 + 6x_4 = -58, \\ -6x_1 + 7x_2 + 3x_3 + 8x_4 = -75, \\ -x_1 + 8x_2 + 5x_3 + x_4 = 1. \end{cases}$$
12.
$$\begin{cases} -x_1 - 8x_2 + 5x_4 = -60, \\ 6x_1 - 6x_2 + 2x_3 + 4x_4 = -10, \\ 9x_1 - 5x_2 - 6x_3 + 4x_4 = 65, \\ -5x_1 - 9x_3 + x_4 = 18. \end{cases}$$
13.
$$\begin{cases} -6x_1 - 5x_2 - 3x_3 - 8x_4 = 101, \\ 5x_1 - x_2 - 5x_3 - 4x_4 = 51, \\ -6x_1 + 5x_3 + 5x_4 = -53, \\ -7x_1 - 2x_2 + 8x_3 + 5x_4 = -63. \end{cases}$$
14.
$$\begin{cases} -x_1 - 3x_2 - 4x_3 = -3, \\ 3x_1 + 7x_2 - 8x_3 + 3x_4 = 30, \\ x_1 - 6x_2 + 2x_3 + 5x_4 = -90, \\ -8x_1 - 4x_2 - x_3 - x_4 = 12. \end{cases}$$
15.
$$\begin{cases} -9x_1 + 8x_2 + 8x_3 + 6x_4 = -81, \\ -7x_1 - 9x_2 + 5x_3 + 4x_4 = -50, \\ -3x_1 - x_2 + 8x_3 = -69, \\ 3x_1 - x_2 - 4x_3 - 5x_4 = 48. \end{cases}$$
16.
$$\begin{cases} -5x_1 - x_2 - 3x_3 - x_4 = 18, \\ -2x_1 + 8x_2 - 4x_4 = -12, \\ -7x_1 - 2x_2 + 2x_3 - 2x_4 = 6, \\ 2x_1 - 4x_2 - 4x_3 + 4x_4 = -12. \end{cases}$$

$$17. \begin{cases} 8x_1 + 8x_2 - 5x_3 - 8x_4 = 13, \\ 8x_1 - 5x_2 + 9x_3 - 8x_4 = 38, \\ 5x_1 - 4x_2 - 6x_3 - 2x_4 = 14, \\ 8x_1 + 3x_2 + 6x_3 + 6x_4 = -95. \end{cases}$$

$$18. \begin{cases} x_1 - x_2 + x_3 - 2x_4 = -20, \\ -9x_1 - x_2 + x_3 + 8x_4 = 60, \\ -7x_1 + 8x_2 - 6x_4 = -60, \\ 3x_1 - 5x_2 + x_3 - 6x_4 = -44. \end{cases}$$

$$19. \begin{cases} -8x_1 + 5x_2 + 8x_3 - 6x_4 = -144, \\ 2x_1 + 7x_2 - 8x_3 - x_4 = 25, \\ -5x_1 - 4x_2 + x_3 - 6x_4 = -21, \\ 5x_1 - 9x_2 - 2x_3 + 8x_4 = 103. \end{cases}$$

$$20. \begin{cases} 7x_1 + 8x_2 + 4x_3 - 6x_4 = -126, \\ -x_1 + 6x_2 - 2x_3 - 6x_4 = -42, \\ 2x_1 + 9x_2 + 6x_3 - 4x_4 = -115, \\ 5x_1 + 9x_2 + x_3 + x_4 = -67. \end{cases}$$

$$21. \begin{cases} -6x_1 - 8x_2 - 2x_3 - 8x_4 = -32, \\ 9x_1 + 8x_3 + 3x_4 = 8, \\ -9x_1 - 5x_3 + 9x_4 = -2, \\ -x_1 + 4x_2 - 8x_3 - 4x_4 = -36. \end{cases}$$

$$22. \begin{cases} 7x_1 - 5x_2 + 6x_3 + 7x_4 = 120, \\ 8x_1 - x_2 - 9x_3 + x_4 = 31, \\ -3x_1 + 8x_2 + 8x_3 + 8x_4 = 6, \\ 2x_1 - 3x_2 + 6x_3 - 4x_4 = 25. \end{cases}$$

$$23. \begin{cases} 2x_1 - 7x_2 + 8x_3 - 4x_4 = 57, \\ -x_1 + 4x_3 - x_4 = 24, \\ 3x_1 - 4x_2 + 2x_3 - x_4 = 28, \\ -9x_1 + x_2 - 4x_3 + 6x_4 = 12. \end{cases}$$

$$24. \begin{cases} -7x_1 - 2x_2 - x_3 - 4x_4 = -12, \\ -4x_1 + 6x_2 - 4x_4 = 22, \\ -8x_1 + 2x_2 - 9x_3 - 3x_4 = 51, \\ -7x_3 + x_4 = 49. \end{cases}$$

$$25. \begin{cases} -5x_1 - 6x_2 + 4x_3 - 2x_4 = 64, \\ 3x_2 - 4x_3 - 6x_4 = -55, \\ 2x_1 + 4x_2 - 4x_3 + 2x_4 = -48, \\ x_1 - 8x_2 + 2x_3 + 8x_4 = 68. \end{cases}$$

$$26. \begin{cases} -2x_1 - 9x_2 - 3x_3 + 7x_4 = -26, \\ -7x_1 + 8x_2 + 2x_3 + 5x_4 = -25, \\ -6x_1 + 2x_2 = -16, \\ -3x_2 + 8x_3 - 3x_4 = -5. \end{cases}$$

$$27. \begin{cases} -2x_1 - x_2 - 9x_3 - 5x_4 = 93, \\ -4x_1 + 4x_2 - 2x_3 + 6x_4 = 16, \\ 5x_1 + 7x_3 - 4x_4 = -80, \\ 9x_2 + 7x_3 + 7x_4 = -119. \end{cases}$$

$$28. \begin{cases} x_1 + 4x_2 - 9x_3 + 7x_4 = -67, \\ 2x_1 - 2x_2 - 2x_3 + 3x_4 = -57, \\ -x_1 + 3x_2 - 9x_3 - x_4 = -26, \\ -5x_1 + 2x_2 + 2x_3 + x_4 = 52. \end{cases}$$

$$29. \begin{cases} 9x_1 - 7x_2 - x_3 + x_4 = 55, \\ 2x_1 + 7x_2 + 3x_3 - 6x_4 = -66, \\ 4x_1 + 7x_2 - 3x_3 - 7x_4 = -43, \\ -9x_1 - 5x_2 - x_3 - 6x_4 = -24. \end{cases}$$

$$30. \begin{cases} -5x_1 + 2x_2 + 4x_3 - 4x_4 = 57, \\ 4x_1 + 9x_2 - 7x_3 - 5x_4 = -23, \\ -5x_1 + 6x_3 + 7x_4 = 23, \\ 4x_1 - 8x_2 - 6x_3 - 6x_4 = -20. \end{cases}$$

1.2. Методом прогонки решить СЛАУ.

1.
$$\begin{cases} -11x_1 - 9x_2 = -122, \\ 5x_1 - 15x_2 - 2x_3 = -48, \\ -8x_2 + 11x_3 - 3x_4 = 14, \\ 6x_3 - 15x_4 + 4x_5 = -50 \\ 3x_4 + 6x_5 = 42. \end{cases}$$
2.
$$\begin{cases} 10x_1 + 5x_2 = -120, \\ 3x_1 + 10x_2 - 2x_3 = -91, \\ 2x_2 - 9x_3 - 5x_4 = 5, \\ 5x_3 + 16x_4 - 4x_5 = -74, \\ -8x_4 + 16x_5 = -56. \end{cases}$$
3.
$$\begin{cases} 13x_1 - 5x_2 = -66, \\ -4x_1 + 9x_2 - 5x_3 = -47, \\ -x_2 - 12x_3 - 6x_4 = -43, \\ 6x_3 + 20x_4 - 5x_5 = -74, \\ 4x_4 + 5x_5 = 14. \end{cases}$$
4.
$$\begin{cases} -14x_1 - 6x_2 = -78, \\ -9x_1 + 15x_2 - x_3 = -73, \\ x_2 - 11x_3 + x_4 = -38, \\ -7x_3 + 12x_4 + 3x_5 = 77, \\ 6x_4 - 7x_5 = 91. \end{cases}$$
5.
$$\begin{cases} 8x_1 + 4x_2 = 48, \\ -5x_1 + 22x_2 + 8x_3 = 125, \\ -5x_2 - 11x_3 + x_4 = -43, \\ -9x_3 - 15x_4 + x_5 = 18, \\ x_4 + 7x_5 = -23. \end{cases}$$
6.
$$\begin{cases} 6x_1 - 5x_2 = -58, \\ -6x_1 + 16x_2 + 9x_3 = 161, \\ 9x_2 - 17x_3 - 3x_4 = -114, \\ 8x_3 + 22x_4 - 8x_5 = -90, \\ 6x_4 - 13x_5 = -55. \end{cases}$$
7.
$$\begin{cases} 15x_1 + 8x_2 = 92, \\ 2x_1 - 15x_2 + 4x_3 = -84, \\ 4x_2 + 11x_3 + 5x_4 = -77, \\ -3x_3 + 16x_4 - 7x_5 = 15, \\ 3x_4 + 8x_5 = -11. \end{cases}$$
8.
$$\begin{cases} -11x_1 - 8x_2 = 99, \\ 9x_1 - 17x_2 + x_3 = -75, \\ -4x_2 + 20x_3 + 9x_4 = 66, \\ -4x_3 - 14x_4 + 3x_5 = 54, \\ -6x_4 + 14x_5 = 8. \end{cases}$$
9.
$$\begin{cases} 8x_1 - 4x_2 = 32, \\ -2x_1 + 12x_2 - 7x_3 = 15, \\ 2x_2 - 9x_3 + x_4 = -10, \\ -8x_3 + 17x_4 - 4x_5 = 133, \\ -7x_4 + 13x_5 = -76. \end{cases}$$
10.
$$\begin{cases} -7x_1 - 6x_2 = -75, \\ 6x_1 + 12x_2 = 126, \\ -3x_2 + 5x_3 = 13, \\ -9x_3 + 21x_4 + 8x_5 = -40, \\ -5x_4 - 6x_5 = -24. \end{cases}$$
11.
$$\begin{cases} -10x_1 - 9x_2 = 7, \\ -5x_1 - 21x_2 - 8x_3 = 29, \\ 7x_2 + 12x_3 + 2x_4 = 31, \\ 8x_4 + 2x_5 = 56, \\ 2x_4 + 10x_5 = -24. \end{cases}$$
12.
$$\begin{cases} -11x_1 + 9x_2 = -114, \\ x_1 - 8x_2 + 3x_3 = 81, \\ -2x_2 - 11x_3 + 5x_4 = -8, \\ 3x_3 - 14x_4 + 7x_5 = -38, \\ 8x_4 + 10x_5 = 144. \end{cases}$$

$$13. \begin{cases} 14x_1 + 9x_2 = 125, \\ -8x_1 + 14x_2 + 6x_3 = -56, \\ -5x_2 - 17x_3 + 8x_4 = 144, \\ x_3 + 5x_4 - 2x_5 = 36, \\ -4x_4 - 10x_5 = 70. \end{cases}$$

$$15. \begin{cases} 16x_1 - 8x_2 = 0, \\ -7x_1 - 16x_2 + 5x_3 = -123, \\ 4x_2 + 12x_3 + 3x_4 = -68, \\ -4x_3 + 12x_4 - 7x_5 = 104, \\ -x_4 + 7x_5 = 20. \end{cases}$$

$$17. \begin{cases} -6x_1 + 5x_2 = 51, \\ -x_1 + 13x_2 + 6x_3 = 100, \\ -9x_2 - 15x_3 - 4x_4 = -12, \\ -x_3 - 7x_4 + x_5 = 47, \\ 9x_4 - 18x_5 = -90. \end{cases}$$

$$19. \begin{cases} 10x_1 - x_2 = 16, \\ -8x_1 + 16x_2 + x_3 = -110, \\ 6x_2 - 16x_3 + 6x_4 = 24, \\ -8x_3 + 16x_4 - 5x_5 = -3, \\ 5x_4 - 13x_5 = 87. \end{cases}$$

$$21. \begin{cases} 7x_1 - 2x_2 = 65, \\ -3x_1 - 7x_2 + 4x_3 = 23, \\ -2x_2 + 15x_3 + 5x_4 = 1, \\ -2x_3 - 12x_4 - 8x_5 = -58, \\ -3x_4 - 10x_5 = -8. \end{cases}$$

$$23. \begin{cases} 7x_1 - 5x_2 = 38, \\ -6x_1 + 19x_2 - 9x_3 = 14, \\ 6x_2 - 18x_3 + 7x_4 = -45, \\ -7x_3 - 11x_4 - 2x_5 = 30, \\ 5x_4 - 7x_5 = 48. \end{cases}$$

$$14. \begin{cases} -x_1 - x_2 = -4, \\ 7x_1 - 17x_2 - 8x_3 = 132, \\ -9x_2 + 19x_3 + 8x_4 = -59, \\ 7x_3 - 20x_4 + 4x_5 = -193, \\ -4x_4 + 12x_5 = -40. \end{cases}$$

$$16. \begin{cases} 18x_1 - 9x_2 = -81, \\ 2x_1 - 9x_2 - 4x_3 = 71, \\ -9x_2 + 21x_3 - 8x_4 = -39, \\ -4x_3 - 10x_4 + 5x_5 = 64, \\ 7x_4 + 12x_5 = 3. \end{cases}$$

$$18. \begin{cases} 8x_1 - 2x_2 = -14, \\ 7x_1 - 19x_2 + 9x_3 = -55, \\ -4x_2 + 21x_3 - 8x_4 = 49, \\ 7x_3 - 23x_4 + 9x_5 = 86, \\ 4x_4 - 7x_5 = 8. \end{cases}$$

$$20. \begin{cases} -6x_1 + 6x_2 = 30, \\ 2x_1 + 10x_2 - 7x_3 = -31, \\ -8x_2 + 18x_3 + 9x_4 = 108, \\ 6x_3 - 17x_4 - 6x_5 = -114, \\ 9x_4 + 14x_5 = 124. \end{cases}$$

$$22. \begin{cases} -14x_1 + 6x_2 = 82, \\ 2x_1 + 7x_2 = -51, \\ -7x_2 - 18x_3 - 9x_4 = -46, \\ 2x_3 - 13x_4 + 2x_5 = 111, \\ -7x_4 - 7x_5 = 35. \end{cases}$$

$$24. \begin{cases} -11x_1 + 9x_2 = -117, \\ -9x_1 + 17x_2 + 6x_3 + x_4 = -97, \\ 5x_2 + 20x_3 + 8x_4 = -6, \\ -6x_3 - 20x_4 + 7x_5 = 59, \\ 2x_4 + 8x_5 = -86. \end{cases}$$

25.
$$\begin{cases} 12x_1 - 5x_2 = 148, \\ -3x_1 - 18x_2 - 8x_3 = 45, \\ -2x_2 - 16x_3 - 9x_4 = -155, \\ -4x_3 + 18x_4 - 7x_5 = 11, \\ 4x_4 - 9x_5 = 3. \end{cases}$$
26.
$$\begin{cases} -12x_1 - 7x_2 = -102, \\ -7x_1 - 11x_2 - 3x_3 = -92, \\ -7x_2 + 21x_3 - 8x_4 = -65, \\ 4x_3 - 13x_4 + 5x_5 = 38, \\ -6x_4 + 14x_5 = -12. \end{cases}$$
27.
$$\begin{cases} -6x_1 + 3x_2 = -33, \\ 6x_1 - 23x_2 - 9x_3 = -107, \\ 2x_2 - 7x_3 - x_4 = 18, \\ 4x_3 + 15x_4 - 9x_5 = -69, \\ 5x_4 - 11x_5 = -31. \end{cases}$$
28.
$$\begin{cases} 16x_1 - 9x_2 = -27, \\ 8x_1 - 13x_2 - 5x_3 = -84, \\ -3x_2 - 21x_3 + 9x_4 = -225, \\ -9x_3 + 16x_4 - 5x_5 = -89, \\ x_4 - 9x_5 = 69. \end{cases}$$
29.
$$\begin{cases} -11x_1 + 9x_2 = -158, \\ -8x_2 - 6x_3 = 66, \\ 6x_2 + 15x_3 - 2x_4 = -45, \\ 4x_3 + 6x_4 - x_5 = 24, \\ -7x_4 - 10x_5 = -1. \end{cases}$$
30.
$$\begin{cases} 6x_1 + 3x_2 = 0, \\ -9x_1 - 17x_2 + 3x_3 = -99, \\ -3x_2 + 12x_3 - 7x_4 = -107, \\ 2x_3 - 9x_4 - 6x_5 = 5, \\ -4x_4 + 5x_5 = -6. \end{cases}$$

1.3. Методом простых итераций и методом Зейделя решить СЛАУ с погрешностью $\varepsilon = 0.01$.

1.
$$\begin{cases} 19x_1 - 4x_2 - 9x_3 - x_4 = 100, \\ -2x_1 + 20x_2 - 2x_3 - 7x_4 = -5, \\ 6x_1 - 5x_2 - 25x_3 + 9x_4 = 34, \\ -3x_2 - 9x_3 + 12x_4 = 69. \end{cases}$$
2.
$$\begin{cases} 24x_1 + 2x_2 + 4x_3 - 9x_4 = -9, \\ -6x_1 - 27x_2 - 8x_3 - 6x_4 = -76, \\ -4x_1 + 8x_2 + 19x_3 + 6x_4 = -79, \\ 4x_1 - 5x_2 - 3x_3 - 13x_4 = -70. \end{cases}$$
3.
$$\begin{cases} -23x_1 - 7x_2 + 5x_3 + 2x_4 = -26, \\ -7x_1 - 21x_2 + 4x_3 + 9x_4 = -55, \\ 9x_1 + 5x_2 - 31x_3 - 8x_4 = -58, \\ x_2 - 2x_3 + 10x_4 = -24. \end{cases}$$
4.
$$\begin{cases} 26x_1 - 9x_2 - 8x_3 + 8x_4 = 20, \\ 9x_1 - 21x_2 - 2x_3 + 8x_4 = -164, \\ -3x_1 + 2x_2 - 18x_3 + 8x_4 = 140, \\ x_1 - 6x_2 - x_3 + 11x_4 = -81. \end{cases}$$
5.
$$\begin{cases} 20x_1 + 5x_2 + 7x_3 + x_4 = -117, \\ -x_1 + 13x_2 - 7x_4 = -1, \\ 4x_1 - 6x_2 + 17x_3 + 5x_4 = 49, \\ -9x_1 + 8x_2 + 4x_3 - 25x_4 = -21. \end{cases}$$
6.
$$\begin{cases} 23x_1 - 6x_2 - 5x_3 + 9x_4 = 232, \\ 8x_1 + 22x_2 - 2x_3 + 5x_4 = -82, \\ 7x_1 - 6x_2 + 18x_3 - x_4 = 202, \\ 3x_1 + 5x_2 + 5x_3 - 19x_4 = -57. \end{cases}$$

7.
$$\begin{cases} 29x_1 + 8x_2 + 9x_3 - 9x_4 = 197, \\ -7x_1 - 25x_2 + 9x_4 = -226, \\ x_1 + 6x_2 + 16x_3 - 2x_4 = -95, \\ -7x_1 + 4x_2 - 2x_3 + 17x_4 = -58. \end{cases}$$

8.
$$\begin{cases} -7x_1 - x_2 + 2x_3 + 2x_4 = -24, \\ 3x_1 - 20x_2 - 8x_4 = -47, \\ -9x_1 + x_2 + 18x_3 - 6x_4 = 28, \\ -x_1 - x_3 - 6x_4 = -50. \end{cases}$$

9.
$$\begin{cases} 12x_1 - 3x_2 - x_3 + 3x_4 = -31, \\ 5x_1 + 20x_2 + 9x_3 + x_4 = 90, \\ 6x_1 - 3x_2 - 21x_3 - 7x_4 = 119, \\ 8x_1 - 7x_2 + 3x_3 - 27x_4 = 71. \end{cases}$$

10.
$$\begin{cases} 28x_1 + 9x_2 - 3x_3 - 7x_4 = -159, \\ -5x_1 + 21x_2 - 5x_3 - 3x_4 = 63, \\ -8x_1 + x_2 - 16x_3 + 5x_4 = -45, \\ -2x_2 + 5x_3 + 8x_4 = 24. \end{cases}$$

11.
$$\begin{cases} 21x_1 + x_2 - 8x_3 + 4x_4 = -119, \\ -9x_1 - 23x_2 - 2x_3 + 4x_4 = 79, \\ 7x_1 - x_2 - 17x_3 + 6x_4 = -24, \\ 8x_1 + 8x_2 - 4x_3 - 26x_4 = -52. \end{cases}$$

12.
$$\begin{cases} 14x_1 - 4x_2 - 2x_3 + 3x_4 = 38, \\ -3x_1 + 23x_2 - 6x_3 - 9x_4 = -195, \\ -7x_1 - 8x_2 + 21x_3 - 5x_4 = -27, \\ -2x_1 - 2x_2 + 8x_3 + 18x_4 = 142. \end{cases}$$

13.
$$\begin{cases} 24x_1 - 7x_2 - 4x_3 + 4x_4 = -190, \\ -3x_1 - 9x_2 - 2x_3 - 2x_4 = -12, \\ 3x_1 + 7x_2 + 24x_3 + 9x_4 = 155, \\ x_1 - 6x_2 - 2x_3 - 15x_4 = -17. \end{cases}$$

14.
$$\begin{cases} -22x_1 - 2x_2 - 6x_3 + 6x_4 = 96, \\ 3x_1 - 17x_2 - 3x_3 + 7x_4 = -26, \\ 2x_1 + 6x_2 - 17x_3 + 5x_4 = 35, \\ -x_1 - 8x_2 + 8x_3 + 23x_4 = -234. \end{cases}$$

15.
$$\begin{cases} -14x_1 + 6x_2 + x_3 - 5x_4 = 95, \\ -6x_1 + 27x_2 + 7x_3 - 6x_4 = -41, \\ 7x_1 - 5x_2 - 23x_3 - 8x_4 = 69, \\ 3x_1 - 8x_2 - 7x_3 + 26x_4 = 27. \end{cases}$$

16.
$$\begin{cases} 21x_1 - 6x_2 - 9x_3 - 4x_4 = 127, \\ -6x_1 + 20x_2 - 4x_3 + 2x_4 = -144, \\ -2x_1 - 7x_2 - 20x_3 + 3x_4 = 236, \\ 4x_1 + 9x_2 + 6x_3 + 24x_4 = -5. \end{cases}$$

17.
$$\begin{cases} -19x_1 + 2x_2 - x_3 - 8x_4 = 38, \\ 2x_1 + 14x_2 - 4x_4 = 20, \\ 6x_1 - 5x_2 - 20x_3 - 6x_4 = 52, \\ -6x_1 + 4x_2 - 2x_3 + 15x_4 = 43. \end{cases}$$

18.
$$\begin{cases} 18x_1 + 8x_2 - 3x_3 + 4x_4 = -84, \\ -7x_1 + 15x_2 - 5x_3 - 2x_4 = -5, \\ -4x_1 + 13x_2 + 4x_4 = -38, \\ -8x_1 - 8x_2 - 6x_3 + 31x_4 = 263. \end{cases}$$

19.
$$\begin{cases} 15x_1 + 7x_2 + 5x_4 = 176, \\ -3x_1 - 14x_2 - 6x_3 + x_4 = -111, \\ -2x_1 + 9x_2 + 13x_3 + 2x_4 = 74, \\ 4x_1 - x_2 + 3x_3 + 9x_4 = 76. \end{cases}$$

20.
$$\begin{cases} 10x_1 - x_2 - 2x_3 + 5x_4 = -99, \\ 4x_1 + 28x_2 + 7x_3 + 9x_4 = 0, \\ 6x_1 + 5x_2 - 23x_3 + 4x_4 = 67, \\ x_1 + 4x_2 + 5x_3 - 15x_4 = 58. \end{cases}$$

21.
$$\begin{cases} 24x_1 + 9x_2 - x_3 - 5x_4 = -24, \\ -x_1 - 14x_2 + x_3 + 9x_4 = 40, \\ -7x_1 + 5x_2 - 21x_3 = -84, \\ x_1 + 4x_2 + 8x_3 - 22x_4 = -56. \end{cases}$$
22.
$$\begin{cases} -18x_1 + 9x_2 - x_3 - 8x_4 = -60, \\ 6x_1 + 22x_2 + 9x_3 = -109, \\ -4x_1 + 2x_2 - 16x_3 + 9x_4 = -103, \\ x_1 + 6x_2 - x_3 - 14x_4 = -33. \end{cases}$$
23.
$$\begin{cases} -24x_1 - 6x_2 + 4x_3 + 7x_4 = 130, \\ -8x_1 + 21x_2 + 4x_3 - 2x_4 = 139, \\ 6x_1 + 6x_2 + 16x_3 = -84, \\ -7x_1 - 7x_2 + 5x_3 + 24x_4 = -165. \end{cases}$$
24.
$$\begin{cases} -25x_1 + 4x_2 - 4x_3 + 9x_4 = 86, \\ -9x_1 + 21x_2 + 5x_3 - 6x_4 = 29, \\ 9x_1 + 2x_2 + 19x_3 - 7x_4 = 28, \\ -7x_1 + 4x_2 - 7x_3 + 25x_4 = 68. \end{cases}$$
25.
$$\begin{cases} 15x_1 - 4x_2 - 6x_3 + 5x_4 = 104, \\ 4x_1 - 14x_2 - x_3 + 4x_4 = 70, \\ 7x_1 - 7x_2 + 27x_3 - 8x_4 = 170, \\ -3x_1 - 3x_2 + 2x_3 - 14x_4 = 48. \end{cases}$$
26.
$$\begin{cases} 18x_1 - 2x_3 + 7x_4 = 50, \\ -x_1 + 14x_2 - 3x_3 + 2x_4 = 2, \\ 5x_1 + 5x_2 + 26x_3 + 7x_4 = 273, \\ -2x_1 - 6x_2 + 9x_3 + 24x_4 = 111. \end{cases}$$
27.
$$\begin{cases} -26x_1 - 7x_2 - 8x_3 - 2x_4 = 51, \\ 2x_1 - 17x_2 - 6x_3 - 2x_4 = 85, \\ -7x_1 - 6x_2 - 23x_3 - 3x_4 = 71, \\ 3x_1 - 2x_2 - 7x_3 - 13x_4 = 91. \end{cases}$$
28.
$$\begin{cases} 10x_1 + 2x_3 + 4x_4 = 110, \\ 2x_1 + 16x_2 - 3x_3 + 8x_4 = 128, \\ x_1 + 5x_2 + 11x_3 - 4x_4 = 102, \\ 8x_1 + x_2 + 6x_3 - 17x_4 = 81. \end{cases}$$
29.
$$\begin{cases} 15x_1 + 3x_2 - 5x_3 - 5x_4 = 36, \\ 7x_1 - 15x_2 - 6x_3 + x_4 = -112, \\ -4x_1 + 7x_2 - 19x_3 - 6x_4 = 19, \\ 3x_1 - 5x_3 + 8x_4 = -23. \end{cases}$$
30.
$$\begin{cases} 22x_1 - 3x_2 - 8x_3 + 7x_4 = -158, \\ -8x_1 - 22x_2 - 4x_3 - 8x_4 = 254, \\ 8x_1 - 2x_2 - 18x_3 + 2x_4 = -108, \\ 7x_1 + 2x_2 - 9x_3 - 24x_4 = -24. \end{cases}$$

1.4. Используя метод вращений, найти собственные значения $\lambda_1, \lambda_2, \lambda_3$ и собственные векторы h_1, h_2, h_3 симметрических матриц с погрешностью $\varepsilon = 0.01$.

$$1. \begin{pmatrix} -7 & 4 & 5 \\ 4 & -6 & -9 \\ 5 & -9 & -8 \end{pmatrix}$$

$$2. \begin{pmatrix} -9 & 7 & 5 \\ 7 & 8 & 9 \\ 5 & 9 & 8 \end{pmatrix}$$

$$3. \begin{pmatrix} 5 & 5 & 3 \\ 5 & -4 & 1 \\ 3 & 1 & 2 \end{pmatrix}$$

$$4. \begin{pmatrix} 8 & 2 & -1 \\ 2 & -5 & -8 \\ -1 & -8 & -5 \end{pmatrix}$$

$$5. \begin{pmatrix} 0 & -7 & 7 \\ -7 & -9 & -5 \\ 7 & -5 & -1 \end{pmatrix}$$

$$6. \begin{pmatrix} 9 & 2 & -7 \\ 2 & -4 & -1 \\ -7 & -1 & 1 \end{pmatrix}$$

$$7. \begin{pmatrix} -6 & 6 & -8 \\ 6 & -4 & 9 \\ -8 & 9 & -2 \end{pmatrix}$$

$$8. \begin{pmatrix} 9 & -2 & 3 \\ -2 & 6 & 8 \\ 3 & 8 & -6 \end{pmatrix}$$

$$9. \begin{pmatrix} 4 & 7 & -1 \\ 7 & -9 & -6 \\ -1 & -6 & -4 \end{pmatrix}$$

$$10. \begin{pmatrix} -7 & -6 & 8 \\ -6 & 3 & -7 \\ 8 & -7 & 4 \end{pmatrix}$$

$$11. \begin{pmatrix} -3 & -2 & -4 \\ -2 & -3 & -7 \\ -4 & -7 & -3 \end{pmatrix}$$

$$12. \begin{pmatrix} 7 & 3 & -1 \\ 3 & -7 & -8 \\ -1 & -8 & -2 \end{pmatrix}$$

$$13. \begin{pmatrix} 8 & 0 & -2 \\ 0 & 5 & 4 \\ -2 & 4 & -6 \end{pmatrix}$$

$$14. \begin{pmatrix} -7 & -5 & -9 \\ -5 & 5 & 2 \\ -9 & 2 & 9 \end{pmatrix}$$

$$15. \begin{pmatrix} -3 & -1 & 3 \\ -1 & 8 & 1 \\ 3 & 1 & 5 \end{pmatrix}$$

$$16. \begin{pmatrix} 8 & -3 & 9 \\ -3 & 8 & -2 \\ 9 & -2 & -8 \end{pmatrix}$$

$$17. \begin{pmatrix} 5 & -3 & -4 \\ -3 & -3 & 4 \\ -4 & 4 & 0 \end{pmatrix}$$

$$18. \begin{pmatrix} 2 & 8 & 7 \\ 8 & 2 & 7 \\ 7 & 7 & -8 \end{pmatrix}$$

$$19. \begin{pmatrix} -8 & 9 & 6 \\ 9 & 9 & 1 \\ 6 & 1 & 8 \end{pmatrix}$$

$$20. \begin{pmatrix} -7 & -9 & 1 \\ -9 & 7 & 2 \\ 1 & 2 & 9 \end{pmatrix}$$

$$21. \begin{pmatrix} -4 & 1 & 7 \\ 1 & 8 & -5 \\ 7 & -5 & 1 \end{pmatrix}$$

$$22. \begin{pmatrix} 2 & -9 & 4 \\ -9 & 0 & 9 \\ 4 & 9 & 6 \end{pmatrix}$$

$$23. \begin{pmatrix} 9 & -5 & -6 \\ -5 & 1 & -8 \\ -6 & -8 & -3 \end{pmatrix}$$

$$24. \begin{pmatrix} -8 & -4 & 8 \\ -4 & -3 & 9 \\ 8 & 9 & -5 \end{pmatrix}$$

$$25. \begin{pmatrix} 5 & -4 & 7 \\ -4 & -3 & 4 \\ 7 & 4 & -1 \end{pmatrix}$$

$$26. \begin{pmatrix} -4 & 1 & -7 \\ 1 & 9 & 1 \\ -7 & 1 & 7 \end{pmatrix}$$

$$27. \begin{pmatrix} -8 & 5 & -7 \\ 5 & 1 & 4 \\ -7 & 4 & 4 \end{pmatrix}$$

$$28. \begin{pmatrix} 3 & 2 & 6 \\ 2 & -3 & -7 \\ 6 & -7 & 3 \end{pmatrix}$$

$$29. \begin{pmatrix} 2 & -1 & -8 \\ -1 & -3 & 4 \\ -8 & 4 & -3 \end{pmatrix}$$

$$30. \begin{pmatrix} -1 & 3 & 4 \\ 3 & 5 & -1 \\ 4 & -1 & -5 \end{pmatrix}$$

1.5. Используя степенной метод, оценить спектральный радиус матриц из задачи 1.4 с погрешностью $\varepsilon = 0.01$.

1.6. Используя QR-алгоритм, найти собственные значения заданных матриц с погрешностью $\varepsilon = 0.01$.

$$1. \begin{pmatrix} 3 & -7 & -1 \\ -9 & -8 & 7 \\ 5 & 2 & -2 \end{pmatrix}$$

$$2. \begin{pmatrix} -6 & -4 & 0 \\ -7 & 6 & -7 \\ -2 & -6 & -7 \end{pmatrix}$$

$$3. \begin{pmatrix} 5 & -5 & -6 \\ -1 & -8 & -5 \\ 2 & 7 & -3 \end{pmatrix}$$

$$4. \begin{pmatrix} -4 & -6 & -3 \\ -1 & 5 & -5 \\ 6 & 2 & 5 \end{pmatrix}$$

$$5. \begin{pmatrix} 5 & 8 & -2 \\ 7 & -2 & -4 \\ 5 & 8 & -1 \end{pmatrix}$$

$$6. \begin{pmatrix} 8 & -1 & -3 \\ -5 & 9 & -8 \\ 4 & -5 & 7 \end{pmatrix}$$

7.
$$\begin{pmatrix} 9 & 0 & 2 \\ -6 & 4 & 4 \\ -2 & -7 & 5 \end{pmatrix}$$

8.
$$\begin{pmatrix} -9 & 2 & 2 \\ -2 & 0 & 7 \\ 8 & 2 & 0 \end{pmatrix}$$

9.
$$\begin{pmatrix} -5 & -8 & 4 \\ 4 & 2 & 6 \\ -2 & 5 & -6 \end{pmatrix}$$

10.
$$\begin{pmatrix} -1 & 4 & -4 \\ 2 & -5 & 0 \\ -8 & -2 & 0 \end{pmatrix}$$

11.
$$\begin{pmatrix} 7 & 6 & -3 \\ 5 & -6 & 8 \\ -7 & 4 & -5 \end{pmatrix}$$

12.
$$\begin{pmatrix} 5 & -1 & -2 \\ -4 & 3 & -3 \\ -2 & -1 & 1 \end{pmatrix}$$

13.
$$\begin{pmatrix} -1 & 2 & 9 \\ 9 & 3 & 4 \\ 8 & -4 & -6 \end{pmatrix}$$

14.
$$\begin{pmatrix} 2 & -4 & 5 \\ -5 & -2 & -3 \\ 1 & -8 & -3 \end{pmatrix}$$

15.
$$\begin{pmatrix} 1 & 7 & -1 \\ -2 & 2 & -2 \\ 9 & -7 & 3 \end{pmatrix}$$

16.
$$\begin{pmatrix} 1 & 2 & 5 \\ -8 & 0 & -6 \\ 7 & -9 & -7 \end{pmatrix}$$

17.
$$\begin{pmatrix} -6 & 1 & -4 \\ -6 & 8 & -2 \\ 2 & -9 & 5 \end{pmatrix}$$

18.
$$\begin{pmatrix} -2 & 7 & -6 \\ -1 & 9 & =4 \\ -1 & 8 & -3 \end{pmatrix}$$

19.
$$\begin{pmatrix} 0 & -1 & 3 \\ -1 & 6 & -3 \\ -8 & 4 & 2 \end{pmatrix}$$

20.
$$\begin{pmatrix} 6 & 5 & -6 \\ -4 & -6 & 9 \\ -6 & 6 & 1 \end{pmatrix}$$

21.
$$\begin{pmatrix} 7 & 4 & 2 \\ -8 & -6 & 5 \\ 9 & 0 & -8 \end{pmatrix}$$

22.
$$\begin{pmatrix} -1 & 8 & 5 \\ 8 & -4 & 4 \\ 2 & 9 & -2 \end{pmatrix}$$

23.
$$\begin{pmatrix} 1 & 5 & -6 \\ 9 & -7 & -9 \\ 6 & -1 & -9 \end{pmatrix}$$

24.
$$\begin{pmatrix} -3 & 1 & -1 \\ 6 & 9 & -4 \\ 5 & -4 & -8 \end{pmatrix}$$

25.
$$\begin{pmatrix} -9 & 9 & -7 \\ -7 & 5 & -1 \\ -4 & 3 & 4 \end{pmatrix}$$

26.
$$\begin{pmatrix} -9 & -9 & -3 \\ -9 & 0 & -2 \\ -5 & -1 & -4 \end{pmatrix}$$

27.
$$\begin{pmatrix} -1 & 0 & 6 \\ 1 & -1 & 5 \\ 8 & 4 & -3 \end{pmatrix}$$

28.
$$\begin{pmatrix} -6 & -3 & 7 \\ 9 & 1 & -6 \\ 3 & -5 & 5 \end{pmatrix}$$

29.
$$\begin{pmatrix} -7 & 6 & 0 \\ 0 & 7 & 3 \\ 1 & 5 & -4 \end{pmatrix}$$

30.
$$\begin{pmatrix} 8 & -9 & -6 \\ -6 & 7 & -5 \\ -1 & -7 & 8 \end{pmatrix}$$

Ответы

1.1.

1. $(2, 4, 2, 3).$

2. $(8, -3, 2, -3).$

3. $(0, -5, 3, -5).$

4. $(-2, -4, 8, 9).$

5. $(-3, -6, 8, -7).$

6. $(6, 6, -1, 6).$

7. $(2, 4, 7, -8)$.
 9. $(-4, 2, -8, -3)$.
 11. $(7, 3, -2, -6)$.
 13. $(-2, -3, -6, -7)$.
 15. $(-1, 0, -9, -3)$.
 17. $(-4, -3, -1, -8)$.
 19. $(9, -6, -6, -1)$.
 21. $(-4, 2, 4, 4)$.
 23. $(1, -5, 7, 9)$.
 25. $(-2, -5, 7, 2)$.
 27. $(-9, -7, -7, -1)$.
 29. $(2, -4, -4, 5)$.

8. $(8, 4, 8, -5)$.
 10. $(8, -9, 2, -5)$.
 12. $(7, 6, -6, -1)$.
 14. $(-3, 6, -3, -9)$.
 16. $(-2, 3, -3, -2)$.
 18. $(-4, 0, -8, -4)$.
 20. $(-4, -5, -7, 5)$.
 22. $(6, -5, 3, 5)$.
 24. $(6, 3, -8, -7)$.
 26. $(3, 1, -1, -2)$.
 28. $(-5, 9, 7, -5)$.
 30. $(-3, 5, 8, 0)$.

1.2.

1. $(7, 5, 4, 6, 4)$.
 3. $(-7, -5, 6, -4, 6)$.
 5. $(3, 6, 1, -2, -3)$.
 7. $(4, 4, 8, -1, -1)$.
 9. $(6, 4, 3, 9, -1)$.
 11. $(2, -3, 3, 8, 4)$.
 13. $(7, 3, -7, 5, -9)$.
 15. $(2, 4, -9, 8, 4)$.
 17. $(-1, 9, -3, -6, -2)$.
 19. $(1, -6, -6, -6, -9)$.
 21. $(7, -8, -3, 6, -1)$.
 23. $(9, 5, 3, -3, -9)$.
 25. $(9, -8, 9, 3, 1)$.
 27. $(9, 7, 0, -4, 1)$.
 29. $(7, -9, 1, 3, -2)$.

2. $(-9, -6, 2, -7, -7)$.
 4. $(6, -1, 4, 7, -7)$.
 6. $(-8, 2, 9, -7, 1)$.
 8. $(-9, 0, 6, -6, -2)$.
 10. $(3, -9, 8, 0, 4)$.
 12. $(3, -9, 6, 8, 8)$.
 14. $(6, -2, -7, 7, -1)$.
 16. $(-8, -7, -6, 3, 2)$.
 18. $(-1, 3, 1, -5, -4)$.
 20. $(-5, 0, 3, 6, 5)$.
 22. $(8, -5, 8, -7, 2)$.
 24. $(9, -2, 3, -7, -9)$.
 26. $(5, 6, -3, -5, -3)$.
 28. $(0, 3, 9, -3, -8)$.
 30. $(-3, 6, -8, -1, -2)$.

1.3. Точные решения СЛАУ

1. $(8, 4, 3, 9)$. 2. $(4, 2, -7, 9)$.
 3. $(1, 2, 3, -2)$. 4. $(2, 8, -9, -4)$.
 5. $(-8, 2, 4, 5)$. 6. $(8, -7, 6, 4)$.
 7. $(7, 6, -9, -3)$. 8. $(7, 1, 7, 6)$.
 9. $(0, 7, -5, -5)$. 10. $(-6, 3, 6, 0)$.
 11. $(-6, -1, -1, 0)$. 12. $(-1, -6, -2, 8)$.
 13. $(-6, 2, 7, -1)$. 14. $(-5, -2, -6, -9)$.
 15. $(-8, -2, -5, 0)$. 16. $(1, -9, -8, 5)$.
 17. $(-2, 2, -4, 1)$. 18. $(-5, -4, -6, 5)$.
 19. $(9, 5, 3, 4)$. 20. $(-8, 4, -5, -5)$.
 21. $(0, 0, 4, 4)$. 22. $(0, -7, 5, -1)$.
 23. $(-9, 3, -3, -8)$. 24. $(-3, 0, 4, 3)$.
 25. $(8, -4, 2, 4)$. 26. $(3, 2, 9, 2)$.
 27. $(4, -3, -3, -4)$. 28. $(9, 7, 6, 2)$.
 29. $(1, 7, 2, -2)$. 30. $(-6, -9, 4, -3)$.

1.4.

$$1. \begin{cases} \lambda_1 = -3.71, \lambda_2 = 2.07, \lambda_3 = -19.36 \\ h_1 = \begin{pmatrix} 0.88 \\ 0.34 \\ 0.3 \end{pmatrix}, h_2 = \begin{pmatrix} -0.05 \\ 0.73 \\ -0.67 \end{pmatrix}, h_3 = \begin{pmatrix} -0.45 \\ 0.58 \\ 0.66 \end{pmatrix}. \end{cases}$$

$$2. \begin{cases} \lambda_1 = -11.69, \lambda_2 = 19.53, \lambda_3 = -0.83 \\ h_1 = \begin{pmatrix} 0.95 \\ -0.28 \\ -0.11 \end{pmatrix}, h_2 = \begin{pmatrix} 0.28 \\ 0.69 \\ 0.66 \end{pmatrix}, h_3 = \begin{pmatrix} -0.11 \\ -0.66 \\ 0.74 \end{pmatrix}. \end{cases}$$

$$3. \begin{cases} \lambda_1 = 8.7, \lambda_2 = -6.23, \lambda_3 = 0.53 \\ h_1 = \begin{pmatrix} 0.83 \\ 0.36 \\ 0.42 \end{pmatrix}, h_2 = \begin{pmatrix} -0.41 \\ 0.9 \\ 0.04 \end{pmatrix}, h_3 = \begin{pmatrix} -0.37 \\ -0.21 \\ 0.9 \end{pmatrix}. \end{cases}$$

$$4. \begin{cases} \lambda_1 = 8.79, \quad \lambda_2 = -13.02, \quad \lambda_3 = 2.22 \\ h_1 = \begin{pmatrix} 0.93 \\ 0.26 \\ -0.22 \end{pmatrix}, h_2 = \begin{pmatrix} -0.03 \\ 0.71 \\ 0.70 \end{pmatrix}, h_3 = \begin{pmatrix} 0.34 \\ -0.65 \\ 0.67 \end{pmatrix}. \end{cases}$$

$$5. \begin{cases} \lambda_1 = 10.30, \quad \lambda_2 = -12.96, \quad \lambda_3 = -7.33 \\ h_1 = \begin{pmatrix} 0.68 \\ -0.41 \\ 0.60 \end{pmatrix}, h_2 = \begin{pmatrix} 0.14 \\ 0.9 \\ 0.13 \end{pmatrix}, h_3 = \begin{pmatrix} -0.59 \\ 0.15 \\ 0.78 \end{pmatrix}. \end{cases}$$

$$6. \begin{cases} \lambda_1 = 13.34, \quad \lambda_2 = -4.30, \quad \lambda_3 = -3.04 \\ h_1 = \begin{pmatrix} 0.85 \\ 0.12 \\ -0.49 \end{pmatrix}, h_2 = \begin{pmatrix} -0.16 \\ 0.98 \\ -0.03 \end{pmatrix}, h_3 = \begin{pmatrix} 0.48 \\ 0.11 \\ 0.86 \end{pmatrix}. \end{cases}$$

$$7. \begin{cases} \lambda_1 = 0.77, \quad \lambda_2 = -19.34, \quad \lambda_3 = 6.57 \\ h_1 = \begin{pmatrix} 0.76 \\ 0.62 \\ -0.17 \end{pmatrix}, h_2 = \begin{pmatrix} -0.59 \\ 0.56 \\ -0.56 \end{pmatrix}, h_3 = \begin{pmatrix} -0.25 \\ 0.53 \\ 0.80 \end{pmatrix}. \end{cases}$$

$$8. \begin{cases} \lambda_1 = 9.35, \quad \lambda_2 = 10.29, \quad \lambda_3 = -10.65 \\ h_1 = \begin{pmatrix} 0.81 \\ 0.43 \\ 0.38 \end{pmatrix}, h_2 = \begin{pmatrix} -0.55 \\ 0.78 \\ 0.28 \end{pmatrix}, h_3 = \begin{pmatrix} -0.17 \\ -0.44 \\ 0.87 \end{pmatrix}. \end{cases}$$

$$9. \begin{cases} \lambda_1 = 8.07, \quad \lambda_2 = -14.62, \quad \lambda_3 = -2.44 \\ h_1 = \begin{pmatrix} 0.84 \\ 0.44 \\ -0.29 \end{pmatrix}, h_2 = \begin{pmatrix} -0.29 \\ 0.84 \\ 0.44 \end{pmatrix}, h_3 = \begin{pmatrix} 0.44 \\ -0.29 \\ 0.84 \end{pmatrix}. \end{cases}$$

$$10. \begin{cases} \lambda_1 = -11.61, \quad \lambda_2 = -3.41, \quad \lambda_3 = 15.02 \\ h_1 = \begin{pmatrix} 0.91 \\ 0.19 \\ -0.37 \end{pmatrix}, h_2 = \begin{pmatrix} 0.09 \\ 0.77 \\ 0.62 \end{pmatrix}, h_3 = \begin{pmatrix} 0.41 \\ -0.60 \\ 0.68 \end{pmatrix}. \end{cases}$$

11. $\begin{cases} \lambda_1 = -1.31, \quad \lambda_2 = -12.02, \quad \lambda_3 = 4.32 \\ h_1 = \begin{pmatrix} 0.87 \\ -0.46 \\ -0.13 \end{pmatrix}, h_2 = \begin{pmatrix} 0.43 \\ 0.61 \\ 0.66 \end{pmatrix}, h_3 = \begin{pmatrix} -0.22 \\ -0.63 \\ 0.73 \end{pmatrix}. \end{cases}$

12. $\begin{cases} \lambda_1 = 8.57, \quad \lambda_2 = -13.05, \quad \lambda_3 = 2.47 \\ h_1 = \begin{pmatrix} 0.87 \\ 0.34 \\ -0.34 \end{pmatrix}, h_2 = \begin{pmatrix} -0.09 \\ 0.81 \\ 0.57 \end{pmatrix}, h_3 = \begin{pmatrix} 0.47 \\ -0.47 \\ 0.74 \end{pmatrix}. \end{cases}$

13. $\begin{cases} \lambda_1 = 8.41, \quad \lambda_2 = 6.12, \quad \lambda_3 = -7.53 \\ h_1 = \begin{pmatrix} 0.95 \\ -0.22 \\ -0.19 \end{pmatrix}, h_2 = \begin{pmatrix} 0.27 \\ 0.92 \\ 0.25 \end{pmatrix}, h_3 = \begin{pmatrix} 0.12 \\ -0.30 \\ 0.94 \end{pmatrix}. \end{cases}$

14. $\begin{cases} \lambda_1 = -11.9, \quad \lambda_2 = 4.21, \quad \lambda_3 = 14.68 \\ h_1 = \begin{pmatrix} 0.90 \\ 0.22 \\ 0.36 \end{pmatrix}, h_2 = \begin{pmatrix} -0.04 \\ 0.89 \\ -0.44 \end{pmatrix}, h_3 = \begin{pmatrix} -0.42 \\ 0.38 \\ 0.81 \end{pmatrix}. \end{cases}$

15. $\begin{cases} \lambda_1 = -4.13, \quad \lambda_2 = 8.30, \quad \lambda_3 = 5.82 \\ h_1 = \begin{pmatrix} 0.94 \\ 0.10 \\ -0.32 \end{pmatrix}, h_2 = \begin{pmatrix} -0.01 \\ 0.95 \\ 0.28 \end{pmatrix}, h_3 = \begin{pmatrix} 0.33 \\ -0.26 \\ 0.90 \end{pmatrix}. \end{cases}$

16. $\begin{cases} \lambda_1 = 14.11, \quad \lambda_2 = 5.94, \quad \lambda_3 = -12.05 \\ h_1 = \begin{pmatrix} 0.78 \\ -0.5 \\ 0.36 \end{pmatrix}, h_2 = \begin{pmatrix} 0.47 \\ 0.86 \\ 0.18 \end{pmatrix}, h_3 = \begin{pmatrix} -0.4 \\ 0.03 \\ 0.91 \end{pmatrix}. \end{cases}$

17. $\begin{cases} \lambda_1 = 9.006, \quad \lambda_2 = -5.77, \quad \lambda_3 = -1.23 \\ h_1 = \begin{pmatrix} 0.78 \\ -0.36 \\ -0.51 \end{pmatrix}, h_2 = \begin{pmatrix} 0.02 \\ 0.82 \\ -0.55 \end{pmatrix}, h_3 = \begin{pmatrix} 0.62 \\ 0.42 \\ 0.65 \end{pmatrix}. \end{cases}$

$$18. \begin{cases} \lambda_1 = -6, \quad \lambda_2 = 14.37, \quad \lambda_3 = -12.37 \\ h_1 = \begin{pmatrix} 0.71 \\ -0.71 \\ 0 \end{pmatrix}, h_2 = \begin{pmatrix} 0.64 \\ 0.64 \\ 0.40 \end{pmatrix}, h_3 = \begin{pmatrix} -0.28 \\ -0.28 \\ 0.91 \end{pmatrix}. \end{cases}$$

$$19. \begin{cases} \lambda_1 = -13.14, \quad \lambda_2 = 14.75, \quad \lambda_3 = 7.38 \\ h_1 = \begin{pmatrix} 0.90 \\ -0.35 \\ -0.23 \end{pmatrix}, h_2 = \begin{pmatrix} 0.42 \\ 0.75 \\ 0.49 \end{pmatrix}, h_3 = \begin{pmatrix} 0.005 \\ -0.54 \\ 0.83 \end{pmatrix}. \end{cases}$$

$$20. \begin{cases} \lambda_1 = -11.55, \quad \lambda_2 = 12.03, \quad \lambda_3 = 8.51 \\ h_1 = \begin{pmatrix} 0.89 \\ 0.44 \\ -0.08 \end{pmatrix}, h_2 = \begin{pmatrix} -0.36 \\ 0.82 \\ 0.42 \end{pmatrix}, h_3 = \begin{pmatrix} 0.25 \\ -0.34 \\ 0.90 \end{pmatrix}. \end{cases}$$

$$21. \begin{cases} \lambda_1 = -9.73, \quad \lambda_2 = 11.07, \quad \lambda_3 = 3.66 \\ h_1 = \begin{pmatrix} 0.77 \\ -0.21 \\ -0.60 \end{pmatrix}, h_2 = \begin{pmatrix} -0.19 \\ 0.81 \\ -0.54 \end{pmatrix}, h_3 = \begin{pmatrix} 0.61 \\ 0.53 \\ 0.58 \end{pmatrix}. \end{cases}$$

$$22. \begin{cases} \lambda_1 = 7.7, \quad \lambda_2 = -12.82, \quad \lambda_3 = 13.12 \\ h_1 = \begin{pmatrix} 0.78 \\ -0.23 \\ 0.57 \end{pmatrix}, h_2 = \begin{pmatrix} 0.54 \\ 0.70 \\ -0.45 \end{pmatrix}, h_3 = \begin{pmatrix} -0.29 \\ 0.67 \\ 0.67 \end{pmatrix}. \end{cases}$$

$$23. \begin{cases} \lambda_1 = 11.89, \quad \lambda_2 = 7.23, \quad \lambda_3 = -12.13 \\ h_1 = \begin{pmatrix} 0.93 \\ -0.25 \\ -0.24 \end{pmatrix}, h_2 = \begin{pmatrix} 0.04 \\ 0.77 \\ -0.62 \end{pmatrix}, h_3 = \begin{pmatrix} 0.34 \\ 0.57 \\ 0.73 \end{pmatrix}. \end{cases}$$

$$24. \begin{cases} \lambda_1 = -2.02, \quad \lambda_2 = 5.63, \quad \lambda_3 = -19.6 \\ h_1 = \begin{pmatrix} 0.76 \\ -0.58 \\ 0.27 \end{pmatrix}, h_2 = \begin{pmatrix} 0.23 \\ 0.64 \\ 0.72 \end{pmatrix}, h_3 = \begin{pmatrix} -0.60 \\ -0.48 \\ 0.63 \end{pmatrix}. \end{cases}$$

25. $\begin{cases} \lambda_1 = 9.73, \quad \lambda_2 = 1.29, \quad \lambda_3 = -10.02 \\ h_1 = \begin{pmatrix} 0.85 \\ -0.11 \\ 0.51 \end{pmatrix}, h_2 = \begin{pmatrix} -0.25 \\ 0.77 \\ 0.57 \end{pmatrix}, h_3 = \begin{pmatrix} -0.46 \\ -0.62 \\ 0.63 \end{pmatrix}. \end{cases}$

26. $\begin{cases} \lambda_1 = -7.51, \quad \lambda_2 = 8.96, \quad \lambda_3 = 10.54 \\ h_1 = \begin{pmatrix} 0.89 \\ -0.08 \\ 0.43 \end{pmatrix}, h_2 = \begin{pmatrix} 0.202 \\ 0.94 \\ -0.24 \end{pmatrix}, h_3 = \begin{pmatrix} -0.39 \\ 0.30 \\ 0.86 \end{pmatrix}. \end{cases}$

27. $\begin{cases} \lambda_1 = -13.84, \quad \lambda_2 = 3.11, \quad \lambda_3 = 7.72 \\ h_1 = \begin{pmatrix} 0.82 \\ -0.38 \\ 0.41 \end{pmatrix}, h_2 = \begin{pmatrix} 0.47 \\ 0.86 \\ -0.13 \end{pmatrix}, h_3 = \begin{pmatrix} -0.30 \\ 0.31 \\ 0.90 \end{pmatrix}. \end{cases}$

28. $\begin{cases} \lambda_1 = 2.47, \quad \lambda_2 = -9.71, \quad \lambda_3 = 10.22 \\ h_1 = \begin{pmatrix} 0.74 \\ 0.61 \\ -0.26 \end{pmatrix}, h_2 = \begin{pmatrix} -0.38 \\ 0.71 \\ 0.57 \end{pmatrix}, h_3 = \begin{pmatrix} 0.54 \\ -0.32 \\ 0.77 \end{pmatrix}. \end{cases}$

29. $\begin{cases} \lambda_1 = 8.76, \quad \lambda_2 = -2.80, \quad \lambda_3 = -9.96 \\ h_1 = \begin{pmatrix} 0.75 \\ -0.26 \\ -0.60 \end{pmatrix}, h_2 = \begin{pmatrix} 0.43 \\ 0.88 \\ 0.15 \end{pmatrix}, h_3 = \begin{pmatrix} 0.49 \\ -0.37 \\ 0.78 \end{pmatrix}. \end{cases}$

30. $\begin{cases} \lambda_1 = 0.66, \quad \lambda_2 = 6.28, \quad \lambda_3 = -7.94 \\ h_1 = \begin{pmatrix} 0.72 \\ -0.36 \\ 0.57 \end{pmatrix}, h_2 = \begin{pmatrix} 0.41 \\ 0.91 \\ 0.06 \end{pmatrix}, h_3 = \begin{pmatrix} -0.55 \\ 0.19 \\ 0.81 \end{pmatrix}. \end{cases}$

1.6.

1. $\begin{cases} \lambda_1 = -13.496, \\ \lambda_2 = 5.248 + 2.864i, \\ \lambda_3 = 5.248 - 2.864i. \end{cases}$

2. $\begin{cases} \lambda_1 = 10.016, \\ \lambda_2 = -11.349, \\ \lambda_3 = -5.667. \end{cases}$

3.
$$\begin{cases} \lambda_1 = -5.416 + 5.831i, \\ \lambda_2 = -5.416 - 5.831i, \\ \lambda_3 = 4.832. \end{cases}$$

4.
$$\begin{cases} \lambda_1 = 7.112, \\ \lambda_2 = -0.556 + 3.821i, \\ \lambda_3 = -0.556 - 3.821i. \end{cases}$$

5.
$$\begin{cases} \lambda_1 = 4.76, \\ \lambda_2 = -5.351, \\ \lambda_3 = 2.591. \end{cases}$$

6.
$$\begin{cases} \lambda_1 = 13.403, \\ \lambda_2 = 8.779, \\ \lambda_3 = 1.819. \end{cases}$$

7.
$$\begin{cases} \lambda_1 = 10.027, \\ \lambda_2 = 3.986 + 6.096i, \\ \lambda_3 = 3.986 - 6.096i. \end{cases}$$

8.
$$\begin{cases} \lambda_1 = -9.071, \\ \lambda_2 = 5.071, \\ \lambda_3 = -4.999. \end{cases}$$

9.
$$\begin{cases} \lambda_1 = -6.224 + 4.709i, \\ \lambda_2 = -6.224 - 4.709i, \\ \lambda_3 = 3.448. \end{cases}$$

10.
$$\begin{cases} \lambda_1 = 5.665, \\ \lambda_2 = -7.548, \\ \lambda_3 = -4.117. \end{cases}$$

11.
$$\begin{cases} \lambda_1 = 8.907, \\ \lambda_2 = -13.982, \\ \lambda_3 = 1.076. \end{cases}$$

12.
$$\begin{cases} \lambda_1 = 6.384, \\ \lambda_2 = 3.839, \\ \lambda_3 = -1.224. \end{cases}$$

13.
$$\begin{cases} \lambda_1 = -13.006, \\ \lambda_2 = 4.503 + 2.803i, \\ \lambda_3 = 4.503 - 2.803i. \end{cases}$$

14.
$$\begin{cases} \lambda_1 = 7.813, \\ \lambda_2 = -5.406 + 1.501i, \\ \lambda_3 = -5.406 - 1.501i. \end{cases}$$

15.
$$\begin{cases} \lambda_1 = 4.134 + 4.655i, \\ \lambda_2 = 4.134 - 4.655i, \\ \lambda_3 = -2.269. \end{cases}$$

16.
$$\begin{cases} \lambda_1 = 7.211, \\ \lambda_2 = -11.932, \\ \lambda_3 = -1.278. \end{cases}$$

17.
$$\begin{cases} \lambda_1 = 8.321, \\ \lambda_2 = -6.268, \\ \lambda_3 = 4.946. \end{cases}$$

18.
$$\begin{cases} \lambda_1 = 4.854, \\ \lambda_2 = -1.854, \\ \lambda_3 = 1. \end{cases}$$

19.
$$\begin{cases} \lambda_1 = 2.311 + 5.103i, \\ \lambda_2 = 2.311 - 5.103i, \\ \lambda_3 = 4.414. \end{cases}$$

20.
$$\begin{cases} \lambda_1 = 9.861, \\ \lambda_2 = -13.276, \\ \lambda_3 = 4.414. \end{cases}$$

21.
$$\begin{cases} \lambda_1 = -6.834 + 2.908i, \\ \lambda_2 = -6.834 - 2.908i, \\ \lambda_3 = 6.669. \end{cases}$$

22.
$$\begin{cases} \lambda_1 = 9.637, \\ \lambda_2 = -10.524, \\ \lambda_3 = -6.113. \end{cases}$$

$$23. \begin{cases} \lambda_1 = -1.827, \\ \lambda_2 = -12.787, \\ \lambda_3 = -0.384. \end{cases}$$

$$24. \begin{cases} \lambda_1 = 10.313, \\ \lambda_2 = -7.861, \\ \lambda_3 = -4.452. \end{cases}$$

$$25. \begin{cases} \lambda_1 = 6.114, \\ \lambda_2 = -3.057 + 2.245i, \\ \lambda_3 = -3.057 - 2.245i. \end{cases}$$

$$26. \begin{cases} \lambda_1 = -15.996, \\ \lambda_2 = 5.536, \\ \lambda_3 = -2.54. \end{cases}$$

$$27. \begin{cases} \lambda_1 = 6.497, \\ \lambda_2 = -10.147, \\ \lambda_3 = -1.35. \end{cases}$$

$$28. \begin{cases} \lambda_1 = 7.443, \\ \lambda_2 = -7.388, \\ \lambda_3 = -0.054. \end{cases}$$

$$29. \begin{cases} \lambda_1 = 8.313, \\ \lambda_2 = -6.156 + 0.681i, \\ \lambda_3 = -6.156 - 0.681i. \end{cases}$$

$$30. \begin{cases} \lambda_1 = 14.697, \\ \lambda_2 = 11.784, \\ \lambda_3 = -3.481. \end{cases}$$

Глава 2

НЕЛИНЕЙНЫЕ УРАВНЕНИЯ И СИСТЕМЫ НЕЛИНЕЙНЫХ УРАВНЕНИЙ

При проведении научных и инженерных расчетов достаточно часто возникает необходимость решения скалярных нелинейных уравнений или систем нелинейных уравнений. В настоящем разделе рассматривается ряд численных методов, позволяющих решать подобные задачи.

2.1. Решение нелинейных уравнений

Численное решение нелинейных (алгебраических или трансцендентных) уравнений вида

$$f(x) = 0 \quad (2.1)$$

заключается в нахождении значений x , удовлетворяющих с заданной точностью данному уравнению, и состоит из следующих основных этапов:

- *отделение* (изоляция, локализация) корней уравнения;
- уточнение с помощью некоторого вычислительного алгоритма конкретного выделенного корня с заданной точностью.

Целью первого этапа является нахождение отрезков из области определения функции $f(x)$, внутри которых содержится только один корень решаемого уравнения. Иногда ограничиваются рассмотрением лишь какой-нибудь части области определения, вызывающей по тем или иным соображениям интерес. Для реализации данного этапа используются *графические* или *аналитические* способы.

При аналитическом способе отделения корней полезна следующая теорема [2].

Теорема 2.1. Непрерывная строго монотонная функция $f(x)$ имеет единственный нуль на отрезке $[a, b]$ тогда и только тогда, когда на его концах она принимает значения разных знаков.

Достаточным признаком монотонности функции $f(x)$ на отрезке $[a, b]$ является сохранение знака производной функции.

Графический способ отделения корней целесообразно использовать в случае, когда имеется возможность построения графика функции $y = f(x)$. Наличие графика исходной функции дает непосредственное представление о количестве и расположении нулей функции, что позволяет определить промежутки, внутри которых содержится только один корень. Если построение графика функции $y = f(x)$ вызывает затруднение, часто оказывается удобным преобразовать уравнение (2.1) к эквивалентному виду $f_1(x) = f_2(x)$ и построить графики функций $y = f_1(x)$ и $y = f_2(x)$. Абсциссы точек пересечения этих графиков будут соответствовать значениям корней решаемого уравнения.

Так или иначе, при завершении первого этапа должны быть определены промежутки, на каждом из которых содержится только один корень уравнения.

Для уточнения корня с требуемой погрешностью обычно применяется какой-либо итерационный метод, заключающийся в построении числовой последовательности $x^{(k)}$ $k = 0, 1, 2, \dots$, сходящейся к искомому корню $x^{(*)}$ уравнения (2.1).

Метод половинного деления. Процесс уточнения корня уравнения (2.1) на отрезке $[a, b]$ при условии, что функция $f(x)$ непрерывна на этом отрезке, заключается в следующем ([1], [3]).

Исходный отрезок делится пополам. Если $f((a + b)/2) = 0$, то корнем уравнения является $x^{(*)} = (a + b)/2$. Если $f((a + b)/2) \neq 0$, то выбирается та из половин $[a, (a + b)/2]$ или $[(a + b)/2, b]$, на концах которой функция $f(x)$ имеет противоположные знаки. Новый суженный отрезок $[a^{(1)}, b^{(1)}]$ снова делится пополам и проводится то же рассмотрение и т.д. В результате на каком-то этапе либо находится точный корень уравнения (2.1), либо имеется последовательность вложенных друг в друга отрезков $[a^{(1)}, b^{(1)}], [a^{(2)}, b^{(2)}], \dots, [a^{(k)}, b^{(k)}]$, для которых $f(a^{(k)})f(b^{(k)}) < 0$, $k = 0, 1, 2, \dots$.

Если требуется найти корень с погрешностью ε , то деление отрезка пополам продолжается до тех пор, пока длина отрезка не станет меньше 2ε . Тогда середина последнего отрезка даст значение корня с требуемой погрешностью.

Метод Ньютона (метод касательных). При нахождении корня уравнения (2.1) методом Ньютона итерационный процесс определяется формулой

$$x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})}, \quad k = 0, 1, 2, \dots . \quad (2.2)$$

Для начала вычислений требуется задание начального приближения $x^{(0)}$. Условия сходимости метода определяются следующей теоремой [2].

Теорема 2.2. Пусть на отрезке $[a, b]$ функция $f(x)$ имеет первую и вторую производные постоянного знака и пусть $f(a)f(b) < 0$. Тогда, если точка $x^{(0)}$ выбрана на $[a, b]$ так, что

$$f(x^{(0)})f'(x^{(0)}) > 0, \quad (2.3)$$

то начатая с нее последовательность $x^{(k)}, k = 0, 1, 2, \dots$, определяемая методом Ньютона (2.2), монотонно сходится к корню $x^{(*)} \in (a, b)$ уравнения (2.1).

В качестве условия окончания итераций в практических вычислениях часто используется правило $|x^{(k+1)} - x^{(k)}| < \varepsilon \Rightarrow x^{(*)} \approx x^{(k+1)}$.

Метод секущих. Использование метода Ньютона предполагает вычисление на каждой итерации значения функции и ее производной. Заменяя производную функции приближенным разностным отношением

$$f'(x^{(k)}) \approx \frac{f(x^{(k)}) - f(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}$$

и подставляя это отношение в (2.2), получаем итерационную формулу метода секущих

$$x^{(k-1)} = x^{(k)} - \frac{f(x^{(k)})(x^{(k)} - x^{(k-1)})}{f(x^{(k)}) - f(x^{(k-1)})}, \quad k = 1, 2, \dots . \quad (2.4)$$

Использование этого метода избавляет от необходимости расчета производной функции в процессе вычислений. Метод является двухшаговым; как видно из формулы (2.4), результат $(k+1)$ -го шага зависит от результатов (k) -го и $(k-1)$ -го шагов.

Для выполнения первой итерации требуется задание двух начальных точек $x^{(0)}$ и $x^{(1)}$. Выбор начальной точки $x^{(0)}$ осуществляется по тому же принципу, что и в методе касательных, например с использованием условия (2.3). Вторая начальная точка $x^{(1)}$

выбирается в непосредственной близости от $x^{(0)}$, желательно между точкой $x^{(0)}$ и искомым корнем.

Окончание счета по методу секущих, учитывая его быструю сходимость, можно контролировать путем проверки на малость или модуля невязки $|f(x^{(k)})|$, или модуля $|x^{(k+1)} - x^{(k)}|$ (см. [2]).

Метод простой итерации. При использовании метода простой итерации уравнение (2.1) заменяется эквивалентным уравнением с выделенным линейным членом

$$x = \varphi(x). \quad (2.5)$$

Решение ищется путем построения последовательности

$$x^{(k+1)} = \varphi(x^{(k)}), \quad k = 0, 1, 2, \dots, \quad (2.6)$$

начиная с некоторого заданного значения $x(0)$. Если $\varphi(x)$ — непрерывная функция, а $x^{(k)}, k = 0, 1, 2, \dots$, — сходящаяся последовательность, то значение $x^{(*)} = \lim_{k \rightarrow \infty} x^{(k)}$ является решением уравнения (2.5). Условия сходимости метода и оценка его погрешности определяются теоремой, доказанной в [2].

Теорема 2.3. Пусть функция $\varphi(x)$ определена и дифференцируема на отрезке $[a, b]$. Тогда, если выполняются условия

$$\begin{aligned} \varphi(x) &\in [a, b] \forall x \in [a, b], \\ \exists q : |\varphi'(x)| &\leq q < 1 \forall x \in (a, b), \end{aligned}$$

то уравнение (2.5) имеет единственный на $[a, b]$ корень $x^{(*)}$ к этому корню сходится определяемая методом простой итерации последовательность $x^{(k)}, k = 0, 1, 2, \dots$, начинающаяся с любого $x^{(0)} \in [a, b]$. При этом справедливы оценки погрешности ($\forall k \in N$):

$$\begin{aligned} |x^{(*)} - x^{(k+1)}| &\leq \frac{q}{1-q} |x^{(k+1)} - x^{(k)}|, \\ |x^{(*)} - x^{(k+1)}| &\leq \frac{q^{k+1}}{1-q} |x^{(1)} - x^{(0)}|. \end{aligned} \quad (2.7)$$

Пример 2.1. С погрешностью $\varepsilon = 10^{-3}$ решить уравнение

$$f(x) = e^{2x} + 3x - 4 = 0. \quad (2.8)$$

Решение. Для локализации корней применим графический способ. Преобразуем исходное уравнение к эквивалентному виду

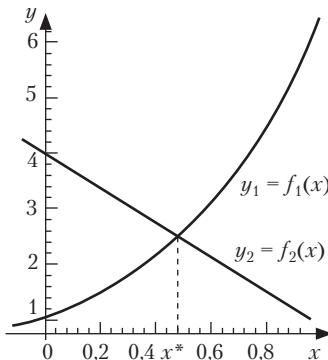


Рис. 2.1

$$e^{2x} = 4 - 3x.$$

Построив графики функций $f_1(x) = e^{2x}$ и $f_2(x) = 4 - 3x$ (рис. 2.1), определяем, что у решаемого уравнения имеется только один корень, который находится в интервале $0.4 < x^* < 0.6$.

Уточним значение корня с требуемой погрешностью, пользуясь приведенными методами.

Метод половинного деления. В качестве исходного отрезка выберем $[0.4, 0.6]$. Результаты дальнейших вычислений, согласно приведенному выше алгоритму, содержатся в таблице 2.1.

Таблица 2.1

k	$a^{(k)}$	$b^{(k)}$	$f(a^{(k)})$	$f(b^{(k)})$	$\frac{a^{(k)} + b^{(k)}}{2}$	$f\left(\frac{a^{(k)} + b^{(k)}}{2}\right)$
0	0.4000	0.6000	-0.5745	1.1201	0.5000	0.2183
1	0.4000	0.5000	-0.5745	0.2183	0.4500	-0.1904
2	0.4500	0.5000	-0.1904	0.2183	0.4750	0.0107
3	0.4500	0.4750	-0.1904	0.0107	0.4625	-0.0906
4	0.4625	0.4750	-0.0906	0.0107	0.4688	-0.0402
5	0.4688	0.4750	-0.0402	0.0107	0.4719	-0.0148
6	0.4719	0.4750	-0.0148	0.0107	0.4734	-0.0020
7	0.4734	0.4750	-0.0020	0.0107	[0.4742]	

Корень уравнения: ~ 0.474 .

Метод Ньютона. Для корректного использования данного метода необходимо в соответствии с теоремой 2.2 определить поведение первой и второй производных функции $f(x)$ на интервале уточнения корня и правильно выбрать начальное приближение $x^{(0)}$.

Для функции $f(x) = e^{2x} + 3x - 4$ имеем $f'(x) = 2e^{2x} + 3$, $f''(x) = 4e^{2x} —$ положительные во всей области определения функции. При решении

уравнения $f(x) = 0$ в качестве начального приближения можно выбрать правую границу интервала $x^{(0)} = 0.6$, для которой выполняется неравенство (2.3):

$$f(0.6)f''(0.6) > 0.$$

Дальнейшие вычисления проводятся по формуле (2.2), где

$$f(x^{(k)}) = e^{2x^{(k)}} + 3x^{(k)} = 4, \quad f'(x^{(k)}) = 2e^{2x^{(k)}} + 3.$$

Итерации завершаются при выполнении условия $|x^{(k+1)} - x^{(k)}| < \varepsilon$.

Результаты вычислений содержатся в таблице 2.2.

Таблица 2.2

k	$x(k)$	$f(x^{(k)})$	$f'(x^{(k)})$	$-f(x^{(k)})/f'(x^{(k)})$
0	0.6000	1.1201	9.6402	-0.1162
1	0.4838	0.0831	8.2633	-0.0101
2	0.4738	0.0005	8.1585	-0.0001
3	[0.4737]			

Корень уравнения: $x^{(*)} \approx 0.474$.

Метод секущих. В качестве начальных точек зададим: $x^{(0)} = 0.6$ и $x^{(1)} = 0.59$. Дальнейшие вычисления проводятся по формуле (2.4), где $f(x^{(k)}) = e^{2x^{(k)}} + 3x^{(k)} - 4$. Итерации завершаются при выполнении

условия $|x^{(k+1)} - x^{(k)}| < \varepsilon$. Результаты вычислений содержатся в таблице 2.3.

Таблица 2.3

k	$x(k)$	$f(x^{(k)})$
0	0.6000	1.1201
1	0.5900	1.0244
2	0.4830	0.0765
3	0.4744	0.0056
4	[0.4737]	

Корень уравнения: $x^{(*)} \approx 0.474$.

Метод простой итерации. Уравнение (2.8) можно записать в виде

$$x = (4 - e^{2x}) / 3 \tag{2.9}$$

или

$$x = (\ln(4 - 3x)) / 2. \tag{2.10}$$

Из двух этих вариантов приемлемым является вариант (2.10), так как, взяв в качестве основного интервал (0.4, 0.55) и положив $\varphi(x) = (\ln(4 - 3x))/2$, будем иметь:

$$\varphi(x) \in [0.4, 0.55] \forall x \in [0.4, 0.55], \quad \varphi'(x) = -\frac{3}{2(4-3x)}.$$

Таким образом, на интервале (0.4, 0.55) получим $|(\varphi'(x)| < 0.64 = q$. Условия теоремы 2.3 выполнены.

В качестве начального приближения положим $x^{(0)} = (0.4 + 0.55)/2 = 0.475$. Вычисляем последовательные приближения $x^{(k)}$ с одним запасным знаком по формуле (2.6), где $\varphi(x^{(k)}) = (\ln(4 - 3x^{(k)}))/2$. В соответствии с (2.7) достижение требуемой точности контролируется условием

$\left| \frac{q}{1-q} |x^{(k+1)} - x^{(k)}| \right| \leq \varepsilon$. Результаты вычислений приведены в таблице 2.4.

Таблица 2.4

k	$x^{(k)}$	$\varphi(x^{(k)})$
0	0.4750	0.4729
1	0.4729	0.4741
2	0.4741	0.4734
3	0.4734	0.4738
4	[0.4738]	

Корень уравнения $x^{(*)} \approx 0.474$.

Замечание. Если непосредственное преобразование уравнения (2.1) к виду (2.5) не позволяет получить уравнение, для которого выполняются условия сходимости метода простой итерации, можно преобразовать уравнение (2.1) к следующему эквивалентному уравнению:

$$x = x - \lambda f(x).$$

Данное уравнение имеет вид (2.5) с $\varphi(x) = x - \lambda f(x)$. Здесь λ — параметр, который подбирается таким образом, чтобы в нужной области выполнялось неравенство $|\varphi'(x)| = |1 - \lambda f'(x)| \leq q \leq 1$ (см. [2], [3]).

2.2. Решение систем нелинейных уравнений

Систему нелинейных уравнений с n неизвестными можно записать в виде

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0, \\ f_2(x_1, x_2, \dots, x_n) = 0, \\ \dots \\ f_n(x_1, x_2, \dots, x_n) = 0 \end{cases} \quad (2.11)$$

или, более коротко, в векторной форме

$$\mathbf{f}(\mathbf{x}) = \mathbf{0}, \quad (2.12)$$

где \mathbf{x} — вектор неизвестных величин, \mathbf{f} — вектор-функция

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \dots \\ f_n(x) \end{pmatrix}, \quad \mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix}.$$

В редких случаях для решения такой системы удается применить метод последовательного исключения неизвестных и свести решение исходной задачи к решению одного нелинейного уравнения с одним неизвестным. Значения других неизвестных величин находятся соответствующей подстановкой в конкретные выражения. Однако в подавляющем большинстве случаев для решения систем нелинейных уравнений используются итерационные методы. В дальнейшем предполагается, что ищется изолированное решение нелинейной системы.

Замечание. Как и в случае одного нелинейного уравнения, локализация решения может осуществляться на основе специфической информации по конкретной решаемой задаче (например, по физическим соображениям) и с помощью методов математического анализа. При решении системы двух уравнений часто удобным является графический способ, когда месторасположение корней определяется как точки пересечения кривых $f_1(x_1, x_2) = 0$, $f_2(x_1, x_2) = 0$ на плоскости (x_1, x_2) .

Метод Ньютона. Если определено начальное приближение $x^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)} \right)^T$, то итерационный процесс нахождения решения системы (2.11) методом Ньютона можно представить в виде

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} + \Delta x_1^{(k)}, \\ x_2^{(k+1)} = x_2^{(k)} + \Delta x_2^{(k)}, \\ \dots \\ x_n^{(k+1)} = x_n^{(k)} + \Delta x_n^{(k)}, \end{cases} \quad k = 0, 1, 2, \dots, \quad (2.13)$$

где значения приращений $\Delta x_1^{(k)}, \Delta x_2^{(k)}, \dots, \Delta x_n^{(k)}$ определяются из решения системы линейных алгебраических уравнений, все

коэффициенты которой выражаются через известное предыдущее приближение $\mathbf{x}^{(k)} = \left(\Delta x_1^{(k)}, \Delta x_2^{(k)}, \dots, \Delta x_n^{(k)}\right)$:

$$\begin{aligned} & f_1(\mathbf{x}^{(k)}) + \frac{\partial f_1(\mathbf{x}^{(k)})}{\partial x_1} \Delta x_1^{(k)} + \frac{\partial f_1(\mathbf{x}^{(k)})}{\partial x_2} \Delta x_2^{(k)} + \dots + \frac{\partial f_1(\mathbf{x}^{(k)})}{\partial x_n} \Delta x_n^{(k)} = 0, \\ & f_2(\mathbf{x}^{(k)}) + \frac{\partial f_2(\mathbf{x}^{(k)})}{\partial x_1} \Delta x_1^{(k)} + \frac{\partial f_2(\mathbf{x}^{(k)})}{\partial x_2} \Delta x_2^{(k)} + \dots + \frac{\partial f_2(\mathbf{x}^{(k)})}{\partial x_n} \Delta x_n^{(k)} = 0, \\ & \dots \\ & f_n(\mathbf{x}^{(k)}) + \frac{\partial f_n(\mathbf{x}^{(k)})}{\partial x_1} \Delta x_1^{(k)} + \frac{\partial f_n(\mathbf{x}^{(k)})}{\partial x_2} \Delta x_2^{(k)} + \dots + \frac{\partial f_n(\mathbf{x}^{(k)})}{\partial x_n} \Delta x_n^{(k)} = 0. \end{aligned} \quad (2.14)$$

В векторно-матричной форме расчетные формулы имеют вид

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \Delta \mathbf{x}^{(k)}, \quad k = 0, 1, 2, \dots, \quad (2.15)$$

где вектор приращений

$$\Delta \mathbf{x}^{(k)} = \begin{pmatrix} \Delta x_1^{(k)} \\ \Delta x_2^{(k)} \\ \dots \\ \Delta x_n^{(k)} \end{pmatrix}$$

находится из решения уравнения

$$\mathbf{f}(\mathbf{x}^{(k)}) + \mathbf{J}(\mathbf{x}^{(k)}) \Delta \mathbf{x}^{(k)} = \mathbf{0}. \quad (2.16)$$

Здесь

$$\mathbf{J}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \dots & \frac{\partial f_1(x)}{\partial x_n} \\ \frac{\partial f_2(x)}{\partial x_1} & \frac{\partial f_2(x)}{\partial x_2} & \dots & \frac{\partial f_2(x)}{\partial x_n} \\ \dots \\ \frac{\partial f_n(x)}{\partial x_1} & \frac{\partial f_n(x)}{\partial x_2} & \dots & \frac{\partial f_n(x)}{\partial x_n} \end{bmatrix}$$

— матрица Якоби первых производных вектора-функции $\mathbf{f}(\mathbf{x})$.

Выражая из (2.16) вектор приращений $\Delta\mathbf{x}^{(k)}$ и подставляя его в (2.15), итерационный процесс нахождения решения можно записать в виде

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \mathbf{J}^{-1}\left(\mathbf{x}^{(k)}\right)\mathbf{f}\left(\mathbf{x}^{(k)}\right), \quad k=0, 1, 2, \dots, \quad (2.17)$$

где $\mathbf{J}^{-1}\left(\mathbf{x}^{(k)}\right)$ — матрица, обратная матрице Якоби. Формула (2.17) есть обобщение формулы (2.2) на случай систем нелинейных уравнений.

При реализации алгоритма метода Ньютона в большинстве случаев предпочтительным является не вычисление обратной матрицы $\mathbf{J}^{-1}\left(\mathbf{x}^{(k)}\right)$, а нахождение из системы (2.14) значений при-

ращений $\Delta x_1^{(k)}, \Delta x_2^{(k)}, \dots, \Delta x_n^{(k)}$ и вычисление нового приближения по (2.13). Для решения таких линейных систем можно привлекать самые разные методы, как прямые, так и итерационные (см. раздел 1.1), с учетом размерности n решаемой задачи и специфики матриц Якоби $\mathbf{J}(\mathbf{x})$ (например, симметрии, разреженности и т.п.).

Использование метода Ньютона предполагает дифференцируемость функций $f_1(x), f_2(x), \dots, f_n(x)$ и невырожденность матрицы Якоби $\left(\det \mathbf{J}(\mathbf{x}^{(k)}) \neq 0\right)$. В случае, если начальное приближение выбрано в достаточно малой окрестности искомого корня, итерации сходятся к точному решению, причем сходимость является квадратичной.

В практических вычислениях в качестве условия окончания итераций обычно используется критерий (см. [2], [5]).

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| \leq \varepsilon, \quad (2.18)$$

где ε — заданная погрешность.

Пример 2.2. Методом Ньютона найти положительное решение системы нелинейных уравнений

$$\begin{cases} f_1(x_1, x_2) = 0.1x_1^2 + x_1 + 0.2x_2^2 - 0.3 = 0, \\ f_2(x_1, x_2) = 0.2x_1^2 + x_2 - 0.1x_1x_2 - 0.7 = 0 \end{cases} \quad (2.19)$$

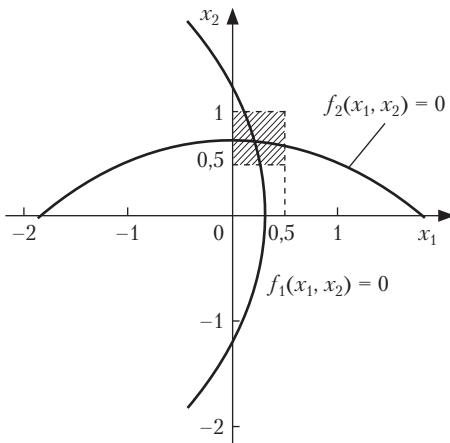


Рис. 2.2

с погрешностью $\varepsilon = 10^{-4}$.

Решение. Для выбора начального приближения применяем графический способ. Построив на плоскости (x_1, x_2) в интересующей нас области кривые $f_1(x_1, x_2) = 0$ и $f_2(x_1, x_2) = 0$ (рис. 2.2), определяем, что положительное решение системы уравнений находится в квадрате $0 < x_1 < 0.5$, $0.5 < x_2 < 1.0$.

В качестве начального приближения примем $x_1^{(0)} = 0.25$, $x_2^{(0)} = 0.75$.

Для системы двух уравнений расчетные формулы (2.13), (2.14) удобно записать в виде, разрешенном относительно $x_1^{(k+1)}$, $x_2^{(k+1)}$,

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} - \frac{\det A_1^{(k)}}{\det J^{(k)}}, \\ x_2^{(k+1)} = x_2^{(k)} - \frac{\det A_2^{(k)}}{\det J^{(k)}}, \end{cases} \quad k = 0, 1, 2, \dots, \quad (2.20)$$

где

$$J^{(k)} = \begin{bmatrix} \frac{\partial f_1(x_1^{(k)}, x_2^{(k)})}{\partial x_1} & \frac{\partial f_1(x_1^{(k)}, x_2^{(k)})}{\partial x_2} \\ \frac{\partial f_2(x_1^{(k)}, x_2^{(k)})}{\partial x_1} & \frac{\partial f_2(x_1^{(k)}, x_2^{(k)})}{\partial x_2} \end{bmatrix},$$

$$A_1^{(k)} = \begin{bmatrix} f_1(x_1^{(k)}, x_2^{(k)}) & \frac{\partial f_1(x_1^{(k)}, x_2^{(k)})}{\partial x_2} \\ f_2(x_1^{(k)}, x_2^{(k)}) & \frac{\partial f_2(x_1^{(k)}, x_2^{(k)})}{\partial x_2} \end{bmatrix},$$

$$A_2^{(k)} = \begin{bmatrix} \frac{\partial f_1(x_1^{(k)}, x_2^{(k)})}{\partial x_1} & f_1(x_1^{(k)}, x_2^{(k)}) \\ \frac{\partial f_2(x_1^{(k)}, x_2^{(k)})}{\partial x_1} & f_2(x_1^{(k)}, x_2^{(k)}) \end{bmatrix}.$$

В рассматриваемом примере имеем

$$f_1(x_1^{(k)}, x_2^{(k)}) = 0.1x_1^{(k)2} + x_1^{(k)} + 0.2x_2^{(k)2} - 0.3,$$

$$f_2(x_1^{(k)}, x_2^{(k)}) = 0.2x_1^{(k)2} + x_2^{(k)} - 0.1x_1^{(k)}x_2^{(k)} - 0.7,$$

$$\frac{\partial f_1(x_1^{(k)}, x_2^{(k)})}{\partial x_1} = 0.2x_1^{(k)} + 1,$$

$$\frac{\partial f_1(x_1^{(k)}, x_2^{(k)})}{\partial x_2} = 0.4x_2^{(k)},$$

$$\frac{\partial f_2(x_1^{(k)}, x_2^{(k)})}{\partial x_1} = 0.4x_1^{(k)} - 0.1x_2^{(k)},$$

$$\frac{\partial f_2(x_1^{(k)}, x_2^{(k)})}{\partial x_2} = 1 - 0.1x_1^{(k)}.$$

Подставляя в правые части соотношений (2.20) выбранные значения $x_1^{(0)}, x_2^{(0)}$, получим приближение $(x_1^{(1)}, x_2^{(1)})$, используемое, в свою очередь, для нахождения $(x_1^{(2)}, x_2^{(2)})$. Итерации продолжаются до выполнения условия (2.18), в котором

$$\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\| = \max_i |x_i^{(k+1)} - x_i^{(k)}|.$$

Результаты вычислений содержатся в таблице 2.5.

Таблица 2.5

k	$x_1^{(k)}$	$f_1(x_1^{(k)}, x_2^{(k)})$	$\frac{\partial f_1(x_1^{(k)}, x_2^{(k)})}{\partial x_1}$	$\frac{\partial f_1(x_1^{(k)}, x_2^{(k)})}{\partial x_2}$	$\det A_1^{(k)}$	$\det A_2^{(k)}$	$\det J^{(k)}$
0	0.25000 0.75000	0.06875 0.04375	1.012361 0.02500	0.300002 0.97500	0.05391	0.04258	0.97969
1	0.29498 0.70654	-0.00138 0.00037	1.00760 0.00734	0.28262 0.98050	-0.00146	0.00038	0.98588
2	0.19646 0.70615	0.00005 0.00000	1.00772 0.00797	0.28246 0.98035	0.00005	0.00000	0.98567
3	0.19641 0.70615						

Корни системы уравнений: $x_1^{(*)} \approx 0.1964, x_2^{(*)} \approx 0.7062$.

Метод простой итерации. При использовании метода простой итерации система уравнений (2.11) приводится к эквивалентной системе специального вида

$$\begin{cases} x_1 = \varphi_1(x_1, x_2, \dots, x_n), \\ x_2 = \varphi_2(x_1, x_2, \dots, x_n), \\ \dots \\ x_n = \varphi_n(x_1, x_2, \dots, x_n), \end{cases} \quad (2.21)$$

или, в векторной форме,

$$\mathbf{x} = \varphi(\mathbf{x}), \quad \varphi(\mathbf{x}) = \begin{pmatrix} \varphi_1(\mathbf{x}) \\ \varphi_2(\mathbf{x}) \\ \dots \\ \varphi_n(\mathbf{x}) \end{pmatrix}, \quad (2.22)$$

где функции $\varphi_1(x), \dots, \varphi_n(x)$ определены и непрерывны в некоторой окрестности искомого изолированного решения $\mathbf{x}^{(*)} = (x_1^{(*)}, \dots, x_n^{(*)})^T$.

Если выбрано начальное приближение $\mathbf{x}^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}\right)^T$,

последующие приближения в методе простой итерации находятся по формулам

$$\begin{cases} x_1^{(k+1)} = \varphi_1\left(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}\right), \\ x_2^{(k+1)} = \varphi_2\left(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}\right), \\ \dots \\ x_n^{(k+1)} = \varphi_n\left(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}\right), \end{cases} \quad k = 0, 1, 2, \dots, \quad (2.23)$$

или, в векторной форме,

$$\mathbf{x}^{(k+1)} = \varphi(\mathbf{x}^{(k)}), \quad k = 0, 1, 2, \dots. \quad (2.24)$$

Если последовательность векторов $\mathbf{x}^{(k)} = \left(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}\right)^T$ сходится, то она сходится к решению $\mathbf{x}^{(*)} = \left(x_1^{(*)}, x_2^{(*)}, \dots, x_n^{(*)}\right)^T$.

Достаточное условие сходимости итерационного процесса (2.23) формулируется следующим образом [3].

Теорема 2.4. Пусть вектор-функция $\varphi(\mathbf{x})$ непрерывна вместе со своей производной

$$\varphi'(\mathbf{x}) = \begin{bmatrix} \frac{\partial \varphi_1(\mathbf{x})}{\partial x_1} & \frac{\partial \varphi_1(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial \varphi_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial \varphi_2(\mathbf{x})}{\partial x_1} & \frac{\partial \varphi_2(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial \varphi_2(\mathbf{x})}{\partial x_n} \\ \dots \\ \frac{\partial \varphi_n(\mathbf{x})}{\partial x_1} & \frac{\partial \varphi_n(\mathbf{x})}{\partial x_2} & \dots & \frac{\partial \varphi_n(\mathbf{x})}{\partial x_n} \end{bmatrix}$$

в ограниченной выпуклой замкнутой области G и

$$\max_{x \in G} \|\varphi'(\mathbf{x})\| \leq q < 1, \quad (2.25)$$

где q — постоянная. Если $\mathbf{x}^{(0)} \in G$ и все последовательные приближения

$$\mathbf{x}^{(k+x)} = \varphi(\mathbf{x}^{(k)}), \quad k = 0, 1, 2, \dots,$$

также содержатся в G , то процесс итерации (2.23) сходится к единственному решению уравнения

$$\mathbf{x} = \varphi(\mathbf{x})$$

в области G и справедливы оценки погрешности, здесь $\forall k \in N$:

$$\begin{aligned} \|\mathbf{x}^{(*)} - \mathbf{x}^{(k+x)}\| &\leq \frac{q^{k+1}}{1-q} \|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|, \\ \|\mathbf{x}^{(*)} - \mathbf{x}^{(k+x)}\| &\leq \frac{q^{k+1}}{1-q} \|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|. \end{aligned} \quad (2.26)$$

Пример 2.2 (продолжение). Найти положительное решение системы (2.19) методом простой итерации с погрешностью $\varepsilon = 10^{-4}$.

Решение. Преобразуем исходную систему уравнений (2.19) к виду

$$\begin{cases} x_1 = 0.3 - 0.1x_1^2 - 0.2x_2^2 \equiv \varphi_1(x_1, x_2), \\ x_2 = 0.7 - 0.2x_1^2 + 0.1x_1x_2 \equiv \varphi_2(x_1, x_2). \end{cases}$$

Проверим выполнение условия (2.25) в области G : $|x_1 - 0.25| \leq 0.25$, $|x_2 - 0.75| \leq 0.25$. Для этого найдем

$$\max_{x \in G} \|\varphi'(x)\| = \max_{x \in G} \left\{ \max_i \sum_{j=1}^n \left| \frac{\partial \varphi_i(x_1, x_2)}{\partial x_j} \right| \right\}. \quad (2.27)$$

Так как

$$\begin{aligned} \frac{\partial \varphi_1(x_1, x_2)}{\partial x_1} &= -0.2x_1, & \frac{\partial \varphi_1(x_1, x_2)}{\partial x_2} &= -0.4x_2, \\ \frac{\partial \varphi_2(x_1, x_2)}{\partial x_1} &= -0.4x_1 + 0.1x_2, & \frac{\partial \varphi_2(x_1, x_2)}{\partial x_2} &= 0.1x_1, \end{aligned}$$

то в области G имеем

$$\begin{aligned} \left| \frac{\partial \varphi_1(x_1, x_2)}{\partial x_1} \right| + \left| \frac{\partial \varphi_1(x_1, x_2)}{\partial x_2} \right| &= |-0.2x_1| + |-0.4x_2| \leq 0.5, \\ \left| \frac{\partial \varphi_2(x_1, x_2)}{\partial x_1} \right| + \left| \frac{\partial \varphi_2(x_1, x_2)}{\partial x_2} \right| &= |-0.4x_1 + 0.1x_2| + |0.1x_1| \leq 0.2, \end{aligned}$$

$$\max_{x \in G} \|\varphi'(x)\| \leq 0.5 \Rightarrow q < 1.$$

Следовательно, если последовательные приближения $(x_1^{(k)}, x_2^{(k)})$

не покинут области G (что легко обнаружить в процессе вычислений), то итерационный процесс будет сходящимся.

В качестве начального приближения примем $x_1^{(0)} = 0.25$, $x_2^{(0)} = 0.75$.

Последующие приближения определяем как

$$\begin{cases} x_1^{(k+1)} = \varphi_1(x_1^{(k)}, x_2^{(k)}), \\ x_2^{(k+1)} = \varphi_2(x_1^{(k)}, x_2^{(k)}), \end{cases} \quad k = 0, 1, 2, \dots,$$

где

$$\begin{aligned} \varphi_1(x_1^{(k)}, x_2^{(k)}) &= 0.3 - 0.1x_1^{(k)2} - 0.2x_2^{(k)2}, \\ \varphi_2(x_1^{(k)}, x_2^{(k)}) &= 0.7 - 0.2x_1^{(k)2} + 0.1x_1^{(k)}x_2^{(k)}. \end{aligned}$$

В соответствии с (2.26) вычисления завершаются при выполнении условия

$$\frac{q}{1-q} \left\| \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \right\| \leq \varepsilon,$$

где

$$\left\| \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)} \right\| = \max_i |x_i^{(k+1)} - x_i^{(k)}|.$$

Результаты вычислений содержатся в таблице 2.6.

Таблица 2.6

k	$x_1^{(k)}$	$x_1^{(k)}$	$\varphi_1(x_1^{(k)}, x_2^{(k)})$	$\varphi_2(x_1^{(k)}, x_2^{(k)})$
0	0.25000	0.7500	0.18125	0.70702
1	0.18125	0.70702	0.19674	0.70617
2	0.19674	0.70617	0.19639	0.70615
3	0.19639	0.70615	0.19641	0.70615
4	0.19641	0.70615		

Корни системы уравнений: $x_1^{(*)} \approx 0.1964$, $x_2^{(*)} \approx 0.7062$.

Замечание. В случае, когда при анализе сходимости конкретной итерационной схемы проверка условия (2.27) является затруднитель-

ной, можно определить норму матрицы $M(\mathbf{x})$ с элементами $m_{ij}(\mathbf{x}) = \max_{x \in G} \frac{\partial \varphi_i(\mathbf{x})}{\partial x_j}$ так, что $\max_{x \in G} \|\varphi'(\mathbf{x})\| \leq \|M(\mathbf{x})\|$. Если $\|M(\mathbf{x})\| \leq q \leq 1$, то последовательные приближения сходятся к решению $\mathbf{x}^{(*)}$ (см. [5], [6]).

Задачи

2.1. Методами простой итерации и Ньютона найти положительный корень нелинейного уравнения; начальное приближение определить графически.

1. $2^x - x^2 - 0.5 = 0$.
2. $\ln(x+2) - x^2 = 0$.
3. $\sqrt{1-x^2} - e^x + 0.1 = 0$.
4. $x^3 + x^2 - x - 0.5 = 0$.
5. $\cos x + 0.25x - 0.5 = 0$.
6. $e^x - 2x - 2 = 0$.
7. $2^x + x^2 - 2 = 0$.
8. $\ln(x+1) - 2x^2 + 1 = 0$.
9. $x^3 + x^2 - 2x - 1 = 0$.
10. $\sin x - 2x^2 + 0.5 = 0$.
11. $e^x - x^3 + 3x^2 - 2x - 3 = 0$.
12. $3^x - 5x^2 + 1 = 0$.
13. $\ln(x+1) - 2x^2 + 0.5 = 0$.
14. $x^3 - 2x^2 - 10x + 15 = 0$.
15. $\sin x - x^2 + 1 = 0$.
16. $xe^x + x^2 - 1 = 0$.
17. $4^x - 5x - 2 = 0$.
18. $\ln(x+1) - x^3 + 1 = 0$.
19. $x^4 - 2x - 1 = 0$.
20. $\operatorname{tg} x - 5x^2 + 1 = 0$, $x \in [-1, 1]$.
21. $3\sqrt{x+1} - e^x + 0.5 = 0$.
22. $10^x - 5x - 2 = 0$.
23. $\ln(x+2) - x^4 + 0.5 = 0$.
24. $x^6 - 5x - 2 = 0$.
25. $\sqrt{x+2} - 2\cos x = 0$.
26. $\lg(x+1) - x + 0.5 = 0$.
27. $x^6 - 5x^3 - 2 = 0$.
28. $\lg(2x+1) - x^3 + 1 = 0$.
29. $x^5 - 7x^2 + 3 = 0$.
30. $x \lg(x+2) + x^2 - 1 = 0$.

2.2. Методом Ньютона или методом простой итерации решить систему нелинейных уравнений (при наличии нескольких решений найти то из них, в котором значения неизвестных являются положительными); начальное приближение определить графически.

Вариант	Значение параметра a	Система уравнений
1	2	$\begin{cases} (x_1^2 + a^2)x_2 - a^3 = 0, \\ (x_1 - a/2)^2 + (x_2 - a/2)^2 - a^2 = 0. \end{cases}$
2	3	
3	4	

Окончание табл.

Вариант	Значение параметра a	Система уравнений
4 5 6	1 2 3	$\begin{cases} x_1 - \cos x_2 = 1, \\ x_2 - \lg(x_1 + 1) = a. \end{cases}$
7 8 9	2 3 4	$\begin{cases} x_1^2 + x_2^2 - a^2 = 0, \\ x_1 - e^{x_2} + a = 0. \end{cases}$
10 11 12	1 2 3	$\begin{cases} x_1 - \cos x_2 = a, \\ x_2 - \sin x_1 = a. \end{cases}$
13 14 15	2 3 4	$\begin{cases} x_1^2 / a^2 + x_2^2 / (a / 2)^2 - 1 = 0, \\ ax_2 - e^{x_1} - x_1 = 0. \end{cases}$
16 17 18	2 3 4	$\begin{cases} ax_1 - \cos x_2 = 0, \\ ax_2 - e^{x_1} = 0. \end{cases}$
19 20 21	1 2 3	$\begin{cases} x_1^2 + 2 \lg x_2 - 1 = 0, \\ x_1^2 - ax_1 x_2 + a = 0. \end{cases}$
22 23 24	1 2 3	$\begin{cases} ax_1^2 - x_1 + x_2^2 - 1 = 0, \\ x_2 - \operatorname{tg} x_1 = 0. \end{cases}$
25 26 27	1 2 3	$\begin{cases} ax_1^2 - x_2 + x_2^2 - a = 0, \\ x_1 - \sqrt{x_2 + a} + 1 = 0. \end{cases}$
28 29 30	4 5 6	$\begin{cases} e^{x_1 x_2} + x_1 - a = 0, \\ x_1^2 - ax_2 - 1 = 0. \end{cases}$

Ответы

2.1. Для всех вариантов $\varepsilon = 10^{-4}$.

1. $-00.4708; 1.5738; 4.1423.$
 3. $-0.9590; 0.0915.$
 5. $-0.7966; 1.4271; 4.1462.$
 7. $-1.2578; 0.6535.$
 9. $-1.8019; -0.4450; 1.2470.$
 11. $-0.5926; 1.0738.$

2. $-0.5876; 1.0571.$
 4. $-1.4516; -0.4030; 0.8546.$
 6. $-0.7680; 1.6784.$
 8. $-0.4492; 0.9070.$
 10. $-0.3112; 0.7744.$
 12. $-0.5556; 0.8380.$

- | | | | |
|-----|--------------------------|-----|--------------------------|
| 13. | -0.8997; 0.4282. | 14. | -3.0000; 1.3820; 3.6180. |
| 15. | -0.6368; 1.4096. | 16. | -1.1675; 0.4782. |
| 17. | -0.2607; 1.6942. | 18. | -0.7648; 1.2154. |
| 19. | -0.4747; 1.3953. | 20. | -0.3548; 0.5738. |
| 21. | -0.9095; 1.4294. | 22. | -0.2997; 0.7654. |
| 23. | -0.8837; 1.1319. | 24. | -0.3992; 1.4487. |
| 25. | -1.0667; 0.6262. | 26. | -0.9658; 0.7407. |
| 27. | -0.7194; 1.7514. | 28. | -0.4600; 1.1494. |
| 29. | -0.6426; 0.6691; 1.8273. | 30. | -1.0000; 0.8011. |

2.2. Для всех вариантов $\varepsilon = 10^{-4}$.

Вариант	x_1	x_2	Вариант	x_1	x_2
1	2.9646	0.6255	16	0.3737	0.7266
2	4.4469	0.9383	17	0.3001	0.4500
3	5.9293	1.2511	18	0.2375	0.3170
4	1.2222	1.3468	19	1.2758	2.0596
5	0.4445	2.1597	20	1.1488	1.4449
6	0.0094	3.0041	21	1.1006	1.2755
7	1.5480	1.2664	22	0.8198	1.0713
8	2.4721	1.6997	23	0.7022	0.8461
9	3.4587	2.0094	24	0.6111	0.7006
10	0.8322	1.7394	25	0.5692	1.4623
11	1.0389	2.8618	26	0.8455	1.4057
12	2.2104	3.8023	27	1.0000	1.0000
13	0.4249	0.9772	28	1.7125	0.4832
14	1.1178	1.3920	29	1.9592	0.5677
15	1.7091	1.8082	30	2.1706	0.6186

Глава 3

ПРИБЛИЖЕНИЕ ФУНКЦИЙ. ЧИСЛЕННОЕ ДИФФЕРЕНЦИРОВАНИЕ И ИНТЕГРИРОВАНИЕ

В настоящем разделе рассмотрены задачи приближения функций с помощью многочленов Лагранжа и Ньютона, с использованием сплайн-интерполяции и метода наименьших квадратов. Приводятся формулы численного дифференцирования, полученные на основе приближения функций и их производных многочленами. Описаны методы численного интегрирования — прямоугольников, трапеций, парабол (метод Симпсона), метод Рунге — Ромберга для оценки точности и уточнения результатов интегрирования. Все методы иллюстрируются подробно разобранными примерами решения характерных задач. Для самостоятельного решения приведены пять задач с ответами; в каждой задаче содержится 30 вариантов условий.

3.1. Интерполяция

Пусть на отрезке $[a, b]$ задано множество несовпадающих точек x_i (интерполяционных узлов), в которых известны значения функции $f_i = f(x_i)$, $i = 0, \dots, n$. Приближающая функция $\phi(x, a)$ такая, что выполняются равенства

$$\phi(x_i, a_0, \dots, a_n) = f(x_i) = f_i, \quad i = 0, \dots, n, \quad (3.1)$$

называется *интерполяционной*.

Наиболее часто в качестве приближающей функции используют многочлены степени n

$$P_n(x) = \sum_{i=0}^n a_i x^i. \quad (3.2)$$

Подставляя в (3.2) значения узлов интерполяции и используя условие $P_n(x_i) = f_i$, получаем систему линейных алгебраических уравнений относительно коэффициентов a_i :

$$\sum_{i=0}^n a_i x^i = f_k, \quad k = 0, \dots, n, \quad (3.3)$$

которая в случае несовпадения узлов интерполяции имеет единственное решение.

Для нахождения интерполяционного многочлена не обязательно решать систему (3.3). Произвольный многочлен может быть записан в виде

$$L_n(x) = \sum_{i=0}^n f_i l_i(x). \quad (3.4)$$

Здесь $l_i(x)$ — многочлены степени n , так называемые *лагранжевы многочлены влияния*, которые удовлетворяют условию

$$l_i(x_j) = \begin{cases} 1, & \text{при } i=j, \\ 0, & \text{при } i \neq j, \end{cases}$$

и, соответственно,

$$l_i(x_j) = \prod_{j=0, j \neq i}^n \frac{(x - x_j)}{(x_i - x_j)},$$

а интерполяционный многочлен (3.4) запишется в виде

$$L_n(x) = \sum_{i=0}^n f_i \prod_{j=0, j \neq i}^n \frac{(x - x_j)}{(x_i - x_j)}. \quad (3.5)$$

Интерполяционный многочлен, записанный в форме (3.5), называется *интерполяционным многочленом Лагранжа*.

Если ввести функцию $\omega_{n+1}(x) = (x - x_0)(x - x_1)\dots(x - x_n) = \prod_{i=0}^n (x - x_i)$, то выражение для интерполяционного многочлена Лагранжа примет вид

$$L_n(x) = \sum_{i=0}^n f_i \frac{\omega_{n+1}(x)}{(x - x_i) \omega'_{n+1}(x_i)}. \quad (3.6)$$

Недостатком интерполяционного многочлена Лагранжа является необходимость полного пересчета всех коэффициентов в случае добавления дополнительных интерполяционных узлов. Чтобы избежать указанного недостатка, используют интерполяционный многочлен в форме Ньютона.

Введем понятие *разделенной разности*. Разделенные разности нулевого порядка совпадают со значениями функции в узлах. Разделенные разности первого порядка обозначаются как $f'(x_i, x_j)$ и определяются через разделенные разности нулевого порядка:

$$f(x_i, x_j) = \frac{f_i - f_j}{x_i - x_j}.$$

Разделенные разности второго порядка определяются через разделенные разности первого порядка:

$$f(x_i, x_j, x_k) = \frac{f(x_i, x_j) - f(x_j, x_k)}{x_i - x_k}.$$

Разделенная разность произвольного порядка определяется соотношениями

$$f(x_i, x_j, x_k, \dots, x_l, x_n) = \frac{f(x_i, x_j, x_k, \dots, x_l) - f(x_j, x_k, \dots, x_n)}{x_i - x_n}. \quad (3.7)$$

Таким образом, для $(n + 1)$ -й точки могут быть построены разделенные разности до n -го порядка; разделенные разности более высоких порядков равны нулю.

Пусть известны значения аппроксимируемой функции $f(x)$ в точках x_0, x_1, \dots, x_n . Интерполяционный многочлен, значения которого в узлах интерполяции совпадают со значениями функции $f(x)$, может быть записан в виде

$$\begin{aligned} P_n(x) = & f(x_0) + (x - x_0)f(x_1, x_0) + (x - x_0)(x - x_1)f(x_0, x_1, x_2) + \\ & + \dots + (x - x_0)(x - x_1)\dots(x - x_n)f(x_0, x_1, \dots, x_n). \end{aligned} \quad (3.8)$$

Запись многочлена в виде (3.8) называется *интерполяционным многочленом Ньютона*. Если функция $f(x)$ не есть многочлен n -й степени, то формула (3.8) для $P_n(x)$ приближает функцию $f(x)$ с некоторой погрешностью. Отметим, что при добавлении новых узлов первые члены многочлена Ньютона остаются неизменными.

Если функция задана в точках x_0, x_1, \dots, x_n , принадлежащих некоторому отрезку $[a, b]$, то при построении интерполяционного многочлена Ньютона удобно пользоваться таблицей разделенных разностей, пример которой для $n = 4$ приведен в таблице 3.1.

Таблица 3.1

x_0	$f(x_0)$	$f(x_0, x_1)$	$f(x_0, x_1, x_2)$	$f(x_0, x_1, x_2, x_3)$	$f(x_0, x_1, x_2, x_3, x_4)$
x_1	$f(x_1)$	$f(x_1, x_2)$	$f(x_1, x_2, x_3)$	$f(x_1, x_2, x_3, x_4)$	
x_2	$f(x_2)$	$f(x_2, x_3)$	$f(x_2, x_3, x_4)$		
x_3	$f(x_3)$	$f(x_3, x_4)$			
x_4	$f(x_4)$				

Замечание. Для повышения точности интерполяции в сумму (3.8) могут быть добавлены новые члены, что требует подключения дополнительных интерполяционных узлов. При этом безразлично, в каком порядке подключаются новые узлы. Этим формула Ньютона выгодно отличается от формулы Лагранжа.

Погрешность интерполяционных многочленов Лагранжа и Ньютона для случая аналитически заданной функции $f(x)$ априорно может быть оценена по формуле, вывод которой приводится, например, в [1]:

$$|\varepsilon_n(x)| = |f(x) - P_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |\omega_{n+1}(x)|, \quad (3.9)$$

где $M_{n+1} = \max \left| f^{(n+1)}(\xi) \right|$, $\xi \in [a, b]$.

Если величину производных аппроксимируемой функции оценить сложно (например, для таблично заданной функции), то используется апостериорная оценка по первому отброшенному члену интерполяционного многочлена Ньютона, в который входят разделенные разности, являющиеся аналогами производных соответствующих порядков.

Использование одной интерполяционной формулы на большом числе узлов нецелесообразно. Интерполяционный многочлен может проявить свои колебательные свойства, при этом его значения между узлами могут существенно отличаться от значений интерполируемой функции. Одна из возможностей преодоления этого недостатка заключается в применении *сплайн-интерполяции*.

Суть сплайн-интерполяции заключается в определении интерполирующей функции по формулам одного типа для различных непересекающихся промежутков и в стыковке значений функции и ее производных на их границах.

Наиболее широко применяемым является случай, когда между любыми двумя точками разбиения исходного отрезка строится многочлен n -й степени:

$$S(x) = \sum_{k=0}^n a_{ik} x^k, \quad x_{i-1} \leq x \leq x_i, \quad i = 1, \dots, n, \quad (3.10)$$

который в узлах интерполяции принимает значения аппроксимируемой функции и непрерывен вместе со своими $(n-1)$ -й производными. Здесь точки x_0, x_1, \dots, x_n упорядочены по возрастанию, т.е.

$$x_0 < x_1 < \dots < x_n, \quad x_0 = a, \quad x_n = b.$$

Такой кусочно-непрерывный интерполяционный многочлен называется *сплайном*. Его коэффициенты находятся из условий равенства в узлах сетки значений сплайна и приближаемой функции, а также равенства $(n - 1)$ -й производных соответствующих многочленов. На практике наиболее часто используется интерполяционный многочлен третьей степени, который удобно представить как

$$S(x) = a_i + b_i(x - x_{i-1}) + c_i(x - x_{i-1})^2 + d_i(x - x_{i-1})^3, \quad (3.11)$$

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

Для построения кубического сплайна необходимо построить n многочленов третьей степени, т.е. определить $4n$ неизвестных a_i, b_i, c_i, d_i . Эти коэффициенты ищутся из условий в узлах сетки:

$$\begin{aligned} S(x_{i-1}) &= a_i = a_{i-1} + b_{i-1}(x_{i-1} - x_{i-2}) + c_{i-1}(x_{i-1} - x_{i-2})^2 \\ &\quad + d_{i-1}(x_{i-1} - x_{i-2})^3 = f_{i-1}, \\ S'(x_{i-1}) &= b_i = b_{i-1} + 2c_{i-1}(x_{i-1} - x_{i-2}) + 3d_{i-1}(x_{i-1} - x_{i-2})^2, \\ S''(x_{i-1}) &= 2c_i = 2c_{i-1} + 6d_{i-1}(x_{i-1} - x_{i-2}), \\ S(x_0) &= a_1 = f_0, \\ S''(x_0) &= c_1 = 0, \\ S(x_n) &= a_n + b_n(x_n - x_{n-1}) + c_n(x_n - x_{n-1})^2 + d_n(x_n - x_{n-1})^3 = f_n, \\ S''(x_n) &= 2c_n + 6d_n(x_n - x_{n-1}) = 0, \quad i = 2, 3, \dots, n. \end{aligned} \quad (3.12)$$

В (3.12) предполагается, что сплайны имеют нулевую кривизну на концах отрезка. В общем случае могут быть использованы и другие условия.

Если ввести обозначение $h_i = x_i - x_{i-1}$ и исключить из системы (3.12) неизвестные a_i, b_i, d_i , то можно получить систему из $(n - 1)$ линейных алгебраических уравнений относительно c_i , $i = 2, \dots, n$, с трехдиагональной матрицей:

$$\begin{aligned} 2(h_1 + h_2)c_2 + h_2c_3 &= 3\left[\left(f_2 - f_1\right)/h_2 - \left(f_1 - f_0\right)/h_1\right], \\ \dots & \\ h_{i-1}c_{i-1} + 2(h_{i-1} + h_i)c_i + h_ic_{i+1} &= 3\left[\left(f_i - f_{i-1}\right)/h_i - \left(f_{i-1} - f_{i-2}\right)/h_{i-1}\right], \quad (3.13) \\ \dots & \\ h_{n-1}c_{n-1} + 2(h_{n-1} + h_n)c_n &= 3\left[\left(f_n - f_{n-1}\right)/h_n - \left(f_{n-1} - f_{n-2}\right)/h_{n-1}\right], \\ i = 3, \dots, n-1. & \end{aligned}$$

Остальные коэффициенты сплайнов могут быть восстановлены по формулам

$$\begin{aligned} a_i &= f_{i-1}, \\ b_i &= \frac{f_i - f_{i-1}}{h_i} = \frac{1}{3} h_i (c_{i+1} + 2c_i), & d_i &= \frac{c_{i+1} - c_i}{3h_i}, \\ b_n &= \frac{f_n - f_{n-1}}{h_n} = \frac{2}{3} h_n c_n, & d_n &= -\frac{c_n}{3h_n}, & i &= 1, \dots, n-1. \end{aligned} \quad (3.14)$$

Пример 3.1. Используя таблицу значений y_i функции $y = f(x)$, вычисленных в точках x_i , $i = 0, \dots, 3$, построить многочлен Лагранжа, проходящий через точки $\{x_i, y_i\}$, вычислить значение погрешности интерполяции в точке X^* для $y = \ln(x)$, $x_i = 0.1, 0.5, 0.9, 1.3$; $X^* = 0.8$ и сравнить ее с оценкой (3.9).

Решение. Функция $y = \ln(x)$ задана в четырех точках, следовательно, искомым является многочлен Лагранжа третьей степени

$$\begin{aligned} L_3(x) &= \sum_{i=0}^3 f_i \frac{\omega_4(x)}{(x-x_i)\omega'_4(x_i)}, \\ \omega_4(x) &= (x-x_0)(x-x_1)(x-x_2)(x-x_3), \\ \omega'_4(x_0) &= (x_0-x_1)(x_0-x_2)(x_0-x_3), \\ \omega'_4(x_1) &= (x_1-x_0)(x_1-x_2)(x_1-x_3), \\ \omega'_4(x_2) &= (x_2-x_0)(x_2-x_1)(x_2-x_3), \\ \omega'_4(x_3) &= (x_3-x_0)(x_3-x_1)(x_3-x_2). \end{aligned}$$

Заполним таблицу:

i	x_i	f_i	$\omega'_4(x_i)$	$f_i / \omega'_4(x_i)$	$X^* - x_i$
0	0.1	-2.30259	-0.384	5.99632	0.7
1	0.5	-0.69315	0.128	-5.41521	0.3
2	0.9	-0.10536	-0.128	0.82313	-0.1
3	1.3	0.26236	0.384	0.68324	-0.5

Искомый многочлен Лагранжа может быть записан в виде

$$\begin{aligned} L_3(x) &= 5.99632(x-0.5)(x-0.9)(x-1.3) - \\ &\quad - 5.41521(x-0.1)(x-0.9)(x-1.3) + \\ &\quad + 0.82313(x-0.1)(x-0.5)(x-1.3) + \\ &\quad + 0.68324(x-0.1)(x-0.5)(x-0.9). \end{aligned}$$

Вычислим значение интерполяционного многочлена и точное значение функции в точке $X^* = 0.8$:

$$L_3(0.8) = -0.20036, \quad y(0.8) = \ln(0.8) = -0.22314.$$

Абсолютная погрешность интерполяции составляет $\Delta(L_3(0.8)) = 0.02278$. Оценка погрешности интерполяции по формуле (3.9) составляет

$$|\epsilon_n(0.8)| \leq \frac{6 \cdot 10^4}{24} \cdot 1.05 \cdot 10^{-2} = 26.25.$$

Пример 3.2. Используя таблицу значений y_i функции $y = f(x)$, вычисленных в точках x_i , $i = 0, \dots, 3$, построить многочлен Ньютона, проходящий через точки $\{x_i, y_i\}$.

Вычислить значение погрешности интерполяции в точке X^* и сравнить ее с оценкой (3.9).

Решение. Имеем $y = \sin(\pi x/6)$, $X_i = 0, 1.0, 2.0, 3.0$; $X^* = 1.5$.

Функция $y = \sin(\pi x/6)$ задана в четырех точках, следовательно, искомым является многочлен Ньютона третьей степени

$$\begin{aligned} P_3(x) = & f(x_0) + (x - x_0)f(x_1, x_0) + (x - x_0)(x - x_1)f(x_0, x_1, x_2) + \\ & + (x - x_0)(x - x_1)(x - x_2)f(x_0, x_1, x_2, x_3). \end{aligned}$$

Заполним таблицу конечных разностей:

i	x_i	f_i	$f(x_i, x_{i+1})$	$f(x_i, x_{i+1}, x_{i+2})$	$f(x_0, x_1, x_2, x_3)$
0	0.0	0.0	0.5		
1	1.0	0.5	0.36603	-0.06699	
2	2.0	0.86603	0.13398	-0.11603	-0.01635
3	3.0	1.0			

Искомый многочлен Ньютона записывается в виде

$$P_3(x) = 0.5x - 0.06699x(x - 1.0) - 0.01635x(x - 1.0)(x - 2.0).$$

Вычислим значение интерполяционного многочлена и точное значение функции в точке $X^* = 1.5$:

$$P_3(1.5) = 0.70589, \quad y(1.5) = \sin(\pi/4) = 0.70711.$$

Абсолютная погрешность интерполяции составляет $\Delta(P_3(1.5)) = 0.00122$, $|\epsilon_4(1.5)| = (0.0751/24) \cdot 0.5625 = 0.0017$.

Пример 3.3. Построить кубический сплайн для функции, заданной в узлах интерполяции, предполагая, что сплайн имеет нулевую кривизну при $x = x_0$ и $x = x_4$; вычислить значение функции $f(1.5)$.

i	0	1	2	3	4
x_i	0.0	1.0	2.0	3.0	4.0
f_i	0.0	1.8415	2.9093	3.1411	3.2432

Решение. Запишем систему уравнений (3.13):

$$\begin{cases} 4c_2 + c_3 = -2.3209, \\ c_2 + 4c_3 + c_4 = -2.5080, \\ c_3 + 4c_4 = -0.38924. \end{cases}$$

Решив данную систему, найдем c_2, c_3, c_4 и, воспользовавшись формулами (3.14), заполним таблицу.

i	$[x_{i-1}, x_i]$	a_i	b_i	c_i	d_i
1	$[0, 1]$	0.0	1.9913	0.0	-0.14983
2	$[1, 2]$	1.8415	1.5418	-0.44949	-0.02450
3	$[2, 3]$	2.9093	0.56934	-0.52299	0.18548
4	$[3, 4]$	3.1411	0.07978	0.03344	-0.01115

Имеем

$$f(x) = a_i + b_i(x - x_{i-1}) + c_i(x - x_{i-1})^2 + d_i(x - x_{i-1})^3,$$

$$x_{i-1} \leq x \leq x_i, \quad i = 1, 2, 3, 4.$$

Вычислим значение функции $f(1.5)$. Точка $x = 1.5$ принадлежит отрезку $[1, 2]$, на этом отрезке таблично заданная функция представляется кубическим сплайном:

$$f(x) = 1.8415 + 1.5418(x - 1) - 0.44949(x - 1)^2 - 0.0245(x - 1)^3,$$

$$1 \leq x \leq 2, \quad i = 2.$$

Получаем $f(1.5) = 2.4969$.

3.2. Метод наименьших квадратов

Пусть задана таблично в узлах x_j функция $y_j = f(x_j), j = 0, 1, \dots, N$. При этом значения функции y_j определены с некоторой погрешностью. Пусть также из физических соображений известен вид функции, которой должны приближенно удовлетворять таблич-

ные точки, например многочлен степени n , у которого неизвестны коэффициенты a_i , $F_n(x) = \sum_{i=0}^n a_i x^i$. Неизвестные коэффициенты будем находить из условия минимума квадратичного отклонения многочлена от таблично заданной функции:

$$\Phi = \sum_{j=0}^N \left[F_n(x_j) - y_j \right]^2. \quad (3.15)$$

Минимума Φ можно добиться только за счет изменения коэффициентов многочлена $F_n(x)$. Необходимые условия экстремума имеют вид

$$\frac{\partial \Phi}{\partial a_k} = 2 \sum_{j=0}^N \left[\sum_{i=0}^n a_i x_j^i - y_j \right] x_j^k = 0, \quad k = 0, 1, \dots, n. \quad (3.16)$$

Эту систему для удобства преобразуют к виду

$$\sum_{i=0}^n a_i \sum_{j=0}^N x_j^{k+1} = \sum_{j=0}^N y_j x_j^k = 0, \quad k = 0, 1, \dots, n. \quad (3.17)$$

Система (3.17) называется *нормальной системой метода наименьших квадратов* (МНК) и представляет собой систему линейных алгебраических уравнений относительно коэффициентов a_i . Решив систему, построим многочлен $F_n(x)$, приближающий функцию $f(x)$ и минимизирующий квадратичное отклонение.

Необходимо отметить, что система (3.17) с увеличением степени n приближающего многочлена становится плохо обусловленной и решение ее связано с потерей точности. Поэтому при использовании метода наименьших квадратов, как правило, используют приближающий многочлен не выше третьей степени.

Пример 3.4. Для таблично заданной функции путем решения нормальной системы МНК найти приближающие многочлены первой и второй степени. Для каждого из приближающих многочленов вычислить сумму квадратов ошибок. Построить графики приближаемой функции и приближающих многочленов.

i	0	1	2	3	4	5
x_i	0.0	1.7	3.4	5.1	6.8	8.5
y_i	0.0	1.3038	1.8439	2.2583	2.6077	2.9155

Решение. Найдем приближающий многочлен первой степени $F_1(x) = a_0 + a_1x$. Для нахождения неизвестных коэффициентов a_0, a_1 запишем нормальную систему МНК (3.17):

$$\begin{cases} a_0(N+1) + a_1 \sum_{j=0}^N x_j = \sum_{j=0}^N y_j, \\ a_0 \sum_{j=0}^N x_j + a_1 \sum_{j=0}^N x_j^2 = \sum_{j=0}^N y_j x_j. \end{cases} \quad (3.18)$$

В данном примере $N = 5$, значения $x_i, y_i, i = 0, \dots, 5$, приведены в таблице. Подставив числовые значения в (3.18), получим

$$\begin{cases} 6a_0 + 25.5a_1 = 10.9292, \\ 25.5a_0 + 158.95a_1 = 62.517. \end{cases} \quad (3.19)$$

Решив систему (3.19), получим $a_0 = 0.4713, a_1 = 0.3177$. Таким образом найден приближающий многочлен первой степени $F_1(x) = 0.4713 + 0.3177x$, значения которого в заданных точках приведены в таблице:

i	0	1	2	3	4	5
x_i	0.0	1.7	3.4	5.1	6.8	8.5
$F_1(x_i)$	0.4713	1.0114	1.5515	2.0916	2.6317	3.1718

Сумма квадратов ошибок составляет $\Phi = \sum_{j=0}^5 [F_1(x_j) - y_j]^2 = 0.4872$.

Найдем приближающий многочлен второй степени $F_2(x) = a_0 + a_1x + a_2x^2$. Для нахождения неизвестных коэффициентов a_0, a_1, a_2 запишем нормальную систему МНК (3.17):

$$\begin{cases} a_0(N+1) + a_1 \sum_{j=0}^N x_j + a_2 \sum_{j=0}^N x_j^2 = \sum_{j=0}^N y_j, \\ a_0 \sum_{j=0}^N x_j + a_1 \sum_{j=0}^N x_j^2 + a_2 \sum_{j=0}^N x_j^3 = \sum_{j=0}^N y_j x_j, \\ a_0 \sum_{j=0}^N x_j^2 + a_1 \sum_{j=0}^N x_j^3 + a_2 \sum_{j=0}^N x_j^4 = \sum_{j=0}^N y_j x_j^2. \end{cases} \quad (3.20)$$

Подставив числовые значения в (3.20), получим

$$\begin{cases} 6a_0 + 25.5a_1 + 158.95a_2 = 10.9292, \\ 25.5a_0 + 158.95a_1 + 1105.43a_2 = 62.517, \\ 158.95a_0 + 1105.43a_1 + 8176.71a_2 = 415.045. \end{cases} \quad (3.21)$$

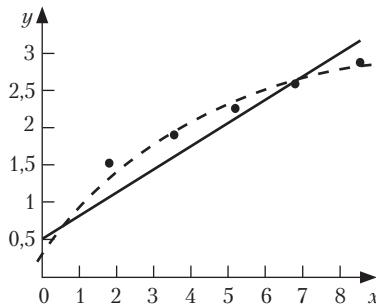


Рис. 3.1

Решив систему (3.21), найдем $a_0 = 0.1295$, $a_1 = 0.6193$, $a_2 = -0.0355$. Таким образом, сформирован приближающий многочлен второй степени $F_2(x) = 0.1295 + 0.6193x - 0.0355x^2$, значения которого в заданных точках приведены в таблице:

i	0	1	2	3	4	5
x_i	0.0	1.7	3.4	5.1	6.8	8.5
$F_2(x_i)$	0.1295	1.0798	1.8250	2.3651	2.7000	2.8299

Сумма квадратов ошибок составляет $\Phi = \sum_{j=0}^5 [F_2(x_j) - y_j]^2 = 0.0946$.

На рис. 3.1 точками обозначены табличные данные, сплошной линией — приближающий многочлен первой степени, пунктирной — приближающий многочлен второй степени.

3.3. Численное дифференцирование

Формулы численного дифференцирования в основном используются при нахождении производных от функции $y = f(x)$, заданной таблично. Исходная функция $y_i = f(x_i)$, $i = 0, 1, \dots, M$, на отрезках $[x_j, x_{j+k}]$ заменяется некоторой приближающей, легко вычисляемой функцией $\phi(x, \mathbf{a})$, $y = \phi(x, \mathbf{a}) + R(x)$, где $R(x)$ — остаточный член приближения, \mathbf{a} — вектор коэффициентов, вообще говоря, различный для каждого из рассматриваемых отрезков, и полагают, что $y'(x) \approx \phi'(x, \mathbf{a})$. Наиболее часто в качестве приближающей функции $\phi(x, \mathbf{a})$ берется интерполяционный многочлен $\phi(x, \mathbf{a}) = P_n(x) = \sum_{i=0}^n a_i x^i$, а производные соответствующих порядков определяются дифференцированием многочлена.

При решении практических задач, как правило, используются аппроксимации первых и вторых производных.

В первом приближении таблично заданная функция может быть аппроксимирована отрезками прямой

$$y(x) \approx \varphi(x) = y_i + \frac{y_{i+1} - y_i}{x_{i+1} - x_i}(x - x_i), \quad x \in [x_i, x_{i+1}].$$

В этом случае производная является кусочно-постоянной функцией и рассчитывается по формуле

$$y'(x) \approx \varphi'(x) = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} = \text{const}, \quad x \in [x_i, x_{i+1}], \quad (3.22)$$

с первым порядком точности в крайних точках интервала, и со вторым порядком точности в средней точке интервала [1]. Правая часть (3.22) совпадает с выражением для разделенной разности $f(x_i, x_{i+1})$ — см. п. 3.1.

При использовании для аппроксимации таблично заданной функции интерполяционного многочлена второй степени имеем

$$\begin{aligned} y(x) \approx \varphi(x) = & y_i + \frac{y_{i+1} - y_i}{x_{i+1} - x_i}(x - x_i) + \\ & + \frac{\frac{y_{i+2} - y_{i+1}}{x_{i+2} - x_{i+1}} - \frac{y_{i+1} - y_i}{x_{i+1} - x_i}}{\frac{x_{i+2} - x_{i+1}}{x_{i+2} - x_i}}(x - x_i)(x - x_{i+1}), \quad x \in [x_i, x_{i+1}], \end{aligned} \quad (3.23)$$

$$\begin{aligned} y'(x) \approx \varphi'(x) = & \frac{y_{i+1} - y_i}{x_{i+1} - x_i} + \\ & + \frac{\frac{y_{i+2} - y_{i+1}}{x_{i+2} - x_{i+1}} - \frac{y_{i+1} - y_i}{x_{i+1} - x_i}}{\frac{x_{i+2} - x_{i+1}}{x_{i+2} - x_i}}(2x - x_i - x_{i+1}), \quad x \in [x_i, x_{i+1}]. \end{aligned} \quad (3.24)$$

При равноотстоящих точках разбиения данная формула обеспечивает второй порядок точности.

Для вычисления второй производной необходимо использовать интерполяционный многочлен как минимум второй степени. После дифференцирования многочлена получаем

$$y''(x) \approx \varphi''(x) = 2 \frac{\frac{y_{i+2} - y_{i+1}}{x_{i+2} - x_{i+1}} - \frac{y_{i+1} - y_i}{x_{i+1} - x_i}}{\frac{x_{i+2} - x_i}{x_{i+2} - x_i}}, \quad x \in [x_i, x_{i+1}]. \quad (3.25)$$

Формулы (3.23)–(3.25) с использованием разделенных разностей могут быть записаны в виде

$$\begin{aligned}y(x) &\approx \varphi(x) = y_i + f(x_i, x_{i+1})(x - x_i) + f(x_i, x_{i+1}, x_{i+2})(x - x_i)(x - x_{i+1}), \\y'(x) &\approx \varphi'(x) = f(x_i, x_{i+1}) + f(x_i, x_{i+1}, x_{i+2})(2x - x_i - x_{i+1}), \\y''(x) &\approx \varphi''(x) = 2f(x_i, x_{i+1}, x_{i+2}).\end{aligned}$$

В случае равноотстоящих узлов с шагом $h = (x_{i+1} - x_i) = x_{i+2} - x_{i+1}$ имеем

$$\begin{aligned}y(x) &\approx \varphi(x) = y_i + \frac{y_{i+1} - y_i}{h}(x - x_i) + \frac{y_{i+2} - 2y_{i+1} - y_i}{2h^2}(x - x_i)(x - x_{i+1}), \\y'(x) &\approx \varphi'(x) = \frac{y_{i+1} - y_i}{h} + \frac{y_{i+2} - 2y_{i+1} - y_i}{2h^2}(2x - x_i - x_{i+1}), \\y''(x) &\approx \varphi''(x) = \frac{y_{i+2} - 2y_{i+1} - y_i}{2h^2}.\end{aligned}$$

Пример 3.5. Вычислить первую и вторую производные от таблично заданной функции $y_i = f(x_i)$, $i = 0, 1, 2, 3, 4$, в точке $x = X^*$; $X^* = 0.2$.

i	0	1	2	3	4
x_i	0.0	0.1	0.2	0.3	0.4
y_i	1.0	1.1052	1.2214	1.3499	1.4918

Решение. Вычислим производную, используя формулу (3.22) и отрезок $[x_1, x_2]$. Так как точка, в которой требуется найти значение производной, совпадает с правой границей отрезка, такую производную еще называют *левосторонней*:

$$y'(0.2) = \frac{y_2 - y_1}{x_2 - x_1} = \frac{1.2214 - 1.1052}{0.2 - 0.1} = 1.1623.$$

Аналогично вычисляется *правосторонняя* производная:

$$y'(0.2) = \frac{y_3 - y_2}{x_3 - x_2} = \frac{1.3499 - 1.2214}{0.3 - 0.2} = 1.2846.$$

Обе эти формулы позволяют вычислить производную с первым порядком точности. Вычислим производную со вторым порядком точности, для чего воспользуемся формулой (3.24):

$$\begin{aligned}y'(0.2) &= \frac{y_2 - y_1}{x_2 - x_1} + \frac{x_3 - x_2}{x_3 - x_1} \frac{x_2 - x_1}{(2 \cdot 0.2 - x_1 - x_2)} (2 \cdot 0.2 - x_1 - x_2) = \frac{1.2214 - 1.1052}{0.2 - 0.1} + \\&+ \frac{\frac{1.3499 - 1.2214}{0.3 - 0.2} - \frac{1.2214 - 1.1052}{0.2 - 0.1}}{\frac{0.3 - 0.2}{0.3 - 0.1}} (2 \cdot 0.2 - 0.1 - 0.2) = 1.2234.\end{aligned}$$

Заметим, что результат вычисления по формуле (3.24) в случае равномерной сетки совпадает с полусуммой левосторонней и правосторонней производных.

Вычислим вторую производную в точке $x = 0.2$, используя соотношение (3.25):

$$y''(0.2) = 2 \frac{\frac{y_3 - y_2}{x_3 - x_2} - \frac{y_2 - y_1}{x_2 - x_1}}{x_3 - x_1} = \frac{\frac{1.3499 - 1.2214}{0.3 - 0.2} - \frac{1.2214 - 1.1052}{0.2 - 0.1}}{0.3 - 0.1} = 1.2224.$$

3.4. Численное интегрирование

Формулы численного интегрирования используются в тех случаях, когда вычислить аналитически определенный интеграл

$F = \int_a^b f(x) dx$ не удается. Отрезок $[a, b]$ разбивают точками $x_0, x_1,$

..., x_N так, что $a = x_0, b = X_N, A < X_1 < \dots < B$ с достаточно мелким шагом $H_I = x_i - X_{I-1}$. При этом на одном или нескольких отрезках H_I подынтегральную функцию $F(X)$ заменяют приближающей $\varphi(x)$ такой, что она, во-первых, близка к $f(x)$, а во-вторых, интеграл от $\varphi(x)$ легко вычисляется. Рассмотрим наиболее простой и часто применяемый способ, когда подынтегральную функцию

заменяют на интерполяционный многочлен $P_n(x) = \sum_{j=0}^n a_j x^j$, при-

чем коэффициенты многочлена a_j , вообще говоря, различны на каждом отрезке $[x_i, x_{i+n}]$ и определяются из условия $\varphi(x_l) = f(x_l), l = i, \dots, i + n$, т.е. многочлен P_n зависит от параметров $a_l = P_n(x, \bar{a}_i)$. В этом случае

$$f(x) = P_n(x, \bar{a}_i) + R_n(x, \bar{a}_i), \quad x \in [x_i, x_{i+k}], \quad (3.26)$$

где $R_n(x, \bar{a}_i)$ — остаточный член интерполяции. Тогда

$$F = \sum_{i=1}^N \int_{x_{i-1}}^{x_i} P_n(x, \bar{a}_i) dx + R,$$

где $R = \sum_{i=1}^N \int_{x_{i-1}}^{x_i} R_n(x, \bar{a}_i) dx$ — остаточный член формулы численного

интегрирования или ее погрешность. Здесь $\bar{a}_i = (a_0, a_1, \dots, a_n)$ — век-

тор, состоящий из коэффициентов интерполяционного многочлена P_n , аппроксимирующего функцию $f(x)$ на отрезке $[x_i, x_{i+n}]$.

При использовании интерполяционных многочленов различной степени получают формулы численного интегрирования различного порядка точности.

Заменив подынтегральную функцию интерполяционным многочленом Лагранжа нулевой степени, проходящим через середину отрезка — точку $x_{i+1/2} = (x_{i+1} + x_i)/2$, получим формулу прямоугольников:

$$\int_a^b f(x) dx \approx \sum_{i=1}^N h_i f\left(\frac{x_{i-1} + x_i}{2}\right). \quad (3.27)$$

В случае постоянного шага интегрирования $h_i = h$, $i = 1, 2, \dots, N$, и существования $f''(x)$, $x \in [a, b]$, имеет место следующая оценка остаточного члена формулы прямоугольников:

$$R \leq \frac{1}{24} h^2 M_2 (b-a), \quad (3.28)$$

где $M_2 = \max|f''(x)|$, $x \in [a, b]$.

Для таблично заданных функций удобно в качестве узлов интерполяции выбрать начало и конец отрезка интегрирования, т.е. заменить функцию $f(x)$ многочленом Лагранжа первой степени

$$F = \int_a^b f(x) dx \approx \frac{1}{2} \sum_{i=1}^N (f_i + f_{i-1}) h_i. \quad (3.29)$$

Эта формула носит название *формулы трапеций*.

В случае постоянного шага интегрирования величина остаточного члена оценивается как

$$R \leq \frac{b-a}{12} h^2 M_2, \quad (3.30)$$

где $M_2 = \max|f''(x)|$, $x \in [a, b]$.

Для повышения порядка точности формулы численного интегрирования заменим подынтегральную кривую параболой — интерполяционным многочленом второй степени, выбрав в качестве узлов интерполяции концы и середину отрезка интегрирования: $x_{i-1}, x_{i+1/2} = (x_{i+1} + x_i)/2, x_i$.

В случае $h_i = (x_i - x_{i-1/2})/2$ получим формулу Симпсона (формулу парабол):

$$F = \int_a^b f(x) dx \approx \frac{1}{3} \sum_{i=1}^N (f_{i-1} + 4f_{i-1/2} + f_i) h_i. \quad (3.31)$$

При постоянном шаге интегрирования $h_i = h$, $i = 1, 2, \dots, N$, формула Симпсона принимает вид

$$F \approx \frac{h}{3} [f_0 + 4f_{1/2} + 2f_1 + 4f_{3/2} + 2f_2 + \dots + 2f_{N-1} + 4f_{N-1/2} + f_N], \quad (3.32)$$

здесь число интервалов, на которое делится отрезок интегрирования, равно $2N$. В случае если существует $f^{IV}(x)$, $x \in [a, b]$, для оценки величины погрешности справедлива мажорантная оценка

$$R \leq \frac{b-a}{180} h^4 M_4, \quad (3.33)$$

где $M_4 = \max |f^{IV}(x)|$, $x \in [a, b]$.

Метод Рунге – Ромберга – Ричардсона позволяет получать более высокий порядок точности вычисления. Если имеются результаты вычисления определенного интеграла на сетке с шагом h – $F = F_h + O(h^p)$ и на сетке с шагом kh – $F = F_{kh} + O((kh)^p)$, то

$$F = \int_a^b f(x) dx \approx F_h + \frac{F_h - F_{kh}}{k^p - 1} + O(h^{p+1}). \quad (3.34)$$

Пример 3.6. Вычислить определенный интеграл $F = \int_{X_0}^{X_1} y dx$ методами

прямоугольников, трапеций, Симпсона с шагами h_1, h_2 . Уточнить полученные значения, используя метод Рунге – Ромберга – Ричардсона. Принять

$$y = \frac{x}{(3x+4)^2}, \quad X_0 = -1, \quad X_k = 1, \quad h_1 = 0.5, \quad h_2 = 0.25.$$

Решение. В случае интегрирования с постоянным шагом формулы принимают следующий вид:

- метод прямоугольников:

$$F = h \left[y\left(\frac{x_0+x_1}{2}\right) + y\left(\frac{x_1+x_2}{2}\right) + \dots + y\left(\frac{x_{N-1}+x_N}{2}\right) \right];$$

- метод трапеций:

$$F = h \left[\frac{y_0}{2} + y_1 + y_2 + \dots + y_{N-1} + \frac{y_N}{2} \right];$$

- метод Симпсона:

$$F = \frac{h}{3} \left[y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \dots + 2y_{N-2} + 4y_{N-1} + y_N \right].$$

В настоящем примере используется формула метода Симпсона, аналогичная (3.32), но без введения полуцелых индексов; в этом случае число узлов N должно быть четным.

Вычислим интеграл с шагом 0.5, результаты занесем в таблицу:

i	x_i	y_i	$\int_{x_0}^{x_i} y dx$		
			Метод прямоугольников	Метод трапеций	Метод Симпсона
0	-1.0	-1.0	0.0	0.0	0.0
1	-0.5	-0.08	-0.12245	-0.27	
2	0.0	0.0	-0.13428	-0.29	
3	0.5	0.01653	-0.12874	-0.28587	-0.22
4	1.0	0.02041	-0.11914	-0.27663	-0.20558

Вычислим интеграл с шагом 0.25, результаты занесем в таблицу:

i	x_i	y_i	$\int_{x_0}^{x_i} y dx$		
			Метод прямоугольников	Метод трапеций	Метод Симпсона
0	-1.0	-1.0	0.0	0.0	0.0
1	-0.75	-0.24490	-0.11570	-0.15561	
2	-0.5	-0.08	-0.15031	-0.19622	
3	-0.25	-0.02367	-0.16165	-0.20918	-0.17163
4	0.0	0.0	-0.16403	-0.21214	-0.18619
5	0.25	0.01108	-0.16239	-0.21076	
6	0.5	0.01653	-0.15882	-0.20731	
7	0.75	0.01920	-0.15430	-0.20284	-0.18112
8	1.0	0.02041	-0.14914	-0.19789	-0.17164

Сравним полученные результаты с оценками (3.28), (3.30), (3.33) и получим:

- для метода прямоугольников

$$R(h=0.5) \leq 1.375, \quad R(h=0.25) \leq 0.34375;$$

- метода трапеций

$$R(h=0.5) \leq 2.75, \quad R(h=0.25) \leq 0.6875;$$

- метода Симпсона

$$R(h=0.5) \leq 8.6625, \quad R(h=0.25) \leq 0.5414.$$

Уточним значение интеграла $F = \int_1^4 [x / (3x + 4)] dx$, используя метод

Рунге — Ромберга — Ричардсона, получим при точном значении x , равном -0.16474 :

- по методу прямоугольников
—0.15937 с абсолютной погрешностью 0.00537;
- по методу трапеций
—0.17164 с абсолютной погрешностью 0.00690;
- по методу Симпсона
—0.16938 с абсолютной погрешностью 0.00464.

Задачи

3.1. Используя таблицу значений Y_i функции $y = f(x)$, вычисленных в точках $X_i, i=0, \dots, 3$, построить интерполяционные многочлены Лагранжа и Ньютона, проходящие через точки (X_i, Y_i) . Вычислить значение погрешности $\Delta(y(X^*))$ интерполяции в точке X^* . Сравнить погрешности метода с оценкой (3.9).

№	Функция	X_i		X^*
		a	b	
1.	$y = \sin x$	0.1π, 0.2π, 0.3π, 0.4π	0.1π, π/6, 0.3π, 0.4π	π/4
2.	$y = \cos x$	0, π/6, 2π/6, 3π/6	0, π/6, 5π/12, π/2	π/4
3.	$y = \operatorname{tg} x$	0, π/8, 2π/8, 3π/8	0, π/8, π/3, 3π/8	3π/16
4.	$y = \operatorname{ctg} x$	π/8, 2π/8, 3π/8, 4π/8,	π/8, 5π/16, 3π/8, π/2	π/3
5.	$y = \ln x$	0.2, 0.6, 1.0, 1.4	0.2, 0.6, 1.2, 1.4	0.8
6.	$y = e^x$	-2, -1, 0, 1	-2, -1, 0.2, 1	-0.5
7.	$y = \sqrt{x}$	0, 1.7, 3.4, 5.1	0, 1.7, 4.0, 5.1	3.0
8.	$y = \arcsin x$	-0.4, -0.1, 0.2, 0.5	-0.4, 0, 0.2, 0.5	0.1

Окончание табл.

№	Функция	X_i		X^*
		a	b	
9.	$y = \arccos x$	-0.4, -0.1, 0.2, 0.5	-0.4, 0, 0.2, 0.5	0.1
10.	$y = \operatorname{arctg} x$	-3, -1, 1, 3	-3, 0, 1, 3	-0.5
11.	$y = \operatorname{arcctg} x$	-3, -1, 1, 3	-3, 0, 1, 3	-0.5
12.	$y = \sin x + x$	0, $\pi/6$, $2\pi/6$, $3\pi/6$	0, $\pi/6$, $5\pi/12$, $\pi/2$	$\pi/4$
13.	$y = \cos x + x$	0, $\pi/6$, $2\pi/6$, $3\pi/6$	0, $\pi/6$, $5\pi/12$, $\pi/2$	$\pi/4$
14.	$y = \operatorname{tg} x + x$	0, $\pi/8$, $2\pi/8$, $3\pi/8$	0, $\pi/8$, $\pi/3$, $3\pi/8$	$3\pi/16$
15.	$y = \operatorname{ctg} x + x$	$\pi/8$, $2\pi/8$, $3\pi/8$, $4\pi/8$	$\pi/8$, $\pi/3$, $3\pi/8$, $\pi/2$	$3\pi/16$
16.	$y = \ln x + x$	0.1, 0.5, 0.9, 1.3	0.1, 0.5, 1.1, 1.3	0.8
17.	$y = e^x + x$	-2, -1, 0, 1	-2, -1, 0.2, 1	-0.5
18.	$y = \sqrt{x} + x$	0, 1.7, 3.4, 5.1	0, 1.7, 4.0, 5.1	3.0
19.	$y = \arcsin x + x$	-0.4, -0.1, 0.2, 0.5	-0.4, 0, 0.2, 0.5	0.1
20.	$y = \operatorname{arccos} x + x$	-0.4, -0.1, 0.2, 0.5	-0.4, 0, 0.2, 0.5	0.1
21.	$y = \operatorname{arctg} x + x$	-3, -1, 1, 3	-3, 0, 1, 3	-0.5
22.	$y = \operatorname{arcctg} x + x$	-3, -1, 1, 3	-3, 0, 1, 3	-0.5
23.	$y = 1/x$	0.1, 0.5, 0.9, 1.3	0.1, 0.5, 1.1, 1.3	0.8
24.	$y = 1/x^2$	0.1, 0.5, 0.9, 1.3	0.1, 0.5, 1.1, 1.3	0.8
25.	$y = 1/x + x$	0.1, 0.5, 0.9, 1.3	0.1, 0.5, 1.1, 1.3	0.8
26.	$y = 1/x^2 + x^2$	0.1, 0.5, 0.9, 1.3	0.1, 0.5, 1.1, 1.3	0.8
27.	$y = x \sin(x)$	0, $\pi/6$, $2\pi/6$, $3\pi/6$	0, $\pi/6$, $5\pi/12$, $\pi/2$	$\pi/4$
28.	$y = x \cos(x)$	0, $\pi/6$, $2\pi/6$, $3\pi/6$	0, $\pi/6$, $5\pi/12$, $\pi/2$	$\pi/4$
29.	$y = xe^x$	-2, -1, 0, 1	-2, -1, 0.2, 1	-0.5
30.	$y = x^2 e^x$	-1.2, -0.7, -0.2, 0.3	-1.2, -0.7, 0, 0.3	-0.5

3.2. Построить кубический сплайн для функции, заданной в узлах интерполяции, предполагая, что сплайн имеет нулевую кривизну при $x = x_0$ и $x = x_4$. Вычислить значение функции в точке $x = X^*$.

$$1. X^* = 1.5$$

i	0	1	2	3	4
x_i	0.0	1.0	2.0	3.0	4.0
f_i	0.0	0.5	0.86603	1.0	0.86603

2. $X^* = 1.5$

i	0	1	2	3	4
x_i	0.0	1.0	2.0	3.0	4.0
f_i	1.0	0.86603	0.5	0.0	-0.5

3. $X^* = 1.5$

i	0	1	2	3	4
x_i	0.0	0.9	1.8	2.7	3.6
f_i	0.0	0.36892	0.85408	1.7856	6.3138

4. $X^* = 2.66666667$

i	0	1	2	3	4
x_i	1.0	1.9	2.8	3.7	4.6
f_i	2.4142	1.0818	0.50953	0.11836	-0.24008

5. $X^* = 0.8$

i	0	1	2	3	4
x_i	0.1	0.5	0.9	1.3	1.7
f_i	-2.3026	-0.69315	-0.10536	0.26236	0.53063

6. $X^* = -0.5$

i	0	1	2	3	4
x_i	-2.0	-1.0	0.0	1.0	2.0
f_i	0.13534	0.36788	1.0	2.7183	7.3891

7. $X^* = 3.0$

i	0	1	2	3	4
x_i	0.0	1.7	3.4	5.1	6.8
f_i	0.0	1.3038	1.8439	2.2583	2.6077

8. $X^* = 0.1$

i	0	1	2	3	4
x_i	-0.4	-0.1	0.2	0.5	0.8
f_i	-0.41152	-0.10017	0.20136	0.52360	0.92730

9. $X^* = 0.1$

i	0	1	2	3	4
x_i	-0.4	-0.1	0.2	0.5	0.8
f_i	1.9823	1.6710	1.3694	1.0472	0.64350

10. $X^* = -0.5$

i	0	1	2	3	4
x_i	-3.0	-1.0	1.0	3.0	5.0
f_i	-1.2490	-0.78540	0.78540	1.2490	1.3734

11. $X^* = -0.5$

i	0	1	2	3	4
x_i	-3.0	-1.0	1.0	3.0	5.0
f_i	2.8198	2.3562	0.78540	0.32175	0.19740

12. $X^* = 0.8$

i	0	1	2	3	4
x_i	0.0	0.5	1.0	1.5	2.0
f_i	0.0	0.97943	1.8415	2.4975	2.9093

13. $X^* = 1.5$

i	0	1	2	3	4
x_i	0.0	1.0	2.0	3.0	4.0
f_i	1.0	1.5403	1.5839	2.01	3.3464

14. $X^* = 1.5$

i	0	1	2	3	4
x_i	0.0	0.9	1.8	2.7	3.6
f_i	0.0	0.72235	1.5609	2.8459	7.7275

15. $X^* = 2.66666667$

i	0	1	2	3	4
x_i	1.0	1.9	2.8	3.7	4.6
f_i	2.8069	1.8279	1.6091	1.5713	1.5663

16. $X^* = 0.8$

i	0	1	2	3	4
x_i	0.1	0.5	0.9	1.3	1.7
f_i	-2.2026	-0.19315	0.79464	1.5624	2.2306

17. $X^* = -0.5$

i	0	1	2	3	4
x_i	-2.0	-1.0	0.0	1.0	2.0
f_i	-1.8647	-0.63212	1.0	3.7183	9.3891

18. $X^* = 3.0$

i	0	1	2	3	4
x_i	0.0	1.7	3.4	5.1	6.8
f_i	0.0	3.0038	5.3439	7.3583	9.4077

19. $X^* = 0.1$

i	0	1	2	3	4
x_i	-0.4	-0.1	0.2	0.5	0.8
f_i	-0.81152	-0.20017	0.40136	1.0236	1.7273

20. $X^* = 0.1$

i	0	1	2	3	4
x_i	-0.4	-0.1	0.2	0.5	0.8
f_i	1.5823	1.5710	1.5694	1.5472	1.4435

21. $X^* = -0.5$

i	0	1	2	3	4
x_i	-3.0	-1.0	1.0	3.0	5.0
f_i	-4.2490	-1.7854	-1.7854	4.2490	6.3734

22. $X^* = -0.5$

i	0	1	2	3	4
x_i	-3.0	-1.0	1.0	3.0	5.0
f_i	-0.18016	1.3562	1.7854	3.3218	5.1974

23. $X^* = 0.8$

i	0	1	2	3	4
x_i	0.1	0.5	0.9	1.3	1.7
f_i	10.0	2.0	1.1111	0.76923	0.58824

24. $X^* = 0.8$

i	0	1	2	3	4
x_i	0.1	0.5	0.9	1.3	1.7
f_i	100.0	4.0	1.2346	0.59172	0.34602

25. $X^* = 0.8$

i	0	1	2	3	4
x_i	0.1	0.5	0.9	1.3	1.7
f_i	10.1	2.5	2.0111	2.0692	2.2882

26. $X^* = 0.8$

i	0	1	2	3	4
x_i	0.1	0.5	0.9	1.3	1.7
f_i	100.01	4.2500	2.0446	2.2817	3.2360

27. $X^* = 1.5$

i	0	1	2	3	4
x_i	0.0	1.0	2.0	3.0	5.0
f_i	0.0	0.26180	0.90690	1.5708	1.3090

28. $X^* = 1.5$

i	0	1	2	3	4
x_i	0.0	1.0	2.0	3.0	5.0
f_i	0.0	0.45345	0.52360	0.0	-2.2672

29. $X^* = -0.5$

i	0	1	2	3	4
x_i	-2.0	-1.0	0.0	1.0	2.0
f_i	-0.27067	-0.36788	0.0	2.7183	14.778

30. $X^* = -0.5$

i	0	1	2	3	4
x_i	-1.2	-0.7	-0.2	0.3	0.8
f_i	0.43372	0.24333	0.03275	0.12149	1.4243

3.3. Для таблично заданной функции путем решения нормальной системы МНК найти приближающие многочлены первой (*а*) и второй (*б*) степени. Для каждого из приближающих многочленов вычислить сумму квадратов ошибок Φ . Построить графики приближаемой функции и приближающих многочленов.

1.

i	0	1	2	3	4	5
x_i	-0.0	0.0	1.0	2.0	3.0	4.0
y_i	-0.5	0.0	0.5	0.86603	1.0	0.86603

2.

i	0	1	2	3	4	5
x_i	-1.0	0.0	1.0	2.0	3.0	4.0
y_i	0.86603	1.0	0.86603	0.50	0.0	-0.50

3.

i	0	1	2	3	4	5
x_i	-0.9	0.0	0.9	1.8	2.7	3.6
y_i	-0.36892	0.0	0.36892	0.85408	1.7856	6.3138

4.

i	0	1	2	3	4	5
x_i	1.0	1.9	2.8	3.7	4.6	5.5
y_i	2.4142	1.0818	0.50953	0.11836	-0.24008	-0.66818

5.

i	0	1	2	3	4	5
x_i	0.1	0.5	0.9	1.3	1.7	2.1
y_i	-2.3026	-0.69315	-0.10536	0.26236	0.53063	0.74194

6.

i	0	1	2	3	4	5
x_i	-3.0	-2.0	-1.0	0.0	1.0	2.0
y_i	0.04979	0.13534	0.36788	1.0	2.7183	7.3891

7.

i	0	1	2	3	4	5
x_i	0.0	0.2	0.4	0.6	0.8	1.0
y_i	0.1	1.0032	1.0512	1.2592	1.8192	3.0

8.

i	0	1	2	3	4	5
x_i	-0.7	-0.4	-0.1	0.2	0.5	0.8
y_i	-0.7754	-0.41152	-0.10017	0.20136	0.5236	0.9273

9.

i	0	1	2	3	4	5
x_i	-0.7	-0.4	-0.1	0.2	0.5	0.8
y_i	2.3462	1.9823	1.671	1.3694	1.0472	0.6435

10.

i	0	1	2	3	4	5
x_i	-5.0	-3.0	-0.1	1.0	3.0	5.0
y_i	-1.3734	-1.249	-0.7854	0.7854	1.249	1.3734

11.

i	0	1	2	3	4	5
x_i	-5.0	-3.0	-0.1	1.0	3.0	5.0
y_i	2.9442	2.8198	2.3562	0.7854	0.32175	0.1974

12.

i	0	1	2	3	4	5
x_i	-1.0	0.0	1.0	2.0	3.0	4.0
y_i	-1.8415	0.0	1.8415	2.9093	2.1411	3.2432

13.

i	0	1	2	3	4	5
x_i	-1.0	0.0	1.0	2.0	3.0	4.0
y_i	-0.4597	1.0	1.5403	1.5839	2.010	3.3464

14.

i	0	1	2	3	4	5
x_i	-0.9	0.0	0.9	1.8	2.7	3.6
y_i	-1.2689	0.0	1.2689	2.6541	4.4856	9.9138

15.

i	0	1	2	3	4	5
x_i	1.0	1.9	2.8	3.7	4.6	5.5
y_i	3.4142	2.9818	3.3095	3.8184	4.3599	4.8318

16.

i	0	1	2	3	4	5
x_i	0.1	0.5	0.9	1.3	1.7	2.1
y_i	-2.2026	-0.19315	0.79464	1.5624	2.2306	2.8419

17.

i	0	1	2	3	4	5
x_i	-3.0	-2.0	-1.0	0.0	1.0	2.0
y_i	-2.9502	-1.8647	-0.63212	1.0	3.7183	9.3891

18.

i	0	1	2	3	4	5
x_i	0.0	1.7	3.4	5.1	6.8	8.5
y_i	0.0	3.0038	5.2439	7.3583	9.4077	11.415

19.

i	0	1	2	3	4	5
x_i	-0.7	-0.4	-0.1	0.2	0.5	0.8
y_i	-1.4754	-0.81152	-0.200017	0.40136	1.0236	1.7273

20.

i	0	1	2	3	4	5
x_i	-0.7	-0.4	-0.1	0.2	0.5	0.8
y_i	1.6462	1.5823	1.571	1.5694	1.5472	1.4435

21.

i	0	1	2	3	4	5
x_i	-5.0	-3.0	-0.1	1.0	3.0	5.0
y_i	-6.3734	-4.249	-1.7854	1.7854	4.249	6.3734

22.

i	0	1	2	3	4	5
x_i	-5.0	-3.0	-0.1	1.0	3.0	5.0
y_i	-2.0558	-0.18016	1.3562	1.7854	3.3218	5.1974

23.

i	0	1	2	3	4	5
x_i	0.1	0.5	0.9	1.3	1.7	2.1
y_i	10	2.0	1.1111	0.76923	0.58824	0.47619

24.

i	0	1	2	3	4	5
x_i	0.1	0.5	0.9	1.3	1.7	2.1
y_i	100	4.0	1.2346	0.59172	0.34602	0.22676

25.

i	0	1	2	3	4	5
x_i	0.1	0.5	0.9	1.3	1.7	2.1
y_i	10.1	2.5	2.0111	2.0692	2.2882	2.5762

26.

i	0	1	2	3	4	5
x_i	0.1	0.5	0.9	1.3	1.7	2.1
y_i	100.01	4.250	2.0446	2.2817	3.236	4.6368

27.

i	0	1	2	3	4	5
x_i	-1.0	0.0	1.0	2.0	3.0	5.0
y_i	0.5	0.0	0.5	1.7321	3.0	2.5

28.

i	0	1	2	3	4	5
x_i	-1.0	0.0	1.0	2.0	3.0	5.0
y_i	-0.86603	0.0	0.86603	1.0	0.0	-4.3301

29.

i	0	1	2	3	4	5
x_i	-3.0	-2.0	-1.0	0.0	1.0	2.0
y_i	-0.14936	-0.27067	-0.36788	0.0	2.7183	14.778

30.

i	0	1	2	3	4	5
x_i	-1.7	-1.2	-0.7	-0.2	0.3	0.8
y_i	0.52796	0.43372	0.24333	0.03275	0.12149	1.4243

3.4. Вычислить первую и вторую производную от таблично заданной функции $y_i = f(x_i)$, $i = 0, 1, 2, 3, 4$, в точке $x = X^*$.

1. $X^* = 1.0$

i	0	1	2	3	4
x_i	-1.0	0.0	1.0	2.0	3.0
y_i	-0.5	0.0	0.50	0.86603	1.0

2. $X^* = 1.0$

i	0	1	2	3	4
x_i	-1.0	0.0	1.0	2.0	3.0
y_i	-0.5	0.0	0.5	0.86603	1.0

3. $X^* = 2.0$

i	0	1	2	3	4
x_i	1.0	1.5	2.0	2.5	3.0
y_i	0.0	0.40547	0.69315	0.91629	1.0985

4. $X^* = 0.2$

i	0	1	2	3	4
x_i	0.0	0.1	0.2	0.3	0.4
y_i	1.0	1.1052	1.2214	1.3499	1.4918

5. $X^* = 2.0$

i	0	1	2	3	4
x_i	0.0	1.0	2.0	3.0	4.0
y_i	0.0	1.0	1.4142	1.7321	2.0

6. $X^* = 0.2$

i	0	1	2	3	4
x_i	-0.2	0.0	0.2	0.4	0.6
y_i	-0.20136	0.0	0.20136	0.41152	0.64350

7. $X^* = 0.2$

i	0	1	2	3	4
x_i	-0.2	0.0	0.2	0.4	0.6
y_i	1.7722	1.5708	1.3694	1.1593	0.9273

8. $X^* = 1.0$

i	0	1	2	3	4
x_i	-1.0	0.0	1.0	2.0	3.0
y_i	-0.7854	0.0	0.7854	1.1071	1.249

9. $X^* = 1.0$

i	0	1	2	3	4
x_i	-1.0	0.0	1.0	2.0	3.0
y_i	2.3562	1.5708	0.7854	0.46365	0.32175

10. $X^* = 1.0$

i	0	1	2	3	4
x_i	0.0	0.5	1.0	1.5	2.0
y_i	0.0	0.97943	1.8415	2.4975	2.9093

11. $X^* = 1.0$

i	0	1	2	3	4
x_i	0.0	0.5	1.0	1.5	2.0
y_i	1.0	1.3776	1.5403	1.5707	1.5839

12. $X^* = 0.2$

i	0	1	2	3	4
x_i	-1.0	-0.4	0.2	0.6	1.0
y_i	-1.4142	-0.55838	0.2787	0.84008	1.4142

13. $X^* = 0.8$

i	0	1	2	3	4
x_i	0.2	0.5	0.8	1.1	1.4
y_i	12.906	5.5273	3.8777	3.2692	3.0319

14. $X^* = 3.0$

i	0	1	2	3	4
x_i	1.0	2.0	3.0	4.0	5.0
y_i	1.0	2.6931	4.0986	5.3863	6.6094

15. $X^* = 0.4$

i	0	1	2	3	4
x_i	0.0	0.2	0.4	0.6	0.8
y_i	1.0	1.4214	1.8918	2.4221	3.0255

16. $X^* = 2.0$

i	0	1	2	3	4
x_i	0.0	1.0	2.0	3.0	4.0
y_i	0.0	2.0	3.4142	4.7321	6.0

17. $X^* = 0.2$

i	0	1	2	3	4
x_i	-0.2	0.0	0.2	0.4	0.6
y_i	-0.40136	0.0	0.40136	0.81152	1.2435

18. $X^* = 0.2$

i	0	1	2	3	4
x_i	-0.2	0.0	0.2	0.4	0.6
y_i	1.5722	1.5708	1.5694	1.5593	1.5273

19. $X^* = 1.0$

i	0	1	2	3	4
x_i	-1.0	0.0	1.0	2.0	3.0
y_i	-1.7854	0.0	1.7854	3.1071	4.249

20. $X^* = 1.0$

i	0	1	2	3	4
x_i	-1.0	0.0	1.0	2.0	3.0
y_i	1.3562	1.5708	1.7854	2.4636	3.3218

21. $X^* = 2.0$

i	0	1	2	3	4
x_i	1.0	1.5	2.0	2.5	3.0
y_i	1.0	0.66667	0.50	0.40	0.33333

22. $X^* = 1.4$

i	0	1	2	3	4
x_i	1.0	1.2	1.4	1.6	1.8
y_i	1.0	0.69444	0.5102	0.39062	0.30864

23. $X^* = 2.0$

i	0	1	2	3	4
x_i	1.0	1.5	2.0	2.5	3.0
y_i	2.0	2.1667	2.5	2.9040	3.3333

24. $X^* = 1.4$

i	0	1	2	3	4
x_i	1.0	1.2	1.4	1.6	1.8
y_i	2.0	2.1344	2.4702	2.9506	3.5486

25. $X^* = 2.0$

i	0	1	2	3	4
x_i	0.0	1.0	2.0	3.0	4.0
y_i	0.0	0.5	1.7321	3.0	3.4641

26. $X^* = 2.0$

i	0	1	2	3	4
x_i	0.0	1.0	2.0	3.0	4.0
y_i	0.0	0.86603	1.0	0.0	-2.0

27. $X^* = 0.0$

i	0	1	2	3	4
x_i	-1.0	-0.5	0.0	0.5	1.0
y_i	-0.36788	-0.30327	0.0	0.82436	2.7183

28. $X^* = 0.4$

i	0	1	2	3	4
x_i	0.0	0.2	0.4	0.6	0.8
y_i	0.0	0.048856	0.23869	0.65596	1.4243

29. $X^* = 1.0$

i	0	1	2	3	4
x_i	-1.0	0.0	1.0	2.0	3.0
y_i	-0.5	0.0	0.5	0.86603	1.0

30. $X^* = 2.0$

i	0	1	2	3	4
x_i	0.0	1.0	2.0	3.0	4.0
y_i	0.0	0.5	0.86603	1.0	0.86603

3.5. Вычислить определенный интеграл $F = \int_{X_0}^{X_1} y dx$ методами прямоугольников, трапеций, Симпсона с шагами h_1, h_2 . Уточнить полученные значения, используя метод Рунге – Ромберга. Оценить погрешность методов интегрирования, используя формулы (3.28), (3.30), (3.33).

№	Функция	X_0	X_1	h_1	h_2
1	$y = \frac{x}{2x+5}$	-1	1	0.5	0.25
2	$y = \frac{x}{(3x+4)^2}$	0	4	1.0	0.5
3	$y = \frac{x}{(3x+4)^3}$	-1	1	0.5	0.25
4	$y = \frac{3x+4}{2x+7}$	-2	2	1.0	0.5
5	$y = \frac{1}{(2x+7)(3x+4)}$	-1	1	0.5	0.25
6	$y = \frac{x}{(2x+7)(3x+4)}$	-1	1	0.5	0.25
7	$y = \frac{1}{3x^2+4x+2}$	-2	2	1.0	0.5
8	$y = \frac{1}{x^2+4}$	-2	2	1.0	0.5
9	$y = \frac{x}{x^2+9}$	0	2	0.5	0.25

Продолжение табл.

№	Функция	X_0	X_1	h_1	h_2
10	$y = \frac{x^2}{x^2 + 16}$	0	2	0.5	0.25
11	$y = \frac{1}{x^3 + 64}$	-2	2	1.0	0.5
12	$y = \frac{x}{x^3 + 8}$	-1	1	0.5	0.25
13	$y = \frac{x^2}{x^3 - 27}$	-2	2	1.0	0.5
14	$y = \frac{1}{x^4 + 16}$	0	2	0.5	0.25
15	$y = \frac{x}{x^4 + 81}$	0	2	0.5	0.25
16	$y = \frac{x^2}{x^4 + 256}$	0	2	0.5	0.25
17	$y = \frac{1}{256 - x^4}$	-2	2	1.0	0.5
18	$y = \frac{1}{16 - x^4}$	-1	1	0.5	0.25
19	$y = \frac{x^2}{625 - x^4}$	0	4	1.0	0.5
20	$y = \frac{\sqrt{x}}{4 + 3x}$	1	5	1.0	0.5
21	$y = \frac{\sqrt{x}}{(1+2x)^2}$	1	5	1.0	0.5
22	$y = x\sqrt{2x + 3}$	-1	1	0.5	0.25
23	$y = \frac{1}{\sqrt{(2x+7)(3x+4)}}$	0	4	1.0	0.5
24	$y = \sqrt{16 - x^2}$	-2	2	1.0	0.5
25	$y = \sqrt{49 - x^2}$	-2	2	1.0	0.5
26	$y = x^2\sqrt{36 - x^2}$	1	5	1.0	0.5

Окончание табл.

№	Функция	X_0	X_1	h_1	h_2
27	$y = \sqrt{9 + x^2}$	1	5	1.0	0.5
28	$y = x^3 \sqrt{4 + x^2}$	1	5	1.0	0.5
29	$y = \sqrt{x^2 - 36}$	6.5	8.5	0.5	0.25
30	$y = x^3 \sqrt{x^2 - 49}$	7.5	9.5	0.5	0.25

Ответы

3.1.

Вариант	α		β	
	$y(X^*)$	$\Delta(y(X^*))$	$y(X^*)$	$\Delta(y(X^*))$
1	0.70695	0.00016	0.70685	0.00026
2	0.70589	0.00122	0.70481	0.00230
3	0.64461	0.02357	0.58525	0.08293
4	0.55798	0.01938	0.57350	0.00385
5	-0.20778	0.01536	-0.19690	0.02624
6	0.59108	0.01545	0.58395	0.02258
7	1.75178	0.01973	1.77471	0.04266
8	0.10006	0.00011	0.10009	0.00007
9	1.47074	0.00011	1.47070	0.00007
10	-0.41000	0.05365	-0.41000	0.05365
11	1.98079	0.05365	1.98079	0.05365
12	1.49129	0.00122	1.48995	0.00255
13	1.49129	0.00122	1.49021	0.00230
14	1.23366	0.02357	1.17430	0.08293
15	2.15155	0.06589	2.20059	0.11494
16	0.59964	0.02278	0.63411	0.05725
17	0.09108	0.01545	0.08395	0.02258
18	4.75178	0.01973	4.77471	0.04266
19	0.20006	0.00011	0.20009	0.00007
20	1.57074	0.00011	1.57070	0.00007
21	-0.91000	0.05365	-0.91000	0.05365
22	1.48079	0.05365	1.48079	0.05365
23	1.02564	0.22436	0.69930	0.55070
24	-1.83213	3.39463	-6.65852	8.22102
25	1.82564	0.22436	1.49930	0.55070
26	-1.19213	3.39463	-6.01852	8.22102
27	0.55922	0.00386	0.56229	0.00693
28	0.54959	0.00577	0.54336	0.01200
29	-0.35991	0.05664	-0.38708	0.08382
30	0.14363	0.00807	0.13721	0.01442

3.2.

Вариант	Отрезок	Коэффициенты сплайна $f(x)$				$f(X^*)$
		a	b	c	d	
1	[1, 2]	0.5	0.451808	-0.0722874	-0.0134956	0.706145
2	[1, 2]	0.866025	-0.239148	-0.157761	0.03008835	0.710872
3	[1, 2]	0.368919	0.550968	0.235096	-0.275910	0.724539
4	[1.9, 2.8]	1.08179	-1.05544	0.708373	-0.269070	0.567736
5	[0.5, 0.9]	-0.693147	2.72502	-4.86964	4.32685	-0.197082
6	[-1.0, 0.0]	0.367879	0.396883	0.246509	-0.0112712	0.626539
7	[1.7, 3.4]	1.30384	0.535474	-0.204257	0.0447927	1.75317
8	[-0.1, 0.2]	-0.100167	1.02012	-0.0885522	0.128102	0.101339
9	[-0.1, 0.2]	1.67096	-1.02012	0.0885522	-0.128102	1.46946
10	[-1, 1]	-0.785398	0.601405	0.277186	-0.0925945	-0.426974
11	[-1, 1]	2.35619	-0.601405	-0.277186	0.0925945	1.99777
12	[0.5, 1]	0.979426	1.87451	-0.253037	-0.0955856	1.51642
13	[1, 2]	1.54030	0.252035	-0.432401	0.223917	1.58621
14	[0.9, 1.8]	0.722349	0.943667	0.235096	-0.275910	1.31359
15	[1.9, 2.8]	1.82792	-0.662744	0.708373	-0.269070	1.61493
16	[0.5, 0.9]	-0.193147	3.72502	-4.86964	4.32685	0.602918
17	[-1, 0]	-0.632121	1.39688	0.246509	-0.0112712	0.126539
18	[1.7, 3.4]	3.00384	1.53547	-0.204257	0.0447927	4.75317
19	[-0.1, 0.2]	-0.200167	2.02012	-0.0885522	0.128102	0.201339
20	[-0.1, 0.2]	1.57096	-0.020121	0.0885522	-0.128102	1.56946
21	[-1, 1]	-1.78540	1.60140	0.277186	-0.0925945	-0.926974
22	[-1, 1]	1.35619	0.398595	-0.277186	0.0925945	1.49777
23	[0.5, 0.9]	2.0	-10.6572	35.0356	-34.8704	1.01454
24	[0.5, 0.9]	4.0	-115.855	465.544	-482.977	-1.89787
25	[0.5, 0.9]	2.5	-9.65719	35.0356	-34.8704	1.81454
26	[0.5, 0.9]	4.25	-114.912	466.830	-483.335	-1.25895
27	[1, 2]	0.261799	0.445714	0.275873	-0.0764867	0.544064
28	[1, 2]	0.453450	0.317090	-0.204539	-0.042402	0.55556
29	[-1, 0]	-0.367879	0.149801	0.370514	-0.152435	-0.219405
30	[-0.7, -0.2]	0.243327	-0.401208	-0.0612669	0.0427454	0.160976

3.3.

Вариант	a			δ			
	a_0	a_1	Φ	a_0	a_1	a_2	Φ
1	0.01836	0.29132	0.27082	0.07353	0.53957	-0.08275	0.01518
2	0.89232	-0.29132	0.27082	0.94749	-0.04307	-0.08275	0.01518
3	-0.19013	1.24620	8.67875	-0.46449	-0.12561	0.50808	2.35562
4	2.57557	-0.62758	0.47014	3.54759	-1.39808	0.11854	0.12596
5	-1.77444	1.37583	0.98535	-2.46043	3.40610	-0.92285	0.17139
6	2.58735	1.28792	11.4547	1.21263	1.80344	0.51552	1.53299
7	0.61813	1.80800	0.81696	1.09051	-1.73486	3.54286	0.06720
8	0.00553	1.10670	0.00394	-0.00699	1.10188	0.04814	0.00324

Окончание табл.

Вариант	<i>a</i>			<i>b</i>			
	<i>a</i> ₀	<i>a</i> ₁	Φ	<i>a</i> ₀	<i>a</i> ₁	<i>a</i> ₂	Φ
9	1.56527	-1.10670	0.00394	1.57779	-1.10188	-0.04814	0.00324
10	0.0	0.32570	0.70071	0.0	0.32570	0.0	0.70071
11	1.57080	-0.32570	0.70071	1.57080	-0.32570	0.0	0.70071
12	0.00974	1.02613	2.80932	0.18993	1.83696	-0.27028	0.08211
13	0.55616	0.63154	0.78843	0.56894	0.68904	-0.01917	0.77471
14	-0.19013	2.24620	8.67875	-0.46450	0.87439	0.50808	2.35562
15	2.57557	0.37242	0.47014	3.54759	-0.39808	0.11854	0.12596
16	-1.77444	2.37583	0.98534	-2.46043	4.40610	-0.92285	0.17139
17	2.58735	2.28792	11.4547	1.21263	2.80344	0.51552	1.53299
18	0.47130	1.31770	0.48722	0.12945	1.61933	-0.03549	0.09457
19	0.00553	2.10670	0.00394	-0.00699	2.10188	0.04814	0.00324
20	1.56527	-0.10670	0.00394	1.57779	-0.10188	-0.04814	0.00324
21	0.0	1.32570	0.70071	0.0	1.32570	0.0	0.70071
22	1.57080	0.67430	0.70071	1.57080	0.67430	0.0	0.70071
23	6.59193	-3.72830	30.2541	10.0988	-14.1074	4.71778	8.98184
24	57.8416	-36.4622	4408.67	98.4497	-156.648	54.6298	1556.37
25	6.59193	-2.72830	30.2541	10.0988	-13.1074	4.71778	8.98184
26	57.0983	-34.2622	4514.05	98.4497	-156.648	55.6298	1556.37
27	0.56886	0.48189	2.03722	0.56886	0.59752	-0.02891	1.96702
28	0.32843	-0.53007	12.8456	0.32843	1.02130	-0.38784	0.21008
29	3.98435	2.39920	78.7506	0.54779	3.68791	1.28871	16.7486
30	0.54968	0.19055	1.12040	0.24489	0.71140	0.57871	0.33895

3.4.

Вариант	Первая производная			Вторая производная
	левосторонняя	правосторонняя	центральная	
1	0.5	0.366025	0.433013	-0.133975
2	0.5	0.366025	0.433013	-0.133975
3	0.575364	0.446287	0.510826	-0.258154
4	1.16232	1.28456	1.22344	1.22242
5	0.414214	0.317837	0.366025	-0.096376
6	1.00679	1.05079	1.02879	0.220025
7	-1.00679	-1.05079	-1.02879	-0.220025
8	0.785398	0.321751	0.553574	-0.463648
9	-0.785398	-0.321751	-0.553574	0.463648
10	1.72409	1.31205	1.51807	-0.824086
11	0.325439	0.060870	0.193155	-0.529139
12	1.39514	1.40344	1.40012	0.016598
13	-5.49885	-2.02839	-3.76362	11.5682
14	1.40547	1.28768	1.34657	-0.117783
15	2.35211	2.65147	2.50179	1.49680
16	1.41421	1.31784	1.36603	-0.096376

Окончание табл.

Вариант	Первая производная			Вторая производная
	левосторонняя	правосторонняя	центральная	
17	2.00679	2.05079	2.02879	0.220025
18	-0.006790	-0.050795	-0.028792	-0.220025
19	1.78540	1.32175	1.55357	-0.463648
20	0.214602	0.678249	0.446426	0.463648
21	-0.333333	-0.2	-0.266667	0.266667
22	-0.921202	-0.597895	-0.759549	1.61653
23	0.666667	0.800000	0.733333	0.266667
24	1.67880	2.40210	2.04045	3.61653
25	1.23205	1.26795	1.25000	0.035898
26	0.13398	-1.0	-0.433013	-1.13397
27	0.606531	1.64872	1.12763	2.08438
28	0.949179	2.08635	1.51777	5.68587
29	0.50	0.366025	0.433013	-0.133975
30	0.366025	0.133975	0.25	-0.232051

3.5. Дается точное значение интеграла с шестью значащими цифрами.

1	-0.0591223	11	0.0626407	21	0.174347
2	0.0706994	12	-0.00629484	22	0.4
3	-0.122449	13	-0.203636	23	0.417972
4	1.77733	14	0.108372	24	15.3058
5	0.104471	15	0.0232347	25	0.0
6	-0.041303	16	0.0101483	26	186.625
7	1.85742	17	0.0158274	27	17.2998
8	0.785398	18	0.0	28	701.285
9	0.183862	19	0.0423871	29	8.85296
10	0.145410	20	0.531219	30	6165.78

Глава 4

ЧИСЛЕННЫЕ МЕТОДЫ РЕШЕНИЯ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

В настоящем разделе рассматриваются различные методы решения задачи Коши и краевой задачи для обыкновенных дифференциальных уравнений. Подробно изучаются метод Эйлера и ошибки, возникающие при реализации метода на ЭВМ, семейство одношаговых методов Рунге — Кутты, многошаговые методы, методы решения начальной задачи для дифференциальных уравнений с запаздывающим аргументом. При решении краевой задачи применяются метод стрельбы и разностный метод.

4.1. Численные методы решения задачи Коши

4.1.1. Задача Коши для одного обыкновенного дифференциального уравнения

Рассматривается задача Коши для одного дифференциального уравнения первого порядка, разрешенного относительно производной

$$y' = f(x, y), \quad y(x_0) = y_0. \quad (4.1)$$

Требуется найти решение на отрезке $[a, b]$, здесь $x_0 = a$.

Введем разностную сетку на отрезке $[a, b]$:

$$\Omega^{(k)} = \{x_k = x_0 + hk\}, \quad h = |b - a| / N, \quad h = 0, 1, \dots, N.$$

Точки x_k называются *узлами* разностной сетки, расстояние h между узлами — *шагом* разностной сетки, а совокупность заданных в узлах сетки значений какой-либо величины называется *сеточной функцией* $y(h) = \{y_k, k = 0, 1, \dots, N\}$.

Приближенное решение задачи Коши (4.1) будем искать численно в виде сеточной функции $y^{(h)}$. Для оценки погрешности приближенного численного решения $y^{(h)}$ будем рассматривать это решение как элемент $(N + 1)$ -мерного линейного векторного пространства с какой-либо нормой. В качестве погрешности

решения принимается норма элемента этого пространства $\delta^{(h)} \quad y^{(h)} \quad [y]^{(h)}$, где $[y]^{(h)}$ – точное решение задачи (4.1) в узлах расчетной сетки. Таким образом, погрешность $\varepsilon_h \quad \|\delta^{(h)}\|$.

4.1.2. Одношаговые методы

Метод Эйлера. Явный метод Эйлера играет важную роль в теории численных методов решения обыкновенных дифференциальных уравнений (ОДУ), хотя и не часто используется в практических расчетах из-за невысокой точности. Вывод расчетных соотношений для этого метода может быть произведен несколькими способами: с помощью геометрической интерпретации, с использованием разложения в ряд Тейлора, конечно-разностным методом (с помощью разностной аппроксимации производной), квадратурным способом (с использованием эквивалентного интегрального уравнения).

Рассмотрим вывод соотношений метода Эйлера геометрическим способом (рис. 4.1). Пусть решение в узле x_0 известно из начальных условий. Рассмотрим процедуру получения решения в узле x_1 .

График функции $y^{(h)}$, которая является решением задачи Коши (4.1), представляет собой гладкую кривую, проходящую через точку (x_0, y_0) согласно условию $y(x_0) = y_0$, и имеет в этой точке касательную. Тангенс угла наклона касательной к оси Ox равен значению производной от решения в точке x_0 и равен значению правой части дифференциального уравнения в точке (x_0, y_0) согласно выражению $y'(x_0) = f(x_0, y_0)$. В случае небольшого шага разностной сетки h график функции и график касательной не успевают существенно разойтись друг от друга и можно в качестве значения решения в узле x_1 принять значение касательной y_1 вместо значения неизвестного точного решения $y_{1 \text{ ист}}$. При этом

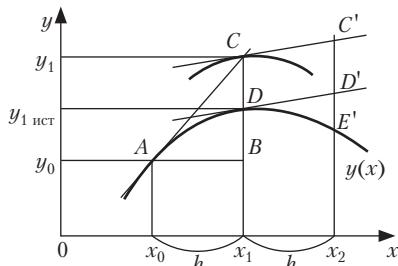


Рис. 4.1

допускается погрешность $|y_1 - y_{1 \text{ ист}}|$, геометрически представленная отрезком CD на рис. 4.1. Из прямоугольного треугольника ABC находим $CB = BA \operatorname{tg}(< CAB)$ или $\Delta y = hy'(x_0)$. Учитывая, что $\Delta y = y_1 - y_0$, и заменяя производную $y'(x_0)$ на правую часть дифференциального уравнения, получаем соотношение

$$y_1 = y_0 + hf(x_0, y_0).$$

Считая теперь точку (x_1, y_1) начальной и повторяя все предыдущие рассуждения, получим значение y_2 в узле x_2 . Переход к произвольным индексам дает формулу метода Эйлера:

$$y_{k+1} = y_k + hf(x_k, y_k). \quad (4.2)$$

Погрешность метода Эйлера. На каждом шаге метода Эйлера допускается *локальная* погрешность по отношению к точному решению, график которого проходит через крайнюю левую точку отрезка. Геометрически локальная погрешность изображается отрезком CD на первом шаге, $C'D'$ — на втором и т.д. Кроме того, на каждом шаге, начиная со второго, накапливается *глобальная* погрешность, представляющая собой разность между численным решением и точным решением исходной начальной задачи (не локальной). Глобальная погрешность на втором шаге изображена отрезком $C'E'$ на рис. 4.1.

Локальная погрешность на каждом шаге выражается соотношением

$$\varepsilon_k^h = \frac{y''(\xi)}{2} h^2,$$

в котором $\xi \in [x_{k-1}, x_k]$. Глобальная погрешность метода Эйлера $\varepsilon_k^h \sim Ch$ в окрестности $h = 0$ ведет себя как линейная функция, и, следовательно, метод Эйлера имеет первый порядок точности относительно шага h .

Неявный метод Эйлера. Если на правой границе интервала использовать точное значение производной от решения (т.е. тангенса угла наклона касательной), то получается неявный метод Эйлера первого порядка точности:

$$y_{k+1} = y_k + hf(x_{k+1}, y_{k+1}), \quad x_{k+1} = x_1 + h. \quad (4.3)$$

В общем случае нелинейное относительно y_{k+1} уравнение (4.3) численно решается с помощью одного из методов, описанных в разделе 2, например методом Ньютона или его модификациями.

Метод Эйлера – Коши. В данном методе на каждом интервале расчет проводится в два этапа. На первом (этап *прогноза*) определяется приближенное решение на правом конце интервала по методу Эйлера, на втором (этап *коррекции*) уточняется значение решения на правом конце с использованием полусуммы тангенсов углов наклона на концах интервала:

$$\begin{aligned}\tilde{y}_{k+1} &= y_k + hf(x_k, y_k), \\ y_{k+1} &= y_k + \frac{h(f(x_k, y_k) + \tilde{f}(x_{k+1}, \tilde{y}_{k+1}))}{2}, \\ x_{k+1} &= x_k + h.\end{aligned}\quad (4.4)$$

Этот метод имеет второй порядок точности.

Неявный метод Эйлера – Коши. Если на правой границе интервала использовать точное значение производной к решению (т.е. тангенса угла наклона касательной), то получается неявный метод Эйлера – Коши (метод трапеций) второго порядка точности:

$$y_{k+1} = y_k + \frac{h(f(x_k, y_k) + f(x_{k+1}, y_{k+1}))}{2}, \quad x_{k+1} = x_k + h. \quad (4.5)$$

Метод Эйлера – Коши с итерационной обработкой. Комбинация формул (4.3), (4.4) и (4.5) дает метод формально второго порядка точности, но более точного в смысле абсолютной величины погрешности приближенного решения, чем метод Эйлера – Коши (4.4):

$$\begin{aligned}y_{k+1}^{(0)} &= y_k + hf(x_k, y_k), \\ y_{k+1}^{(i)} &= y_k + \frac{h(f(x_k, y_k) + f(x_{k+1}, y_{k+1}^{(i-1)}))}{2}, \\ x_{k+1} &= x_k + h.\end{aligned}\quad (4.6)$$

В формуле (4.6) правые верхние индексы в круглых скобках обозначают номер итерации, при этом начальное приближение $y_{k+1}^{(0)}$ определяется по методу Эйлера. Метод Эйлера – Коши с итерационной обработкой представляет собой реализацию метода простой итерации для решения нелинейного уравнения (4.5) в неявном методе Эйлера – Коши. Выполнять простые итерации до полной сходимости нет смысла, поэтому рекомендуется выполнять 3–4 итерации.

Первый улучшенный метод Эйлера. Данный метод использует расчет приближенного значения производной от решения в точке на середине расчетного интервала. Значение производной в середине получают применением явного метода Эйлера на половинном шаге по x :

$$\begin{aligned} x_{k+1/2} &= x_k + \frac{h}{2}, & y_{k+1/2} &= y_k + \frac{h}{2} f(x_k, y_k), \\ x_{k+1} &= x_k + h, & y_{k+1} &= y_k + h f(x_{k+1/2}, y_{k+1/2}). \end{aligned} \quad (4.7)$$

Данная модификация метода Эйлера имеет второй порядок точности.

Метод Рунге – Кутты. Все рассмотренные выше явные методы являются вариантами методов Рунге – Кутты. Семейство явных методов Рунге – Кутты p -го порядка записывается в виде совокупности формул:

$$\begin{aligned} y_{k+1} &= y_k + \Delta y_k, & y_{k+1} &= \sum_{i=1}^p c_i K_i^k, \\ K_i^k &= h f \left(x_k + a_i h, \quad y_k + h \sum_{j=1}^{i-1} b_{ij} K_j^k \right), & i &= 2, 3, \dots, p. \end{aligned} \quad (4.8)$$

Параметры a_i , b_{ij} , c_i подбираются так, чтобы значение y_{k+1} рассчитанное по соотношению (4.8), совпадало со значением разложения в точке x_{k+1} точного решения в ряд Тейлора с погрешностью $O(h^{p+1})$.

Метод Рунге – Кутты третьего порядка точности. Один из методов Рунге – Кутты третьего порядка $\begin{pmatrix} p & 3, a_1 & 0, a_2 & \frac{1}{3}, \end{pmatrix}$

$a_3 = \frac{2}{3}$, $b_{21} = \frac{1}{3}$, $b_{31} = 0$, $b_{32} = \frac{2}{3}$, $c_1 = \frac{1}{4}$, $c_2 = 0$, $c_3 = \frac{3}{4}$ имеет вид

$$\begin{aligned} y_{k+1} &= y_k + \Delta y_k, & \Delta y_k &= \frac{1}{4} (K_1^k + 3K_3^k), \\ K_1^k &= h f(x_k, y_k), \\ K_2^k &= h f \left(x_k + \frac{1}{3} h, \quad y_k + \frac{1}{3} K_1^k \right), \\ K_3^k &= h f \left(x_k + \frac{2}{3} h, \quad y_k + \frac{2}{3} K_2^k \right). \end{aligned} \quad (4.9)$$

Метод Рунге – Кутты четвертого порядка точности. Метод Рунге – Кутты четвертого порядка $\begin{pmatrix} p & 4, a_1 & 0, a_2 & \frac{1}{2}, a_3 & \frac{1}{2}, \\ a_4 = 1, b_{21} = \frac{1}{2}, b_{31} = 0, b_{32} = \frac{1}{2}, b_{41} = 0, b_{42} = 0, b_{43} = \frac{1}{2}, c_1 = \frac{1}{6}, c_2 = \frac{1}{3}, \\ c_3 = \frac{1}{3}, c_3 = \frac{1}{6} \end{pmatrix}$ является одним из самых широко используемых методов для решения задачи Коши: $y_{k+1} = y_k + \Delta y_k$,

$$\begin{aligned} \Delta y_k &= \frac{1}{6} (K_1^k + 2K_2^k + 2K_3^k + K_4^k), \\ K_1^k &= hf(x_k, y_k), \\ K_2^k &= hf\left(x_k + \frac{1}{2}h, y_k + \frac{1}{2}K_1^k\right), \\ K_3^k &= hf\left(x_k + \frac{2}{3}h, y_k + \frac{1}{2}K_2^k\right), \\ K_4^k &= hf(x_k + h, y_k + K_3^k). \end{aligned} \quad (4.10)$$

Контроль точности с выбором шага h . Основным способом контроля точности получаемого численного решения при решении задачи Коши является применение методов, основанных на принципе Рунге – Ромберга – Ричардсона.

Пусть y^h – решение задачи Коши (4.1), полученное методом Рунге – Кутты p -го порядка точности с шагом h в точке $x + 2h$. Пусть y^{2h} – решение той же задачи в точке $x + 2h$, полученное тем же методом, но с шагом $2h$. Тогда выражение

$$\tilde{y} = y^h + \frac{y^h - y^{2h}}{2^p - 1} \quad (4.11)$$

аппроксимирует точное решение $y(x + 2h)$ в точке $x + 2h$ с $(p + 1)$ -м порядком. Второе слагаемое в выражении (4.11) оценивает главный член в погрешности решения y^h , т.е. главный член $R^h = (y^h - y^{2h})/(2^p - 1)$. Контроль точности может быть организован следующим образом. Выбирается значение шага h и дважды рассчитывается решение в точке $(x + 2h)$: один раз с шагом h , другой раз – с шагом $2h$. Рассчитывается величина R^h и сравнивается с заданной погрешностью ϵ . Если $R^h < \epsilon$, то можно продолжать вычисления с тем же шагом, в противном случае необходимо вернуться к решению в точке x , уменьшить шаг h и повторить вычисления.

Вычислительная «стоимость» такого контроля достаточно велика, особенно для многостадийных методов. Поэтому можно использовать более грубый способ контроля правильности выбора шага h . В случае метода Рунге – Кутты четвертого порядка точности следует на каждом шаге h рассчитывать параметр

$$\Theta^k = \left| \frac{K_2^k - K_3^k}{K_1^k - K_2^k} \right|. \quad (4.12)$$

Если величина Θ^k порядка нескольких сотых единицы, то расчет продолжается с тем же шагом, если $\Theta^k > 0.1$, то шаг следует уменьшить, если же $\Theta^k < 0.01$, то шаг можно увеличить. Таким образом с помощью определения величин Θ^k или R^h можно организовать алгоритм выбора шага h для явного метода Рунге – Кутты.

4.1.3. Решение задачи Коши для системы обыкновенных дифференциальных уравнений

Рассматривается задача Коши для системы дифференциальных уравнений первого порядка, разрешенных относительно производной:

$$\begin{cases} y'_1 = f_1(x, y_1, y_2, \dots, y_n), \\ y'_2 = f_2(x, y_1, y_2, \dots, y_n), \\ \dots \\ y'_n = f_n(x, y_1, y_2, \dots, y_n), \\ y_1(x_0) = y_{01}, \\ y_2(x_0) = y_{02}, \\ \dots \\ y_n(x_0) = y_{0n}. \end{cases} \quad (4.13)$$

Система (4.13) в более компактном виде записывается в векторной форме следующим образом:

$$\begin{aligned} \mathbf{y}' &= \mathbf{F}(x, \mathbf{y}), \\ \mathbf{y}(x_0) &= \mathbf{y}_0. \end{aligned} \quad (4.14)$$

Здесь $\mathbf{y}(x) = (y_1, y_2, \dots, y_n)^T$ — вектор-столбец неизвестных функций, $\mathbf{F} = (f_1, f_2, \dots, f_n)^T$ — вектор-функция правых частей. К векторному дифференциальному уравнению (4.14) можно применить все методы, рассмотренные выше в данном разделе (благодаря линейной структуре всех рассмотренных методов). При

этом в формулах (4.2)–(4.11) все величины векторные, кроме переменной x и шага h .

Рассмотрим задачу Коши для системы двух ОДУ первого порядка, в которой уравнения записаны в развернутом виде

$$\begin{cases} y' = f(x, y, z), \\ z' = g(x, y, z), \\ y(x_0) = y_0, \\ z(x_0) = z_0. \end{cases} \quad (4.15)$$

Приведем формулы метода Рунге – Кутты четвертого порядка точности для решения системы (4.15):

$$\begin{aligned} y_{k+1} &= y_k + \Delta y_k, & z_{k+1} &= z_k + \Delta z_k, \\ \Delta y_k &= \frac{1}{6} (K_1^k + 2K_2^k + 2K_3^k + K_4^k), & \Delta z_k &= \frac{1}{6} (L_1^k + 2L_2^k + 2L_3^k + L_4^k), \\ K_1^k &= hf(x_k, y_k, z_k), \\ L_1^k &= hg(x_k, y_k, z_k), \\ K_2^k &= hf\left(x_k + \frac{1}{2}h, y_k + \frac{1}{2}K_1^k, z_k + \frac{1}{2}L_1^k\right), \\ L_2^k &= hg\left(x_k + \frac{1}{2}h, y_k + \frac{1}{2}K_1^k, z_k + \frac{1}{2}L_1^k\right), \\ K_3^k &= hf\left(x_k + \frac{2}{2}h, y_k + \frac{1}{2}K_2^k, z_k + \frac{1}{2}L_2^k\right), \\ L_3^k &= hg\left(x_k + \frac{2}{2}h, y_k + \frac{1}{2}K_2^k, z_k + \frac{1}{2}L_2^k\right), \\ K_4^k &= hf(x_k + h, y_k + K_3^k, z_k + L_3^k), \\ L_4^k &= hg(x_k + h, y_k + K_3^k, z_k + L_3^k). \end{aligned} \quad (4.16)$$

Контроль правильности выбора шага h в случае использования метода Рунге – Кутты четвертого порядка точности для системы (4.15) может быть организован с помощью вычисления на каждом шаге h параметров

$$\Theta_1^k = \left| \frac{K_2^k - K_3^k}{K_1^k - K_2^k} \right|, \quad \Theta_2^k = \left| \frac{L_2^k - L_3^k}{L_1^k - L_2^k} \right|. \quad (4.17)$$

Если величины $\Theta_i^k, i = 1, 2$, порядка нескольких сотых единицы, то расчет продолжается с тем же шагом, если $\Theta_i^k > 0.1$, то шаг следует уменьшить, если же $\Theta_i^k < 0.01$, то шаг можно увеличить.

4.1.4. Решение задачи Коши для ОДУ второго и более высоких порядков

Задача Коши для ОДУ n -го порядка ставится следующим образом:

$$\begin{aligned} y^{(n)} &= f(x, y, y', y'', \dots, y^{(n-1)}), \\ y(x_0) &= y_0, \\ y'(x_0) &= y_{01}, \\ y''(x_0) &= y_{02}, \\ &\dots \\ y^{(n-1)}(x_0) &= y_{0(n-1)}; \end{aligned} \tag{4.18}$$

здесь $y^{(m)} = \frac{d^m y}{dx^m}$ — производная m -го порядка от решения, $m = 1, 2, \dots, n$.

Основной прием, используемый при решении задач типа (4.18), заключается в введении новых переменных и сведении задачи (4.8) для ОДУ высокого порядка к решению системы ОДУ первого порядка (4.13).

Введем новые переменные

$$\begin{aligned} z_1 &= y', \\ z_2 &= y'', \\ &\dots \\ z_{n-1} &= y^{(n-1)}. \end{aligned}$$

С ними задачу (4.8) можно переписать в виде системы n ОДУ первого порядка:

$$\begin{cases} y' = z_1, \\ z'_1 = z_2, \\ z'_2 = z_3, \\ \dots \\ z'_{n-2} = z_{n-1}, \\ z'_{n-1} = f(x, y, z_1, \dots, z_{n-1}), \\ y(x_0) = y_0, \\ z_1(x_0) = y_{01}, \\ z_2(x_0) = y_{02}, \end{cases} \tag{4.19}$$

Полученная система, состоящая из n ОДУ первого порядка с соответствующими начальными условиями, решается любым из описанных методов. Пусть необходимо решить задачу Коши для ОДУ второго порядка:

$$\begin{aligned} y'' &= f(x, y, y'), \\ y(x_0) &= y_0, \\ y'(x_0) &= y_{01}. \end{aligned} \quad (4.20)$$

Путем введения замены $z = y'$ приведем (4.20) к системе

$$\begin{cases} y' = z, \\ z' = f(x, y, z), \\ y(x_0) = y_0, \\ z(x_0) = y_{01}, \end{cases} \quad (4.21)$$

которую можно решить, например, с использованием метода Рунге – Кутты четвертого порядка точности (4.16).

Пример 4.1. Явным методом Эйлера с шагом $h = 0.1$ получить численное решение дифференциального уравнения $y' = (y + x)^2$ с начальными условиями $y(0) = 0$ на интервале $[0, 0.5]$. Численное решение сравнить с точным решением $y = \operatorname{tg}(x) - x$.

Решение. Исходя из начальной точки $x_0 = 0, y_0 = 0$, рассчитаем значение y_1 в узле $x_1 = 0.1$ по формулам (4.2):

$$y_1 = y_0 + hf(x_0, y_0) = 0 + 0.1 \cdot (0 + 0)^2 = 0.$$

Аналогично получим решение в следующем узле $x_2 = 0.2$:

$$y_2 = y_1 + hf(x_1, y_1) = 0 + 0.1 \cdot (0 + 0.1)^2 = 0.001.$$

Продолжим вычисления и, введя обозначения $\Delta y_k = hf(x_0, y_0)$ и $\varepsilon_k = |y_{\text{ист}}(x_k) - y_k|$, где $y_{\text{ист}}(x_k)$ — точное решение в узловых точках, получаемые результаты занесем в таблицу 4.1.

Таблица 4.1

	x	y	Δy_k	$y_{\text{ист}}$	ε_k
0	0.0000000000	0.0000000000	0.0000000000	0.0000000000	0.0000
1	0.1000000000	0.0000000000	0.0010000000	0.000334672	0.3347E-03*
2	0.2000000000	0.0010000000	0.0040401000	0.002710036	0.1710E-02
3	0.3000000000	0.0050401000	0.009304946	0.009336250	0.4296E-02
4	0.4000000000	0.014345046	0.017168182	0.022793219	0.8447E-02
5	0.5000000000	0.031513228		0.046302490	0.1479E-01

*Здесь и далее запись вида $0.3347\text{E}-03$ означает $0.3347 \cdot 10^{-3}$.

Решением задачи является табличная функция — таблица 4.2 (оставлены 5 значащих цифр в каждом числе).

Таблица 4.2

k	0	1	2	3	4	5
x_k	0.00000	0.1000	0.200000	0.3000000	0.400000	0.500000
y_k	0.00000	0.0000	0.001000	0.0050401	0.014345	0.031513

Пример 4.5. Решить задачу из примера 4.1 методом Эйлера — Коши (4.4).

Решение. Исходя из начальных значений $x_0 = 0$, $y_0 = 0$, рассчитаем значение y_1 в узле $x_1 = 0.1$ по формулам (4.4):

$$\tilde{y}_1 = y_0 + hf(x_0, y_0) = 0 + 0.1 \cdot (0 + 0)^2 = 0,$$

$$f(x_1, \tilde{y}_1) = (0 + 0.1)^2 = 0.01,$$

$$y_1 = y_0 + 0.5h(f(x_0, y_0) + f(x_1, \tilde{y}_1)) = 0 + 0.5 \cdot 0.1 \cdot (0 + 0.01) = 0.0005.$$

Аналогично получим решение в остальных узлах. Продолжим вычисления и, введя обозначение $\Delta y_k = 0.5h(f(x_k, y_k) + f(x_{k+1}, \tilde{y}_{k+1}))$, получаемые результаты занесем в таблицу 4.3.

Таблица 4.3

k	x_k	y_k	\tilde{y}_k	Δy_k	$y_{\text{ист}}$	ε_k
0	0.0	0.0000000000		0.000500000	0.0000000000	0.0000000000
1	0.1	0.0000000000	0.00000	0.002535327	0.000334672	0.1653E-03
2	0.2	0.003035327	1.510025E-003	0.006778459	0.002710036	0.3253E-03
3	0.3	0.009813786	7.157661E-003	0.013594561	0.009336250	0.4775E-03
4	0.4	0.023408346	1.941224E-002	0.023615954	0.022793219	0.6151E-03
5	0.5	0.047024301	4.133581E-002		0.046302490	0.7218E-03

Решением задачи является табличная функция — таблица 4.4 (оставлены 5 значащих цифр в каждом числе).

Таблица 4.4

k	0	1	2	3	4	5
x_k	0.00000	0.100000	0.2000000	0.3000000	0.4000000	0.500000
y_k	0.00000	0.000500	0.0030353	0.0098138	0.023408	0.047024

Пример 4.3. Решить задачу из примера 4.1 первым улучшенным методом Эйлера (4.7).

Решение. Стартуем из начальной точки $x_0 = 0$, $y_0 = 0$ и рассчитаем значение $y_{1/2}$ в узле $x_{1/2} = x_0 + h/2 = 0.05$ по формулам (4.4):

$$y_{1/2} = y_0 + \frac{h}{2} f(x_0, y_0) = 0 + \frac{0.1}{2} \cdot (0 + 0)^2 = 0.$$

Затем определим величину правой части (величину производной от решения) в середине интервала $[x_0, x_1]$:

$$f(x_{1/2}, y_{1/2}) = (0 + 0.05)^2 = 0.0025.$$

Окончательно рассчитаем значение функции в узле x_1 :

$$y_1 = y_0 + hf(x_{1/2}, y_{1/2}) = 0 + 0.1 \cdot 0.0025 = 0.00025.$$

Аналогично получим решение в остальных узлах. Продолжим вычисления и, введя обозначение $\Delta y_k = hf(x_{k+1/2}, y_{k+1/2})$, получаемые результаты занесем в таблицу 4.5.

Таблица 4.5

k	x_k	y_k	$y_{k+1/2}$	Δy_k	$y_{\text{ист}}$	ε_k
0	0.0	0.000000000	0.000000000	0.000250000	0.000000000	0.000000000
1	0.1	0.000250000	0.0007525031	0.002272632	0.000334672	0.8467E-04
2	0.2	0.002522632	0.0045734025	0.006480762	0.002710036	0.1874E-03
3	0.3	0.009003393	0.0137775483	0.013233410	0.009336250	0.3329E-03
4	0.4	0.022236804	0.0311509998	0.023150628	0.022793219	0.5564E-03
5	0.5	0.045387432			0.046302490	0.9151E-03

Решением задачи является табличная функция — таблица 4.6 (оставлены 5 значащих цифр в каждом числе).

Таблица 4.6

k	0	1	2	3	4	5
x_k	0.00000	0.10000	0.20000	0.30000	0.40000	0.50000
y_k	0.00000	0.000250	0.0025226	0.0090033	0.0227932	0.045387

Пример 4.4. Решить задачу из примера 4.1 методом Рунге—Кутты четвертого порядка (4.10).

Решение. Вычислим значения вспомогательных величин:

$$K_1^0 = hf(x_0, y_0) = 0.1 \cdot (0 + 0)^2 = 0;$$

$$y_0^1 = y_0 + \frac{1}{2} K_1^0 = 0 + \frac{1}{2} \cdot 0 = 0,$$

$$K_2^0 = hf\left(x_0 + \frac{1}{2}h, y_0 + \frac{1}{2}K_1^0\right) = 0.1 \cdot \left(0 + \frac{1}{2} \cdot 0 + 0 + \frac{1}{2} \cdot 0.1\right)^2 = 0.00025;$$

$$y_0^2 = y_0 + \frac{1}{2} K_2^0 = 0 + \frac{1}{2} \cdot 0.00025 = 0.000125,$$

$$K_3^0 = hf\left(x_0 + \frac{2}{3}h, y_0 + \frac{1}{3}K_2^0\right) = 0.1 \cdot \left(0 + \frac{1}{3} \cdot 0.00025 + 0 + \frac{1}{3} \cdot 0.1\right)^2 = 0.000251251;$$

$$y_0^3 = y_0 + K_3^0 = 0 + 0.000251251 = 0.000251251,$$

$$K_4^0 = hf(x_0 + h, y_0 + K_3^0) = 0.1 \cdot (0 + 0.000251251 + 0 + 0.1)^2 = 0.001005031.$$

Найдем приращение функции на первом интервале

$$\begin{aligned}\Delta y_0 &= \frac{1}{6} (K_1^0 + 2K_2^0 + 2K_3^0 + K_4^0) = \\ &= \frac{1}{6} (0 + 2 \cdot 0.00025 + 2 \cdot 0.000251251 + 0.001005031) = 0.000334588\end{aligned}$$

и значение функции в первом узле

$$y_1 = y_0 + \Delta y_0 = 0 + 0.000334588 = 0.000334588.$$

Аналогично получим решение в остальных узлах — таблица 4.7.

Таблица 4.7

k	x_k	y_k^i	K_i^k	$\mathcal{O}y_k$	Θ^k	$y_{\text{ист}}$	ε_k
0/1	0.0	0.0000000	0.000000000			0.000000	0.0000000
0/2	0.05	0.0000000	0.000250000				
0/3	0.05	0.0001250	0.000251252				
0/4	0.1	0.00025125	0.001005031	0.000334589	0.005006		
1/1	0.1	0.000334589	0.001006703			0.00033467	0.8301E-07
1/2	0.15	0.000837941	0.002275208				
1/3	0.15	0.001472193	0.002294383				
1/4	0.2	0.002628972	0.004105850	0.002375289	0.015116		
2/1	0.2	0.002709878	0.004109129			0.002710036	0.1573E-06
2/2	0.25	0.004764443	0.006490492				
2/3	0.25	0.005955124	0.006551303				
2/4	0.3	0.009261181	0.009564248	0.006626161	0.025535		
3/1	0.3	0.009336039	0.009568879			0.009336250	0.2103E-06
3/2	0.35	0.014120479	0.013258372				
3/3	0.35	0.015965225	0.013393055				
3/4	0.4	0.022729094	0.017869989	0.013456954	0.036504		
4/1	0.4	0.022792993	0.017875391			0.022793219	0.2259E-06
4/2	0.45	0.031730689	0.023206446				
4/3	0.45	0.034396216	0.023463969				
4/4	0.5	0.046256962	0.029839667	0.023509315	0.048306		
5	0.5	0.046302308				0.046302490	0.1823E-06

Решением задачи является табличная функция — таблица 4.8 (оставлены 7 значащих цифр в каждом числе).

Таблица 4.8

k	0	1	2	3	4	5
x_k	0.00000	0.1000	0.200000	0.3000000	0.400000	0.500000
y_k	0.00000	0.000334589	0.002709878	0.009336039	0.02279299	0.04630231

Пример 4.5. На интервале $[0, 1]$ с шагом $h = 0.2$ решить задачу Коши методом Рунге — Кутты четвертого порядка:

$$\begin{aligned} (x^2 + 1)y'' &= 2xy', \\ y(0) &= 1, \\ y'(0) &= 3. \end{aligned}$$

Численное решение сравнить с аналитическим решением

$$y_{\text{ист}} = x^3 + 3x + 1.$$

Решение. Аналогично (4.18)–(4.21) введением новой переменной $z = y'$ решение исходной начальной задачи для дифференциального уравнения второго порядка сводится к решению системы двух дифференциальных уравнений первого порядка:

$$\begin{cases} y' = z, \\ z' = \frac{2xz}{x^2 + 1}, \\ y(0) = 1, \\ z(0) = 3. \end{cases}$$

Данную систему решим методом Рунге — Кутты с использованием формул (4.16). Вычислим значения вспомогательных величин:

$$K_1^0 = hf(x_0, y_0, z_0) = hz_0 = 0.2 \cdot 3 = 0.6;$$

$$L_1^0 = hg(x_0, y_0, z_0) = h \frac{2x_0 z_0}{x_0^2 + 1} = 0.2 \cdot \frac{2 \cdot 0 \cdot 3}{0^2 + 1} = 0;$$

$$K_2^0 = hf\left(x_0 + \frac{1}{2}h, y_0 + K_1^0, z_0 + L_1^0\right) = 0.2 \cdot \left(3 + \frac{1}{2} \cdot 0\right) = 0.6;$$

$$\begin{aligned} L_2^0 &= hg\left(x_0 + \frac{1}{2}h, y_0 + \frac{1}{2}K_1^0, z_0 + \frac{1}{2}L_1^0\right) = 0.2 \cdot \frac{2 \cdot (0 + 0.1)(3 + (1/2) \cdot 0)}{(0 + 0.1)^2 + 1} = \\ &= 0.11881188; \end{aligned}$$

$$\begin{aligned} K_3^0 &= hf\left(x_0 + \frac{2}{2}h, y_0 + \frac{1}{2}K_2^0, z_0 + \frac{1}{2}L_2^0\right) = 0.2 \cdot \left(3 + \frac{1}{2} \cdot 0.1881188\right) = \\ &= 0.611881188; \end{aligned}$$

$$\begin{aligned}
 L_3^0 &= hg \left(x_0 + \frac{2}{2}h, y_0 + \frac{1}{2}K_2^0, z_0 + \frac{1}{2}L_2^0 \right) = \\
 &= 0.2 \cdot \frac{2 \cdot (0+0.1)(3+(1/2) \cdot 0.11881188)}{(0+0.1)^2 + 1} = 0.121164592; \\
 K_4^0 &= hf(x_0 + h, y_0 + K_3^0, z_0 + L_3^0) = 0.2 \cdot (3 + 0.12116459) = 0.62423292; \\
 L_4^0 &= hg(x_0 + h, y_0 + K_3^0, z_0 + L_3^0) = 0.2 \cdot \frac{(0+0.2)(3+0.121164592)}{(0+0.2)^2 + 1} = \\
 &= 0.240089584.
 \end{aligned}$$

Найдем приращения функций на первом интервале:

$$\begin{aligned}
 \Delta y_0 &= \frac{1}{6}(K_1^0 + 2K_2^0 + 2K_3^0 + K_4^0) = \\
 &= \frac{1}{6}(0.6 + 2 \cdot 0.6 + 2 \cdot 0.611881188 + 0.62423292) = 0.607999216, \\
 \Delta z_0 &= \frac{1}{6}(L_1^0 + 2L_2^0 + 2L_3^0 + L_4^0) = \\
 &= \frac{1}{6}(0.0 + 2 \cdot 0.11881188 + 2 \cdot 0.121164592 + 0.240089584) = 0.1200071
 \end{aligned}$$

и значения функций в первом узле

$$\begin{aligned}
 y_1 &= y_0 + \Delta y_0 = 1 + 0.607999216 = 1.607999216, \\
 z_1 &= z_0 + \Delta z_0 = 3 + 0.1200071 = 3.1200071.
 \end{aligned}$$

Аналогично получим решения в остальных узлах, результаты вычислений занесем в таблицу 4.9.

Таблица 4.9

k	x_k	y_k	z_k	Δy_k	Δz_k	$y_{\text{ист}}$	ε_k
0	0.0	1.0000000	3.000000000	0.607999216	0.1200	1.000000000	0.00000
1	0.2	1.607999216	3.120007088	0.655995430	0.3600	1.607999216	0.784E-06
2	0.4	2.263994646	3.480019051	0.751991317	0.6000	2.263994646	0.535E-05
3	0.6	3.015985963	4.080024218	0.895987662	0.8400	3.015985963	0.140E-04
4	0.8	3.911973624	4.920018746	1.087984366	1.080	3.911973624	0.264E-04
5	1.0	4.999957990	6.000004180			5.000000000	0.420E-04

Решением задачи является табличная функция — таблица 4.10 (оставлены 5 значащих цифр в каждом числе).

Таблица 4.10

k	0	1	2	3	4	5
x_k	0.00000	0.200000	0.4000000	0.6000000	0.8000000	1.000000
y_k	1.0000000	1.607999216	2.263994646	3.015985963	3.911973624	4.99995799

4.1.5. Решение дифференциальных уравнений с запаздывающим аргументом

Многие процессы в живой и неживой природе описываются моделями, представленными дифференциальными уравнениями с *запаздывающими аргументами*. Наиболее часто такие модели используют при исследовании динамики развития популяций, процесса кроветворения, динамики различных автогенераторов, механизмов изменения рыночной конъюнктуры и т.п. Решение подобных уравнений обладает определенной спецификой.

Рассмотрим для простоты случай одного дифференциального уравнения с единственным запаздывающим аргументом a :

$$\begin{aligned} y' &= f(x, y(x), y(x-a)), \\ y(x_0) &= y_0. \end{aligned} \quad (4.22)$$

Пусть имеется решение в точке $y_k = y(x_k)$. Опишем процедуру нахождения решения в точке $x_k = x_k + h$ модифицированным методом Эйлера (4.7) второго порядка точности. В этом методе надо использовать значение решения в точке x_k и предварительное решение в точке $x_{k+1/2} = x_k + h/2$. Соответственно от этих точек следует брать запаздывание a , т.е. необходимо найти значение решения в точках $(x_k - a)$, $(x_k + h/2 - a)$.

Для примера опишем процедуру определения значения $y(x_k - a)$. Если $(x_k - a)$ лежит левее начальной точки x_0 , то $y(x_k - a)$ определяется из начальных условий — в этом случае должно быть задано поведение решения на интервале левее точки x_0 , достаточном для определения значения в точке $(x_k - a)$. Если $(x_k - a)$ совпадает с одним из узлов правее x_0 , то $y(x_k - a)$ принимает значение функции в этом узле. Если величина $(x_k - a)$ не совпадает ни с одним узловым значением x_m , $m = 0, 1, 2, \dots$, то она лежит внутри некоторого отрезка $[x_j, x_{j+1}]$ и можно по значениям y в трех узлах, например в x_{j-1}, x_j, x_{j+1} построить интерполяционный многочлен $P_2(x)$ для определения приближенного значения $y(x_k - a) \approx P_2(x_k - a)$. Таким образом, схема расчета значения решения в новой точке для системы (4.22) будет выглядеть следующим образом:

$$\begin{aligned} y_{k+1/2} &= y_k + \frac{h}{2} f(x_k, y_k, y(x_k - a)), \\ x_{k+1} &= x_k + h, \\ x_{k+1/2} &= x_k + h/2, \\ y_{k+1} &= y_k + h f(x_{k+1/2}, y_{k+1/2}, y(x_{k+1/2} - a)). \end{aligned} \quad (4.23)$$

Пример 4.6. Улучшенным методом Эйлера с шагом $h = 0.1$ получить численное решение дифференциального уравнения

$$y' = A_1 y(x) \left(1 - \frac{y(x - A_2)}{A_3} \right)$$

с начальными условиями $y(0) = 2.0$ на интервале $[0, 4]$ с шагом $h = 0.4$ при $A_1 = 1.6$, $A_2 = 0.5$, $A_3 = 10$; здесь A_2 — константа, характеризующая запаздывание аргумента.

Данное уравнение может описывать динамику одновидовой популяции (в этом случае A_1 — коэффициент экспоненциального роста, A_3 — емкость среды обитания, A_2 — возраст производителей, x — время). Смысл модели в следующем: скорость роста популяции зависит не только от общей численности $y(x)$ в любой момент времени x , определяемой емкостью среды обитания A_3 , но и от количества взрослых особей в момент времени $(x - A_2)$. Данное уравнение может также описывать цикличность деловой активности на фондовом рынке.

Решение. Решение будем проводить с использованием формул (4.23). Значение решения в точке

$$x_0 - A_2 = 0.0 - 0.5 = -0.5$$

и во всех точках, лежащих левее точки x_0 , примем равным начальному значению $y_0 = 2.0$, т.е. $y(x_0 - A_2) = 2.0$. Определим величину функции в точке

$$x_{1/2} = x_0 + h / 2 = 0.0 + 0.1 = 0.1$$

по методу Эйлера:

$$\begin{aligned} y_{1/2} &= y_0 + \frac{h}{2} f(x_0, y_0, y(x_0 - A_2)) = 2.0 + 0.1 \cdot 1.6 \cdot 2.0 \cdot (1.0 - 2.0 / 10.0) = \\ &= 2.256. \end{aligned}$$

В середине первого шага считаем значение функции с запаздывающим аргументом

$$x_0 + h / 2 - A_2 = 0.0 + 0.1 - 0.5 = -0.4, \quad y(x_0 + h / 2 - A_2) = 2.0$$

и затем значение решения в точке 1:

$$\begin{aligned} y_1 &= y_0 + hf(x_{1/2}, y_{1/2}, y(x_{1/2} - A_2)) = \\ &= 2.0 + 0.2 \cdot 1.6 \cdot 2.256 \cdot (1.0 - 2.0 / 10.0) = 2.577536. \end{aligned}$$

Продолжая таким образом вычисления и используя квадратичную интерполяцию (например, с помощью многочлена Лагранжа) для нахождения значения функции с запаздывающим аргументом, когда значения $(x_k - A_2)$ или $(x_k + h/2 - A_2)$ будут лежать правее, чем точка x_0 ,

получим решения в последующих точках. Результаты вычислений занесены в таблицу 4.11, в которой для удобства использованы следующие обозначения:

$$\Delta \tilde{y}_k = \frac{h}{2} f(x_k, y_k, y(x_k - A_2)), \quad x_{k+1/2} = x_k + \frac{h}{2}, \quad y_{k+1/2} = y_k + \Delta \tilde{y}_k,$$

$$\tilde{x}_k = x_k + \frac{h}{2} - A_2, \quad \tilde{y}_k = y\left(x_k + \frac{h}{2} - A_2\right), \quad \Delta y_k = h f\left(x_k + \frac{h}{2}, y_k + \Delta \tilde{y}_k, \tilde{y}_k\right).$$

Таблица 4.11

k	x_k	y_k	$x_k - A_2$	$y(x_k - A_2)$	\tilde{y}_k	$x_{k+1/2}$	$y_{k+1/2}$	\tilde{x}_k	\tilde{y}_k	Δy_k
0	0.0	2.0	-0.5	2.0	0.256000	0.1	2.25600	-0.4	2.0	0.577536
1	0.2	2.57754	-0.3	2.0	0.329925	0.3	2.90746	-0.2	2.0	0.744310
2	0.4	3.32185	-0.1	2.0	0.425196	0.5	3.74704	0.0	2.0	0.959243
3	0.6	4.28109	0.1	2.26792	0.529627	0.7	4.81072	0.2	2.57754	1.142636
4	0.8	5.42372	0.3	2.92282	0.614154	0.9	6.03788	0.4	3.32185	1.290300
5	1.0	6.71402								

Решением задачи является табличная функция — таблица 4.12 (оставлены 6 значащих цифр).

Таблица 4.12

k	0	1	2	3	4	5
x_k	0.00000	0.200000	0.4000000	0.6000000	0.8000000	1.000000
y_k	2.0	2.57754	3.32185	4.28109	5.42372	6.71402

Замечание. Как правило, в отличие от примера 4.6, в задачах с запаздывающим аргументом интересуются поведением решения на достаточно больших интервалах времени. При этом выполняется от сотен до тысяч шагов по времени, что приводит к необходимости использования компьютера.

4.1.6. Многошаговые методы. Метод Адамса

Многошаговые методы решения задачи Коши характеризуются тем, что решение в текущем узле зависит от данных не в одном предыдущем узле, как это имеет место в одношаговых методах, а в нескольких предыдущих. Многие многошаговые методы различного порядка точности можно конструировать с помощью квадратурного способа, т.е. с использованием эквивалентного интегрального уравнения.

Решение дифференциального уравнения $y' = f(x, y)$ удовлетворяет интегральному соотношению

$$y_{k+1} = y_k + \int_{x_k}^{x_{k+1}} f(x, y(x)) dx, \quad (4.24)$$

Если решение задачи Коши получено в узлах вплоть до k -го, то можно аппроксимировать подынтегральную функцию, например интерполяционным многочленом какой-либо степени. Вычислив интеграл от построенного многочлена на отрезке $[x_k, x_{k+1}]$, получим ту или иную формулу Адамса. В частности, если использовать многочлен нулевой степени (т.е. заменить подынтегральную функцию ее значением на левом конце отрезка в точке x_k), то получим явный метод Эйлера. Если проделать то же самое, но подынтегральную функцию аппроксимировать значением на правом конце в точке x_{k+1} , то получим неявный метод Эйлера.

Метод Адамса. При использовании интерполяционного многочлена третьей степени, построенного по значениям подынтегральной функции в последних четырех узлах, получим метод Адамса четвертого порядка точности:

$$y_{k+1} = y_k + \frac{h}{24}(55f_k - 59f_{k-1} + 37f_{k-2} - 9f_{k-3}), \quad (4.25)$$

где f_k — значение подынтегральной функции в узле x_k . Метод Адамса (4.25), как и все многошаговые методы, не является самостартующим, т.е. для того, чтобы использовать метод Адамса, необходимо иметь решения в первых четырех узлах. В узле x_0 решение y_0 известно из начальных условий, а в других трех узлах x_1, x_2, x_3 решения y_1, y_2, y_3 можно получить с помощью подходящего одношагового метода, например метода Рунге — Кутты четвертого порядка (4.10).

Метод Адамса — Бэшфорта — Моултона. Данный метод типа *предиктор — корректор* позволяет повысить точность вычислений метода Адамса за счет двойного вычисления значения функции $f(x, y)$ при определении y_{k+1} на каждом новом шаге по x .

Этап предиктор. Аналогично методу Адамса по значениям в узлах $x_{k-3}, x_{k-2}, x_{k-1}, x_k$ рассчитывается «предварительное» значение решения в узле x_{k+1} :

$$\tilde{y}_{k+1} = y_k + \frac{h}{24}(55f_k - 59f_{k-1} + 37f_{k-2} - 9f_{k-3}). \quad (4.26)$$

С помощью полученного значения \tilde{y}_{k+1} рассчитывается «предварительное» значение функции $f_{k+1} = f(x_k + 1, \tilde{y}_{k+1})$ в новой точке.

Этап корректор. На корректирующем этапе по методу Адамса четвертого порядка по значениям в узлах $x_{k-2}, x_{k-1}, x_k, x_{k+1}$ рассчитывается окончательное значение решения в узле x_{k+1} :

$$y_{k+1} = y_k + \frac{h}{24}(9f_{k+1} + 19f_k - 5f_{k-1} + f_{k-2}). \quad (4.27)$$

Пример 4.7. Методом Адамса с шагом $h = 0.1$ получить численное решение дифференциального уравнения

$$y' = (y + x)^2$$

с начальными условиями $y(0) = 0$ на интервале $[0, 1.0]$. Численное решение сравнить с точным решением $y = \operatorname{tg}(x) - x$.

Решение. Данная задача на первой половине интервала совпадает с задачей из примера 4.4. Поэтому для нахождения решения в первых узлах будем использовать результаты решения этой задачи методом Рунге – Кутты четвертого порядка (4.10), приведенные в примере 4.4. Решением задачи является табличная функция, располагающаяся во втором и третьем столбцах таблицы 4.13, в пятом столбце приведено точное решение и в последнем — погрешность численного решения.

Таблица 4.13

k	x_k	y_k	$f(x_k, y_k)$	$y_{\text{ист}}$	ε_k
0	0.0	0.0000000	0.000000000	0.000000	0.0000000
1	0.1	0.000334589	0.010067030	0.00033467	0.8301E-07
2	0.2	0.002709878	0.041091295	0.002710036	0.1573E-06
3	0.3	0.009336039	0.095688785	0.009336250	0.2103E-06
4	0.4	0.022715110	0.178688064	0.022793219	0.781090E-04
5	0.5	0.046098359	0.298223418	0.046302490	0.204131E-03
6	0.6	0.083724841	0.467479658	0.084136808	0.411968E-03
7	0.7	0.141501753	0.708125200	0.142288380	0.786628E-03
8	0.8	0.228133669	1.057058842	0.229638557	0.150489E-02
9	0.9	0.357181945	1.580506443	0.360158218	0.297627E-02
10	1.0	0.551159854	2.406096892	0.557407725	0.624787E-02

Пример 4.8. Методом Адамса – Бэшфорта – Моултона с шагом $h = 0.1$ получить численное решение начальной задачи из примера 4.7.

Решение. Как и в предыдущем примере, в первых трех узлах после начального решения получаем методом Рунге – Кутты четвертого порядка. Начиная с четвертого узла ($k = 4$), на каждом шаге в расчетах значений y_{k+1} используем соотношения (4.26), (4.27). Решением задачи является табличная функция, располагающаяся во втором и четвертом столбцах таблицы 4.14, в шестом столбце приведено точное решение и в последнем — погрешность численного решения. Решение, полученное методом Адамса – Бэшфорта – Моултона, несколько точнее, чем решение методом Адамса.

Таблица 4.14

k	x_k	\tilde{y}_k	y_k	$f(x_k, y_k)$	$y_{\text{ист}}$	ε_k
0	0.0	—	0.0000000	0.000000000	0.000000	0.0000000
1	0.1	—	0.000334589	0.010067030	0.00033467	0.8301E-07
2	0.2	—	0.002709878	0.041091295	0.002710036	0.1573E-06

Окончание табл. 4.14

k	x_k	\tilde{y}_k	y_k	$f(x_k, y_k)$	$y_{\text{ист}}$	ε_k
3	0.3	—	0.009336039	0.005688785	0.009336250	0.2103E-06
4	0.4	0.022715110	0.02279808	0.17875822	0.022793219	0.4863E-05
5	0.5	0.046197407	0.04631491	0.29845998	0.046302490	0.1242E-04
6	0.6	0.083978353	0.08416105	0.46807634	0.084136808	0.2424E-04
7	0.7	0.142027364	0.142331883	0.70952300	0.142288380	0.4350E-04
8	0.8	0.229171282	0.229714203	1.06031134	0.229638557	0.7565E-04
9	0.9	0.359247335	0.360288001	1.58832585	0.360158218	0.1298E-03
10	1.0	0.555451403	0.557625580	2.42619745	0.557407725	0.2179E-03

4.2. Численные методы решения краевой задачи для обыкновенных дифференциальных уравнений

Примером краевой задачи является двухточечная краевая задача для обыкновенного дифференциального уравнения второго порядка:

$$y'' = f(x, y, y') \quad (4.28)$$

с граничными условиями, заданными на концах отрезка $[a, b]$,

$$\begin{aligned} y(a) &= y_a, \\ y(b) &= y_b. \end{aligned} \quad (4.29)$$

Следует найти такое решение $y(x)$ на этом отрезке, которое принимает на концах отрезка значения y_a, y_b . Если функция $f(x, y, y')$ линейна по аргументам y, y' , то задача (4.28), (4.29) — линейная краевая задача, в противном случае — нелинейная.

Кроме граничных условий (4.29), называемых граничными условиями первого рода, используются еще условия на производные от решения на концах — граничные условия второго рода:

$$\begin{aligned} y'(a) &= \hat{y}_a, \\ y'(b) &= \hat{y}_b \end{aligned} \quad (4.30)$$

или линейная комбинация решений и производных — граничные условия третьего рода:

$$\begin{cases} \alpha y(a) + \beta y'(a) = \hat{y}_a, \\ \delta y(b) + \gamma y'(b) = \hat{y}_b, \end{cases} \quad (4.31)$$

где $\alpha, \beta, \delta, \gamma$ — числа такие, что $|\alpha| + |\beta| \neq 0, |\delta| + |\gamma| \neq 0$. Возможно на разных концах отрезка использовать условия различных типов.

Здесь рассматриваются два приближенных метода решения краевой задачи:

- метод стрельбы (пристрелки);
- конечно-разностный метод.

4.2.1. Метод стрельбы

Суть метода заключена в многократном решении задачи Коши для приближенного нахождения решения краевой задачи.

Пусть необходимо решить краевую задачу (4.28), (4.29) на отрезке $[a, b]$. Вместо исходной задачи формулируется задача Коши с уравнением (4.28) и с начальными условиями

$$\begin{aligned} y(a) &= y_a, \\ y'(a) &= \eta, \end{aligned} \tag{4.32}$$

где η — некоторое значение тангенса угла наклона касательной к решению в точке $x = a$.

Положим сначала некоторое начальное значение параметру $\eta = \eta_0$, после чего решим каким-либо методом задачу Коши (4.28), (4.32). Пусть $y = y(x, y_0, \eta_0)$ — решение этой задачи на интервале $[a, b]$. Тогда, сравнивая значение функции $y(b, y_0, \eta_0)$ со значением y_1 на правом конце отрезка, можно получить информацию для корректировки угла наклона касательной к решению в левом конце отрезка. Решая задачу Коши для нового значения $\eta = \eta_1$, получим другое решение со значением $y(b, y_0, \eta_1)$ на правом конце. Таким образом, значение решения на правом конце $y(b, y_0, \eta)$ будет являться функцией одной переменной η .

Задачу можно сформулировать следующим образом: требуется найти значение переменной η^* такое, чтобы решение $y(b, y_0, \eta^*)$ на правом конце отрезка совпало со значением y_1 из (4.29). Другими словами, решение исходной задачи эквивалентно нахождению корня уравнения

$$\Phi(\eta) = 0, \tag{4.33}$$

в котором $\Phi(\eta) = y(b, y_0, \eta) - y_1$.

Уравнение (4.33) является *алгоритмическим* уравнением, так как левая часть его задается с помощью алгоритма численного решения соответствующей задачи Коши. Но методы решения уравнения (4.33) аналогичны методам решения нелинейных уравнений, изложенным в разделе 2. Отметим, что поскольку невозможно вычислить производную функции $\Phi(\eta)$, вместо метода Ньютона следует использовать метод секущих, в котором производная от функции заменена ее разностным аналогом. Данный

разностный аналог легко вычисляется по двум приближениям, например η_j и η_{j+1} . Следующее значение искомого корня определяется с помощью соотношения

$$\eta_{j+2} = \eta_{j+1} - \frac{\eta_{j+1} - \eta_j}{\Phi(\eta_{j+1}) - \Phi(\eta_j)} \Phi(\eta_{j+1}). \quad (4.34)$$

Итерации по формуле (4.34) выполняются до удовлетворения заданной точности.

Пример 4.9. Методом стрельбы решить краевую задачу

$$y'' = e^x + \sin y$$

с граничными условиями первого рода на отрезке $[0, 1]$:

$$y(0) = 1, \quad y(1) = 2.$$

Решение. Заменой переменных $z = y'$ сведем дифференциальное уравнение второго порядка к системе двух дифференциальных уравнений первого порядка

$$\begin{cases} y' = z, \\ z' = e^x + \sin y. \end{cases}$$

Задачу Коши для системы с начальными условиями на левом конце $y(0) = 1$, $y'(0) = \eta$ будем решать методом Рунге — Кутты четвертого порядка точности с шагом $h = 0.1$ до удовлетворения на правом конце условия

$$|y(1.0, 1.0, \eta_k) - 2.0| = |\Phi(\eta_k)| \leq \varepsilon,$$

где $\varepsilon = 0.0001$ и $y(1.0, 1.0, \eta_k)$ — значение решения задачи Коши на правом конце отрезка при $b = 1.0$, $y(0) = y_0 = 1.0$, η_k — значение первой производной к решению на левом конце отрезка на k -й итерации.

Примем в качестве первых двух значений параметра η следующие: $\eta_0 = 1.0$, $\eta_1 = 0.8$. Дважды решив задачу Коши с этими параметрами методом Рунге — Кутты с шагом $h = 0.1$, получим два решения:

$$y(1.0, 1.0, \eta_0) = 3.168894836, \quad y(1.0, 1.0, \eta_1) = 2.97483325.$$

Вычислим новое приближение параметра η по формуле (4.34):

$$\eta_2 = 0.8 - \frac{0.8 - 1.0}{2.97483325 - 3.168894836} \cdot (2.97483325 - 2.0) = -0.204663797.$$

Решая задачу Коши с параметром η_2 , получим решение $y(1.0, 1.0, \eta_2) = 1.953759449$ и т.д.:

$$\begin{aligned} \eta_3 &= -0.204663797 - \frac{-0.204663797 - 0.8}{1.953759449 - 2.97483325} \cdot (1.953759449 - 2.0) = \\ &= -0.159166393; \end{aligned}$$

$$y(1.0, 1.0, \eta_3) = 2.001790565; |\Phi(\eta_3)| = 0.001790565 \geq \epsilon;$$

$$\eta_4 = -0.159166393 - \frac{-0.159166393 - (-0.204663797)}{2.001790565 - 1.953759449} \cdot (2.001790565 - 2.0) = \\ = -0.160862503;$$

$$y(1.0, 1.0, \eta_4) = 2.000003115; |\Phi(\eta_4)| = 0.000003115 \leq \epsilon.$$

Вычисления заносим в таблицу 4.15.

Таблица 4.15

j	η_j	$y(1.0, 1.0, \eta_j)$	$ \Phi(\eta_j) $
0	+1.000000000	3.168894836	1.168894836
1	+0.800000000	2.974483325	0.974483325
2	-0.204663797	1.953759449	0.046240551
3	-0.159166393	2.001790565	0.001790565
4	-0.160862503	2.000003115	0.000003115

Приближенным решением краевой задачи будем считать приведенную в таблице 4.16 табличную функцию, полученную в результате решения задачи Коши с параметром η_4 .

Таблица 4.16

k	0	1	2	3	4
x_k	0.0	0.10000	0.20000	0.30000	0.40000
y_k	1.0	0.99328	1.00601	1.03942	1.09497
5	6	7	8	9	10
0.50000	0.60000	0.70000	0.80000	0.90000	1.000
1.17434	1.27944	1.41236	1.57528	1.77045	2.000

4.2.2. Конечно-разностный метод решения краевой задачи

Рассмотрим двухточечную краевую задачу для линейного дифференциального уравнения второго порядка на отрезке $[a, b]$ с соответствующими краевыми условиями

$$y'' + p(x)y' + q(x)y = f(x), \quad (4.35)$$

$$y(a) = y_a, \quad y(b) = y_b. \quad (4.36)$$

Введем на отрезке $[a, b]$ разностную сетку $\Omega^{(h)} = \{x_k = x_a + hk\}$, $h = (b - a)/N$, $k = 0, 1, \dots, N$. Решение задачи (4.35), (4.36) будем искать в виде сеточной функции $y^{(h)} = \{y_k\}$, $k = 0, 1, \dots, N\}$,

предполагая, что решение существует и единственно. Введем разностную аппроксимацию производных ([1], параграф 2.10):

$$\begin{aligned} y'_k &= \frac{y_{k+1} - y_{k-1}}{2h} + O(h^2); \\ y''_k &= \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} + O(h^2). \end{aligned} \quad (4.37)$$

Подставляя аппроксимации производных из (4.37) в (4.35), (4.36), получим систему уравнений для нахождения y_k :

$$\left\{ \begin{array}{l} y_0 = y_a, \\ \dots \\ \frac{y_{k+1} - y_k + y_{k-1}}{h^2} + p(x_k) \frac{y_{k+1} - y_{k-1}}{2h} + q(x_k) y_k = f(x_k), \\ \dots \\ y_N = y_b, \end{array} \right. \quad k = 1, 2, \dots, N-1. \quad (4.38)$$

Приводя подобные и учитывая, что при задании граничных условий первого рода две неизвестные y_0, y_N уже фактически определены, получим систему линейных алгебраических уравнений с трехдиагональной матрицей коэффициентов:

$$\left\{ \begin{array}{l} \left(-2 + h^2 q(x_1) \right) y_1 + \left(1 + \frac{p(x_1)h}{2} \right) y_2 = h^2 f(x_1) - \left(1 - \frac{p(x_1)h}{2} \right) y_2, \\ \left(1 - \frac{p(x_k)h}{2} \right) y_{k-1} + \left(-2 + h^2 q(x_k) \right) y_k + \left(1 + \frac{p(x_k)h}{2} \right) y_{k+1} = h^2 f(x_k), \\ \left(1 - \frac{p(x_{N-1})h}{2} \right) y_{N-1} + \left(-2 + h^2 q(x_{N-1}) \right) y_N = \\ = h^2 f(x_{N-1}) - \left(1 + \frac{p(x_{N-1})h}{2} \right) y_b, \end{array} \right. \quad k = 2, \dots, N-2. \quad (4.39)$$

Для системы (4.39) при достаточно малых шагах сетки h и $q(x_k) < 0$ выполнены условия преобладания диагональных элементов

$$\left\{ \begin{array}{l} \left| -2 + h^2 q(x_1) \right| > \left| 1 + \frac{p(x_1)h}{2} \right|, \\ \left| -2 + h^2 q(x_k) \right| > \left| 1 - \frac{p(x_k)h}{2} \right| + \left| 1 + \frac{p(x_k)h}{2} \right|, \quad k = 2, \dots, N-2, \\ \left| -2 + h^2 q(x_{N-1}) \right| > \left| 1 - \frac{p(x_{N-1})h}{2} \right|, \end{array} \right. \quad (4.40)$$

что гарантирует устойчивость счета и корректность применения метода прогонки для решения этой системы.

В случае использования граничных условий второго и третьего рода аппроксимация производных проводится с помощью односторонних разностей первого или второго порядков.

$$\begin{cases} y'_0 = \frac{y_1 - y_0}{h} + O(h), \\ y'_n = \frac{y_N - y_{N-1}}{h} + O(h), \end{cases} \quad (4.41)$$

$$\begin{cases} y'_0 = \frac{-3y_0 + y_1 - y_2}{2h} + O(h^2), \\ y'_N = \frac{y_{N-2} - 4y_{N-1} + 3y_N}{2h} + O(h^2). \end{cases} \quad (4.42)$$

В случае использования формул (4.41) линейная алгебраическая система аппроксимирует дифференциальную задачу в целом только с первым порядком (из-за аппроксимации в граничных точках), однако сохраняется трехдиагональная структура матрицы коэффициентов. При использовании формул (4.42) второй порядок аппроксимации сохраняется везде, но матрица линейной системы не трехдиагональная.

Пример 4.10. Решить краевую задачу с шагом $h = 0.2$:

$$\begin{cases} y'' + xy' - y = 0, \\ y(0) = 1, \\ y'(1) + 2y(1) = 0. \end{cases}$$

Решение. Здесь

$$\begin{aligned} p(x) &= x, & q(x) &= -1, & f(x) &= 0, & N &= 5, \\ x_0 &= 0, x_1 = 0.2, x_2 = 0.4, x_3 = 0.6, x_4 = 0.8, x_5 = 1.0. \end{aligned}$$

Во всех внутренних узлах отрезка $[0, 1]$ после замены производных их разностными аналогами получим

$$(1 - 0.1x_k)y_{k-1} + (-2.04)y_k + (1 + 0.1x_k)y_{k+1} = 0, \quad k = 1, \dots, 4.$$

На левой границе $y_0 = 1$, на правой границе аппроксимируем производную односторонней разностью первого порядка:

$$\frac{y_5 - y_4}{0.2} + 2y_5 = 0.$$

С помощью группировки слагаемых, приведения подобных членов и подстановки значений x_k и с учетом граничного условия $y_0 = 1$ получим систему линейных алгебраических уравнений:

$$\begin{cases} -2.04y_1 + 1.02y_2 &= -0.98, \\ 0.96y_1 - 2.04y_2 + 1.04y_3 &= 0, \\ 0.94y_2 - 2.04y_3 + 1.06y_4 &= 0, \\ 0.92y_3 - 2.04y_4 + 1.08y_5 &= 0, \\ y_4 - 1.4y_5 &= 0. \end{cases}$$

В данной трехдиагональной системе выполнено условие преобладания диагональных элементов, и можно использовать метод прогонки (раздел 1.1.2). В результате решения системы методом прогонки получим следующие значения:

$$y_5 = 0.2233205, y_4 = 0.31265, y_3 = 0.43111, y_2 = 0.58303, y_1 = 0.77191.$$

Решением краевой задачи является табличная функция, приведенная в таблице 4.17.

Таблица 4.17

k	0	1	2	3	4	5
x_k	0	0.2	0.4	0.6	0.8	1.0
y_k	1.0	0.77191	0.58303	0.43111	0.31265	0.22332

Задачи

4.1. Решить задачу Коши методом по выбору преподавателя для обычного дифференциального уравнения первого порядка на указанном отрезке с заданным шагом h . Полученное численное решение сравнить с точным. Определить погрешность решения.

№	Задача Коши	Точное решение
1	$y' = -y + e^x$, $y(0) = 1.5$, $x \in [0, 1]$, $h = 0.1$	$y = \frac{e^x}{2} + e^{-x}$
2	$y' = -2xy + xe^{-x^2}$, $y(0) = 0$, $x \in [0, 1]$, $h = 0.1$	$y = \frac{e^{-x^2}}{2}$

Продолжение табл.

№	Задача Коши	Точное решение
3	$y' = -y^2 - \frac{1}{x^4}$, $y(1) = 1$, $x \in [1, 2]$, $h = 0.1$	$y = \frac{x - \operatorname{tg}\left(\frac{-1+x}{x}\right)}{x^2}$
4	$y' = -y \cos(x) + \frac{1}{2} \sin(2x)$, $y(0) = 0$, $x \in [0, 1]$, $h = 0.1$	$y = \sin x - 1 + e^{\sin x}$
5	$y' = \frac{-3yx + 8y - x^2}{x^2 - 5x + 6}$, $y(0) = -1$, $x \in [0, 1]$, $h = 0.1$	$y = \frac{-(1/4)x^4 + (2/3)x^3 + 12}{(x-2)^2(x-3)}$
6	$y' = -ytgx + \sin 2x$, $y(0) = -1$, $x \in [0, 1]$, $h = 0.1$	$y = -2\cos^2 x + \cos x$
7	$y' = \frac{(x-1)y}{x^2}$, $y(1) = e$, $x \in [1, 2]$, $h = 0.1$	$y = xe^{1/x}$
8	$y' = y \sin x - \cos x \sin x$, $y(0) = 0.367879$, $x \in [0, 1]$, $h = 0.1$	$y = \cos x - 1 + e^{\cos x}$
9	$y' = \frac{y^2 + x^2 y}{x^3}$, $y(1) = 0.5$, $x \in [1, 2]$, $h = 0.1$	$y = \frac{x^2}{1+x}$
10	$y' = (y+x)^2$, $y(0) = 0$, $x \in [0, 1]$, $h = 0.1$	$y = \operatorname{tg} x - x$

Продолжение табл.

№	Задача Коши	Точное решение
11	$y' = 2 \frac{x^2 - xy}{x^2 + 1},$ $y(1) = 0.6666667,$ $x \in [1, 2], h = 0.1$	$y = \frac{2}{3} \left(\frac{x^3 + 1}{x^2 + 1} \right)$
12	$y' = \frac{y^2 + yx}{x^2},$ $y(1) = 1,$ $x \in [1, 2], h = 0.1$	$y = -\frac{x}{\ln x - 1}$
13	$y' = x^2(y^2 + 1),$ $y(0) = 0,$ $x \in [0, 1], h = 0.1$	$y = \operatorname{tg} \left(\frac{1}{3} x^3 \right)$
14	$y' = -\frac{1}{2} \frac{y}{x} + x^2,$ $y(1) = 1,$ $x \in [1, 2], h = 0.1$	$y = \frac{2}{7} x^3 + \frac{5}{7\sqrt{x}}$
15	$y' = \frac{y}{x^2} + e^{x-1/x},$ $y(1) = 1.367879,$ $x \in [1, 2], h = 0.1$	$y = e^{-1/x} (e^x + 1)$
16	$y' = e^x - y + e^x,$ $y(0) = 0.541325,$ $x \in [0, 1], h = 0.1$	$y = \ln(-1 + e^{e^x})$
17	$y' = (-y + x \sin x) / x,$ $y(1) = 0.3011687 = \sin 1 - \cos 1,$ $x \in [1, 2], h = 0.1$	$y = \frac{\sin x - x \cos x}{x}$
18	$y' = \frac{y}{x} + \frac{1}{\ln x },$ $y(2) = -0.733026,$ $x \in [2, 3], h = 0.1$	$y = x \ln \ln x $

Продолжение табл.

№	Задача Коши	Точное решение
19	$y' = \frac{y}{x} + x^2 \sin x,$ $y(1) = 1.3011687 = \sin 1 - \cos 1 + 1,$ $x \in [1, 2], h = 0.1$	$y = x \sin x - x^2 \cos x + x$
20	$y' = \frac{-yx + x(x^2 + 1)}{x^2 + 1},$ $y(0) = 1,$ $x \in [0, 1], h = 0.1$	$y = \frac{1}{3}x^2 + \frac{1}{3} + \frac{2}{3\sqrt{x^2 + 1}}$
21	$y' = -\frac{y}{x} - x^2,$ $y(1) = 0.75,$ $x \in [1, 2], h = 0.1$	$y = \frac{-\frac{1}{4}x^4 + 1}{x}$
22	$y' = \frac{1-yx}{x^2+1},$ $y(1) = 0.623225,$ $x \in [1, 2], h = 0.1$	$y = \frac{\ln(x + \sqrt{x^2 + 1})}{\sqrt{x^2 + 1}}$
23	$y' = -y^2 + \frac{y}{x},$ $y(1) = 2/3,$ $x \in [1, 2], h = 0.1$	$y = 2 \frac{x}{x^2 + 2}$
24	$y' = -\left(\frac{y}{x}\right)^2 - \frac{y}{x} - 1,$ $y(1) = 0,$ $x \in [1, 2], h = 0.1$	$y = -\frac{x \ln x}{\ln x + 1}$
25	$y' = \frac{y^2 \ln x - y}{x},$ $y(1) = 0.5,$ $x \in [1, 2], h = 0.1$	$y = \frac{1}{\ln x + 1 + x}$

Окончание табл.

№	Задача Коши	Точное решение
26	$y' = \frac{y(2y \ln x - 1)}{x}$, $y(1) = 0.3333333333$, $x \in [1, 2]$, $h = 0.1$	$y = \frac{1}{2(\ln x + 1) + x}$
27	$y' = \frac{y^2 + xy}{x^2}$, $y(1) = 1$, $x \in [1, 2]$, $h = 0.1$	$y = \frac{x}{\ln x - 1}$
28	$y' = e^{y/x} + \frac{y}{x} + 1$, $y(1) = -0.541325$, $x \in [1, 2]$, $h = 0.1$	$y = x \left(\ln \left(-\frac{x}{-1 + x/e} \right) - 1 \right)$
29	$y' = \frac{y \ln y}{x}$, $y(1) = e$, $x \in [1, 2]$, $h = 0.1$	$y = e^x$
30	$y' = \frac{y(\ln(xy) - 1)}{x}$, $y(1) = e$, $x \in [1, 2]$, $h = 0.1$	$y = \frac{1}{x} e^x$

4.2. Решить задачу Коши для обыкновенного дифференциального уравнения второго порядка на указанном отрезке с заданным шагом h . Полученное численное решение сравнить с точным. Определить погрешность решения.

№	Задача Коши	Точное решение
1	$y'' + y - \sin 3x = 0$, $y(0) = 1$, $y'(0) = 1$, $x \in [0, 1]$, $h = 0.1$	$y = \cos x + \frac{11}{8} \sin x - \frac{\sin 3x}{8}$

Продолжение табл.

№	Задача Коши	Точное решение
2	$y'' + y - 2\cos x = 0,$ $y(0) = 1,$ $y'(0) = 1,$ $x \in [0, 1], \quad h = 0.1$	$y = \sin x + \cos x$
3	$y'' - 2y - 4x^2 e^{x^2} = 0,$ $y(0) = 3,$ $y'(0) = 0,$ $x \in [0, 1], \quad h = 0.1$	$y = e^{x^2} + e^{x\sqrt{2}} + e^{-x\sqrt{2}}$
4	$x^2 y'' - x(x^2 - 1)y' - (x^2 + 1)y = 0,$ $y(1) = 1 + e^{1/2},$ $y'(1) = 2e^{1/2} - 1,$ $x \in [1, 2], \quad h = 0.1$	$y = \frac{1}{x} (1 + e^{x^2/2})$
5	$y'' - (1 + 2\tg^2 x)y - \sin 3x = 0,$ $y(0) = 1,$ $y'(0) = 2,$ $x \in [0, 1], \quad h = 0.1$	$y = \frac{1}{\cos x} + \sin x + \frac{x}{\cos x}$
6	$y'' + 4xy' + (4x^2 + 2)y = 0,$ $y(0) = 1,$ $y'(0) = 1,$ $x \in [0, 1], \quad h = 0.1$	$y = (1+x)e^{-x^2}$
7	$y'' - 4xy' + (4x^2 - 2)y = 0,$ $y(0) = 1,$ $y'(0) = 1,$ $x \in [0, 1], \quad h = 0.1$	$y = (1+x)e^{x^2}$

Продолжение табл.

№	Задача Коши	Точное решение
8	$y'' - 4xy' + (4x^2 - 3)y = e^{x^2}$, $y(0) = 1$, $y'(0) = 0$,	$y = (e^x + e^{-x} - 1)e^{x^2}$
9	$x \in [0, 1]$, $h = 0.1$ $y'' - \left(\frac{1}{x^{1/2}}\right)y' +$ $+ \left(\frac{1}{4x^2}\right)(x + x^{1/2} - 8)y = 0$, $y(1) = 2e$, $y'(1) = 2e$,	$y = \left(x^2 + \frac{1}{x}\right)e^{x^{1/2}}$
10	$x \in [1, 2]$, $h = 0.1$ $y'' + y'\operatorname{tg}x + y\cos^2 x = 0$, $y(0) = 0$, $y'(0) = 1$,	$y = \cos x(\sin x) + \sin x(\cos x)$
11	$x \in [0, 1]$, $h = 0.1$ $y'' + y'\operatorname{tg}x - y\cos^2 x = 0$, $y(0) = 2$, $y'(0) = 0$,	$y = e^{\sin x} + e^{-\sin x}$
12	$x \in [0, 1]$, $h = 0.1$ $(x^2 + 1)y'' - 2xy' + 2y = 0$, $y(0) = 1$, $y'(0) = 1$,	$y = x - x^2 + 1$
13	$x \in [0, 1]$, $h = 0.1$ $y'' - 2(\operatorname{tg}x)y' + 3y = 0$, $y(0) = 1$, $y'(0) = 2$,	$y = \frac{\cos 2x + \sin 2x}{\cos x}$
	$x \in [0, 1]$, $h = 0.1$	

Продолжение табл.

№	Задача Коши	Точное решение
14	$y'' + 2y' \operatorname{ctgx} x + 3y = 0,$ $y(0) = 1,$ $y'(0) = 1,$ $x \in [1, 2], \quad h = 0.1$	$y = \frac{-0.9783 \cos 2x + 0.4776 \sin 2x}{\sin x}$
15	$xy'' + y' = 0,$ $y(1) = 1,$ $y'(1) = 1,$ $x \in [1, 2], \quad h = 0.1$	$y = 1 + \ln x $
16	$(x^2 - 1)y'' - 2xy' + 2y = 0,$ $y(2) = 7,$ $y'(2) = 5,$ $x \in [2, 3], \quad h = 0.1$	$y = x^2 + x + 1$
17	$xy'' - (x+1)y' + y = 0,$ $y(1) = 2 + e,$ $y'(1) = 1 + e,$ $x \in [1, 2], \quad h = 0.1$	$y = x + 1 + e^x$
18	$y'' - \frac{x+1}{x}y' - 2\frac{x-1}{x}y = 0,$ $y(1) = 1,$ $y'(1) = 1,$ $x \in [1, 2], \quad h = 0.1$	$y = \frac{e^{2x}}{3e^2} + \frac{(3x+1)e^{-x}}{3e}$
19	$y'' + \frac{1}{x}y' + \frac{2}{x}y = 0,$ $y(1) = 1,$ $y'(1) = 1,$ $x \in [1, 2], \quad h = 0.1$	$y = (\cos 2 - \sin 2)\cos(2x^{1/2}) + (\cos 2 + \sin 2)\sin(2x^{1/2})$

Продолжение табл.

№	Задача Коши	Точное решение
20	$x(x-1)y'' + \frac{1}{2}y' - \frac{3}{4}y = 0,$ $y(2) = \sqrt{2},$ $y'(2) = \frac{3}{2}\sqrt{2},$	$y = x ^{3/2}$
21	$x \in [2, 3], h = 0.1$ $x^2y'' - 12y = 0,$ $y(1) = 2,$ $y'(1) = 1,$	$y = x^4 + x^{-3}$
22	$x \in [1, 2], h = 0.1$ $x^2y'' + (x^2 - 2)y = 0,$ $y(1) = 1,$ $y'(1) = 0,$	$y = \sin(x-1) + \frac{1}{x}\cos(x-1)$
23	$x \in [1, 2], h = 0.1$ $x^2y'' - xy' - y - 3x^2 = 0,$ $y(1) = 3,$ $y'(1) = 2,$	$y = x^2 + x + \frac{1}{x}$
24	$x \in [1, 2], h = 0.1$ $x^2y'' + (x+1)y' - y = 0,$ $y(1) = 2 + e,$ $y'(1) = 1,$	$y = x + 1 + xe^{1/x}$
25	$x \in [1, 2], h = 0.1$ $(x-2)^2y'' - (x-2)y' - 3y = 0,$ $y(3) = 2,$ $y'(3) = 2,$ $x \in [3, 4], h = 0.1$	$y = (x-2)^3 + \frac{1}{x-2}$

Окончание табл.

№	Задача Коши	Точное решение
26	$x^4y'' + 2x^3y' + y = 0,$ $y(1) = 1,$ $y'(1) = 1,$	$y = (\sin 1 + \cos 1)\cos \frac{1}{x} + (\sin 1 - \cos 1)\sin \frac{1}{x}$
27	$x \in [1, 2], h = 0.1$ $x^2y'' - 2xy' + (x^2 + 2)y = 0,$ $y(\pi/2) = \pi/2,$ $y'(\pi/2) = 1 - \pi/2,$	$y = x \cos x + x \sin x$
28	$x \in [\pi/2, \pi/2+1], h = 0.1$ $x^2y'' - 3y' + 4y - 5x = 0,$ $y(1) = 6,$ $y'(1) = 8,$	$y = 5x + x^2 + x^2 \ln x $
29	$x \in [1, 2], h = 0.1$ $x^2y'' - 3xy' - 5y - x^2 \ln x = 0,$ $y(1) = 1,$ $y'(1) = 1,$	$y = \frac{19}{54}x^5 + \frac{35}{54} \cdot \frac{1}{x} - \frac{x^2}{9} \ln x $
30	$x \in [1, 2], h = 0.1$ $x^2(x+1)y'' - x(2x+1)y' + (2x+1)y = 0,$ $y(1) = 2,$ $y'(1) = 4,$	$y = x^2 + x + x \ln x$

4.3. Решить задачу Коши для обыкновенного дифференциального уравнения первого порядка с запаздывающим аргументом на отрезке $[0, 1]$ с заданным шагом $h = 0.1$ улучшенным методом Эйлера.

№	Уравнения	A_1	A_2	A_3	A_4	$y(x_0)$	$z(x_0)$
1		0.1	0.2	0.27	10	1.1	—
2		0.1	0.3	0.17	5	1.2	—
3	$y' = \frac{A_1 \cdot y(x) + A_2 \cdot y(x - A_3)}{1 + (y(x + A_3))^{A_4}}$	0.15	0.2	0.15	10	1.3	—
4		0.15	0.3	0.25	5	1.4	—
5		0.1	0.2	0.25	10	1.5	—

Окончание табл.

№	Уравнения	A_1	A_2	A_3	A_4	$y(x_0)$	$z(x_0)$
6		3	0.125	0.125	0.5	1.6	—
7		4	0.125	0.15	0.6	1.7	—
8	$y' = y(x) \cdot (A_1 + 3 \cdot (A_2 \cdot y(x - A_4) -$	3.5	0.15	0.125	0.7	1.8	—
9	$-y(x - A_3)) / (1 - A_2)$	2.5	0.15	0.15	0.45	1.9	—
10		5	0.15	0.125	0.68	2.0	—
11		1	1	0.11	—	0.005	—
12		1.2	2	0.13	—	0.1	—
13	$y' = A_2 - A_1 \cdot y(x) - y^2(x - A_3)$	1.5	1.5	0.21	—	0.2	—
14		1.3	2	0.23	—	0.3	—
15		1.1	1	0.25	—	0.4	—
16		1	0.75	0.41	—	1	—
17		1	0.5	0.32	—	1.5	—
18	$y' = -A_1 \cdot y(x) +$	1.5	0.3	0.24	—	2	—
19	$+ A_2 \cdot \{1 - [\sin y(x - A_3)]\}$	1.8	0.9	0.18	—	2.5	—
20		1.2	0.1	0.25	—	3	—
21		1.25	0.15	10	—	1.4	—
22		1.5	0.25	1	—	2.4	—
23	$y' = A_1 \cdot y(x) \cdot (1 - y(x - A_2) / A_3)$	2.0	0.27	100	—	3.4	—
24		1.6	0.32	10	—	4.4	—
25		1.8	0.17	1	—	5.4	—
26		0.35	0.35	—	—	1.0	1.0
27		0.25	0.35	—	—	2.0	1.0
28	$y' = y(x) \cdot [2 - 2 \cdot y(x - A_1) - z(x)]$	0.35	0.25	—	—	1.0	2.0
29	$z' = z(x) \cdot [2 - 2 \cdot z(x - A_2) - y(x)]$	0.23	0.4	—	—	2.0	2.0
30		0.4	0.24	—	—	2.0	0.5

4.4. Решить краевую задачу для обыкновенного дифференциального уравнения второго порядка на указанном отрезке методом стрельбы и конечно-разностным методом.

№	Краевая задача	Точное решение
1	$xy'' + 2y' - xy = 0,$ $y'(1) = 0,$ $1.5y(2) + y'(2) = e^2$	$y(x) = \frac{e^x}{x}$
2	$xy'' + 2y' - xy = 0,$ $y'(1) = e^{-1},$ $y(2) = 0.5e^{-2}$	$y(x) = \frac{e^{-x}}{x}$
3	$x^2(x+1)y'' - 2y' = 0,$ $y'(1) = -1,$ $2y(2) - 4y'(2) = 4$	$y(x) = \frac{1}{x} + 1$

Продолжение табл.

№	Краевая задача	Точное решение
4	$x^2(x+1)y'' - 2y' = 0,$ $y'(1) = 1 + 4 \ln 2,$ $y(2) = -1 + 3 \ln 2$	$y(x) = -1 + \frac{2}{x} + \frac{2(x+1)}{x} \ln x+1 $
5	$y'' - \frac{2}{\cos^2 x} y = 0,$ $y'(0) = 1,$ $y'\left(\frac{\pi}{4}\right) = 1$	$y(x) = \operatorname{tg} x$
6	$y'' - 2\left(1 + (\operatorname{tg} x)^2\right)y = 0,$ $y(0) = 0,$ $y\left(\frac{\pi}{4}\right) = \sqrt{3}$	$y(x) = -\operatorname{tg} x$
7	$(2x+1)y'' + 4xy' - 4y = 0,$ $y'(1) + 2y(1) = 3,$ $y'(0) = 1$	$y(x) = x + e - 2x$
8	$(2x+1)y'' + 4xy' - 4y = 0,$ $y'(-2) + 2y(-2) = -9,$ $y'(0) = 1$	$y(x) = 3x + e - 2x$
9	$xy'' - (2x+1)y' + (x+1)y = 0,$ $y'(0) = 1,$ $y'(1) - 2y(1) = 0$	$y(x) = e^x(x^2 + 1)$
10	$xy'' - (2x+1)y' + (x+1)y = 0,$ $y'(1) = 3e,$ $y'(2) - 2y(2) = 0$	$y(x) = e^x x^2$
11	$x(x-1)y'' - xy' + y = 0,$ $y'(1) = 2,$ $2y'(2) - y(2) = 1$	$y(x) = 1 + x + x \ln x $

Продолжение табл.

№	Краевая задача	Точное решение
12	$x(x-1)y'' - xy' + y = 0,$ $y'(1) = 3,$ $y'(3) - 3y(3) = 4$	$y(x) = 2 + x + 2x \ln x $
13	$(e^x + 1)y'' - 2y' - e^x y = 0,$ $y'(0) = \frac{3}{4},$ $y'(1) = \frac{e^2(e+2)}{(e+1)^2}$	$y(x) = e^x - 1 + \frac{1}{e^x + 1}$
14	$(e^x + 1)y'' - 2y' - e^x y = 0,$ $y'(0) = 1,$ $y'(1) - y(1) = 1$	$y(x) = e^x - 1$
15	$x^2 \ln x y'' - xy' + y = 0,$ $y'(-1) = 0,$ $y'(1) - y(1) = 0$	$y(x) = 1 + x + x \ln x$
16	$y'' - \operatorname{tg} x y' + 2y = 0,$ $y'(0) = 2,$ $y'(\frac{\pi}{6}) = 2.5 - 0.5 \cdot \ln 3$	$y(x) = \sin x + 2 -$ $- \sin x \cdot \ln\left(\frac{1+\sin x}{1-\sin x}\right)$
17	$(x^2 - 1)y'' - (x-3)y' - y = 0,$ $y'(0) = 0,$ $y'(1) + y(1) = -0.75$	$y(x) = x - 3 + \frac{1}{x+1}$
18	$xy'' - (x+1)y' - 2(x-1)y = 0,$ $y'(0) = 4,$ $y'(1) - 2y(1) = -9e^{-1}$	$y(x) = e^{2x} + (3x+1)e^{-x}$
19	$y'' + 4xy' + (4x^2 + 2)y = 0,$ $y'(0) = 1,$ $4y(2) - y'(2) = 23e^{-4}$	$y(x) = (1+x)e^{-x^2}$

Продолжение табл.

№	Краевая задача	Точное решение
20	$xy'' - (2x+1)y' + 2y = 0,$ $y'(0) = 4,$ $y'(1) - 2y(1) = -4$	$y(x) = 2x + 1 + e^{2x}$
21	$x(2x+1)y'' - 2(x+1)y' - 2y = 0,$ $y'(1) = 0,$ $y(3) - y'(3) = \frac{31}{9}$	$y(x) = x + 1 + \frac{1}{x}$
22	$x(x+4)y'' - (2x+4)y' + 2y = 0,$ $y'(0) = 1,$ $y(2) - y'(2) = 3$	$y(x) = x^2 + x + 2$
23	$x(x^2+6)y'' - 4(x^2+3)y' + 6xy = 0,$ $y'(0) = 0,$ $y(4) - y'(4) = 26$	$y(x) = x^3 + x^2 + 2$
24	$(x^2+1)y'' - 2y = 0,$ $y'(0) = 2,$ $y(1) = 3 + \frac{\pi}{2}$	$y(x) = x^2 + x + 1 + (x^2 + 1)\arctg(x)$
25	$2x(x+2)y'' + (2-x)y' + y = 0,$ $y'(1) = \frac{3}{2},$ $y'(4) + y(4) = \frac{21}{4}$	$y(x) = \sqrt{ x } + x - 2$
26	$x(x+1)y'' + (x+2)y' - y = x + \frac{1}{x},$ $y'(1) = \frac{3}{2},$ $4y'(2) + y(2) = 13 + 4\ln 2$	$y(x) = x + \frac{7}{2} + \frac{1}{x} + \left(\frac{x}{2} + 1\right)\ln x $

Окончание табл.

№	Краевая задача	Точное решение
27	$(2x+1)y'' + (2x-1)y' - 2y = x^2 + x,$ $y'(0) = 1,$ $y'(1) + y(1) = 5$	$y(x) = 2x - 1 + e^{-x} + \frac{x^2 + 1}{2}$
28	$xy'' - (2x+1)y' + 2y = 0,$ $y'(0) = 2,$ $y(1) = e^2$	$y(x) = e^{2x}$
29	$(x^2 - 1)y'' + (x - 3)y' - y = 0,$ $y(0) = 18,$ $y(3) = 0$	$y(x) = 6x - 18$
30	$(x^2 + 1)y'' - 2y = 0,$ $y'(0) = 0,$ $y(2) - y'(2) = 1$	$y(x) = x^2 + 1$

Глава 5

ЧИСЛЕННОЕ РЕШЕНИЕ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ С ЧАСТНЫМИ ПРОИЗВОДНЫМИ

В настоящем разделе рассматривается метод конечных разностей, который является одним из наиболее распространенных численных методов решения уравнений с частными производными — уравнений математической физики.

5.1. Численное решение уравнений параболического типа.

Понятие о методе конечных разностей.

Основные определения и конечно-разностные схемы

5.1.1. Постановка задач для уравнений параболического типа

Классическим примером уравнения параболического типа является уравнение теплопроводности (диффузии). В одномерном по пространству случае однородное (без источников энергии) уравнение теплопроводности имеет вид

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < l, \quad t > 0. \quad (5.1)$$

Если на границах $x = 0$ и $x = l$ заданы значения искомой функции $u(x, t)$ в виде

$$u(0, t) = \varphi_0(t), \quad t > 0; \quad (5.2)$$

$$u(l, t) = \varphi_l(t), \quad t > 0, \quad (5.3)$$

т.е. *граничные условия первого рода*, и, кроме того, заданы *начальные условия* при $t = 0$

$$u(x, 0) = \psi(x), \quad 0 < x < l, \quad (5.4)$$

то задачу (5.1)–(5.4) называют *первой начально-краевой задачей для уравнения теплопроводности* (5.1).

В терминах теории теплообмена функция $u(x, t)$ выражает распределение температуры в пространственно-временной области

$$\Omega \times T = \{0 \leq x \leq l; 0 \leq t \leq T\},$$

a^2 — коэффициент температуропроводности, а (5.2), (5.3) с помощью функций $\varphi_0(t)$, $\varphi_l(t)$ задают температуру на границах $x = 0$ и $x = l$.

Если на границах $x = 0$ и $x = l$ заданы значения производных искомой функции по пространственной переменной

$$\frac{\partial u(0, t)}{\partial x} = \varphi_0(t), \quad t > 0; \quad (5.5)$$

$$\frac{\partial u(l, t)}{\partial x} = \varphi_l(t), \quad t > 0, \quad (5.6)$$

т.е. *граничные условия второго рода*, то задачу (5.1), (5.5), (5.6), (5.4) называют *второй начально-краевой задачей для уравнения теплопроводности* (5.1). В терминах теории теплообмена на границах в этом случае заданы тепловые потоки.

Если на границах заданы линейные комбинации искомой функции и ее производной по пространственной переменной

$$\alpha \frac{\partial u(0, t)}{\partial x} + \beta u(0, t) = \varphi_0(t), \quad t > 0; \quad (5.7)$$

$$\gamma \frac{\partial u(l, t)}{\partial x} + \delta u(l, t) = \varphi_l(t), \quad t > 0, \quad (5.8)$$

т.е. *граничные условия третьего рода*, то задачу (5.1), (5.7), (5.8), (5.4) называют *третьей начально-краевой задачей для уравнения теплопроводности* (5.1). В терминах теории теплообмена граничные условия (5.7), (5.8) задают теплообмен между газообразной или жидкой средой и границами расчетной области с неизвестными температурами $u(0, t)$, $u(l, t)$.

Для *пространственных задач теплопроводности* первая начально-краевая задача в области $\bar{\Omega} = \Omega + \Gamma$ имеет вид

$$\begin{cases} \frac{\partial u}{\partial t} = a^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right), & M(x, y, z) \in \Omega, \quad t > 0; \\ u(M, t) = \varphi(M, t), & M(x, y, z) \in \Gamma, \quad t > 0; \\ u(M, 0) = \psi(M), & M(x, y, z) \in \bar{\Omega}. \end{cases} \quad (5.9)$$

$$(5.10)$$

$$(5.11)$$

Аналогично ставятся вторая и третья начально-краевые задачи для пространственного уравнения задачи (5.9)–(5.11).

На практике часто ставятся начально-краевые задачи теплопроводности со смешанными краевыми условиями, когда на границах задаются граничные условия различных родов.

5.1.2. Понятие о методе конечных разностей.

Применение метода конечных разностей к решению уравнений параболического типа

Основные определения, связанные с методом конечных разностей, рассмотрим на примере конечно-разностного решения первой начально-краевой задачи для уравнения теплопроводности (5.1)–(5.4).

Нанесем на пространственно-временную область $0 \leq x \leq l$, $0 \leq t \leq T$ конечно-разностную сетку ω_{ht}

$$\omega_{ht} = \{x_j = jh, \quad j = 0, 1, \dots, N; \quad t^k = k\tau, \quad k = 0, 1, \dots, K\} \quad (5.12)$$

с пространственным шагом $h = l/N$ и шагом по времени $\tau = T/K$ (рис. 5.1).

Введем два временных слоя: нижний $t^k = k\tau$, на котором распределение искомой функции $u(x_j, t^k)$, $j = 0, 1, 2, \dots, N$ известно (при $k = 0$ распределение определяется начальным условием (5.4) вида $u(x_j, t^0) = \psi(x_j)$), и верхний временной слой $t^{k+1} = (k + 1)\tau$, на котором распределение искомой функции $u(x_j, t^{k+1})$, $j = 0, 1, 2, \dots, N$, подлежит определению.

Сеточной функцией задачи (5.1)–(5.4), имеющей обозначение u_j^k , назовем однозначное отображение целых аргументов j, k

в значения функции $u_j^k = u(x_j, t^k)$.

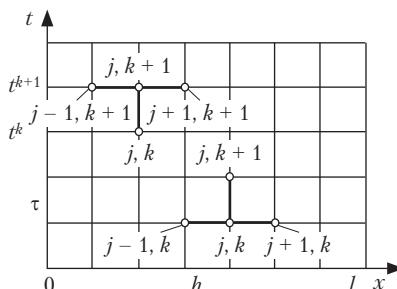


Рис. 5.1

На сетке (5.12) введем сеточные функции u_j^k , u_j^{k+1} , первая из которых известна, вторая — подлежит определению. Для ее определения в задаче (5.1)–(5.4) заменим (аппроксимируем) дифференциальные операторы отношением конечных разностей (см. раздел 3) и получим

$$\frac{\partial u}{\partial t} \Big|_j^k = \frac{u_j^{k+1} - u_j^k}{\tau} + O(\tau), \quad (5.13)$$

$$\frac{\partial^2 u}{\partial x^2} \Big|_j^k = \frac{u_{j+1}^k - 2u_j^k + u_{j-1}^k}{h^2} + O(h^2). \quad (5.14)$$

Подставляя (5.13), (5.14) в задачу (5.1)–(5.4), получим *явную конечно-разностную схему* для этой задачи в форме

$$\begin{aligned} \frac{u_j^{k+1} - u_j^k}{\tau} &= a^2 \frac{u_{j+1}^k - 2u_j^k + u_{j-1}^k}{h^2} + O(\tau + h^2), \\ k &= 0, 1, \dots, K-1, \quad j = 1, \dots, N-1, \\ u_0^k &= \varphi_0(t^k), \quad u_N^k = \varphi_l(t^k), \quad u_j^0 = \psi(x_j), \\ k &= 0, 1, \dots, K-1, \quad j = 1, \dots, N, \end{aligned} \quad (5.15)$$

где для каждого j -го уравнения все значения сеточной функции известны, за исключением одного — u_j^{k+1} , которое может быть определено явно из соотношений (5.15). В соотношения (5.15) краевые условия ($j = 0, j = N$) входят при значениях $j = 1$ и $j = N-1$, а начальное условие — при $k = 0$.

Если в (5.14) дифференциальный оператор по пространственной переменной аппроксимировать отношением конечных разностей на верхнем временном слое

$$\frac{\partial^2 u}{\partial x^2} \Big|_j^{k+1} = \frac{u_{j+1}^{k+1} - 2u_j^{k+1} + u_{j-1}^{k+1}}{h^2} + O(h^2), \quad (5.16)$$

то после подстановки (5.13), (5.16) в задачу (5.1)–(5.4) получим *неявную конечно-разностную схему* для этой задачи

$$\begin{aligned} \frac{u_j^{k+1} - u_j^k}{\tau} &= a^2 \frac{u_{j+1}^{k+1} - 2u_j^{k+1} + u_{j-1}^{k+1}}{h^2} + O(\tau + h^2), \\ k &= 0, 1, \dots, K-1, \quad j = 1, \dots, N-1, \\ u_0^{k+1} &= \varphi_0(t^{k+1}), \quad u_N^{k+1} = \varphi_l(t^{k+1}), \quad u_j^0 = \psi(x_j), \\ k &= 0, 1, \dots, K-1, \quad j = 1, \dots, N, \end{aligned} \quad (5.17)$$

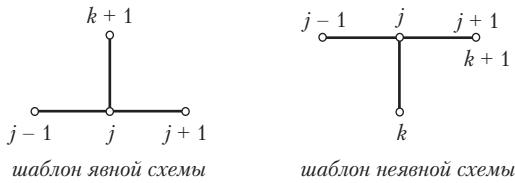


Рис. 5.2

Теперь сеточную функцию и u_j^{k+1} на верхнем временном слое

можно получить из решения системы линейных алгебраических уравнений (5.17) с трехдиагональной матрицей. Эта система в форме, пригодной для использования метода прогонки, имеет вид

$$\begin{cases} b_1 u_1^{k+1} + c_1 u_2^{k+1} = d_1, & j=1, \\ a_j u_{j-1}^{k+1} + b_j u_j^{k+1} + c_j u_{j+1}^{k+1} = d_j, & j=2, \dots, N-2, \\ a_{N-1} u_{N-2}^{k+1} + b_{N-1} u_{N-1}^{k+1} = d_{N-1}, & j=N-1, \end{cases}$$

где

$$a_1 = 0,$$

$$a_j = \sigma, j = 1, \dots, N-1;$$

$$b_j = -(1+2\sigma), j = 1, \dots, N-1;$$

$$c_j = \sigma, j = 1, \dots, N-1;$$

$$c_{N-1} = 0;$$

$$d_j = -u_j^k, j = 2, \dots, N-2;$$

$$d_1 = -\left(u_j^k + \sigma\varphi_0(t^{k+1})\right);$$

$$d_{N-1} = -\left(u_{N-1}^k + \sigma\varphi_l(t^{k+1})\right);$$

$$\sigma = \frac{a^2\tau}{h^2}.$$

Шаблоном конечно-разностной схемы называют ее геометрическую интерпретацию на конечно-разностной сетке.

На рис. 5.2 приведены шаблоны для явной (5.15) и неявной (5.17) конечно-разностных схем при аппроксимации задачи (5.1)–(5.4).

Явная конечно-разностная схема (5.15), записанная в форме

$$u_j^{k+1} = \sigma u_{j+1}^k + (1-2\sigma)u_j^k + \sigma u_{j-1}^k, \sigma = \frac{a^2\tau}{h^2}, \quad (5.18)$$

$$j = 1, \dots, N-1, \quad k = 0, 1, 2, \dots,$$

обладает тем достоинством, что решение на верхнем временном слое t^{k+1} получается сразу (без решения системы линейных алгебраических уравнений) по значениям сеточных функций на нижнем временном слое t^k , где решение известно. При $k = 0$ значения сеточной функции формируются из начального условия (5.4). Но эта же схема обладает существенным недостатком, поскольку она является условно устойчивой с условием $\sigma \frac{a^2\tau}{h^2} \leq \frac{1}{2}$, накладываемым на сеточные характеристики τ и h .

С другой стороны, неявная конечно-разностная схема (5.17), записанная в форме

$$a_j u_j^{k+1} + b_j u_j^{k+1} + c_j u_{j+1}^{k+1} = d_j, \quad j = 2, \dots, N-1, \quad k = 0, 1, 2, \dots, \quad (5.19)$$

приводит к необходимости решать систему линейных алгебраических уравнений, но зато эта схема абсолютно устойчива.

Проанализируем схемы (5.18), (5.19). Пусть точное решение, которое не известно, возрастает по времени, т.е. $u_j^{k+1} > u_j^k$. Тогда, в соответствии с явной схемой (5.18), разностное решение будет заниженным по сравнению с точным, так как u_j^{k+1} определяется по меньшим значениям сеточной функции на предыдущем временном слое, поскольку решение является возрастающим по времени. Для неявной схемы (5.19) на *возрастающем* решении, наоборот, решение завышено по сравнению с точным, поскольку оно определяется по значениям сеточной функции на верхнем временном слое. На *убывающем* решении картина изменяется противоположным образом: явная конечно-разностная схема завышает решения, а неявная — занижает, это проиллюстрировано рис. 5.3.

На основе этого анализа возникла идея о построении более точной неявно-явной конечно-разностной схемы с весами при пространственных конечно-разностных операторах, причем в случае измельчения шагов τ и h точное (неизвестное) решение может быть взято в *вилку* сколь угодно узкую, так как если явная и неявная схемы аппроксимируют дифференциальную задачу и эти схемы устойчивы, то при стремлении сеточных характеристик τ и h к нулю, решения по явной и неявной схемам стремятся к точному решению с разных сторон.

Рассмотрим неявно-явную схему с весами для простейшего уравнения теплопроводности

$$\frac{u_j^{k+1} - u_j^k}{\tau} = \theta a^2 \frac{u_{j+1}^{k+1} - 2u_j^{k+1} + u_{j-1}^{k+1}}{h^2} + (1-\theta) a^2 \frac{u_{j+1}^k - 2u_j^k + u_{j-1}^k}{h^2}, \quad (5.20)$$

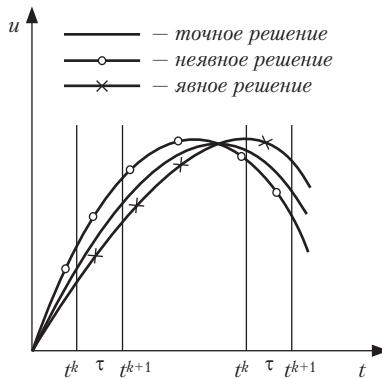


Рис. 5.3

где θ — вес неявной части конечно-разностной схемы, $(1 - \theta)$ — вес явной части, причем $0 \leq \theta \leq 1$. При $\theta = 1$ имеем полностью неявную схему, при $\theta = 0$ — полностью явную схему и при $\theta = 1/2$ — схему Кранка — Николсона.

Для схемы Кранка — Николсона порядок аппроксимации составляет $O(\tau^2 + h^2)$, что на один порядок по времени выше по сравнению с обычными явной или неявной схемами. Неявно-явная схема с весами (5.20) абсолютно устойчива при $1/2 \leq \theta \leq 1$ и условно устойчива с условием $\sigma \leq 1/2$ при $0 \leq \theta \leq 1/2$. Таким образом, получаем следующий вывод.

- Схема Кранка — Николсона (5.20) при $\theta = 1/2$ абсолютно устойчива и имеет второй порядок аппроксимации по времени и пространственной переменной x .

5.1.3. Аппроксимация граничных условий, содержащих производные

В задачах математической физики вообще и в задачах теплопроводности в частности граничные условия первого рода аппроксимируются точно в узлах на границе расчетной области. Граничные условия второго и третьего рода отличаются тем, что в них присутствует производная первого порядка искомой функции по пространственной переменной. Поэтому для замыкания конечно-разностной схемы необходима их аппроксимация. Простейшим вариантом является аппроксимация производных направленными разностями первого порядка:

$$\left. \frac{\partial u}{\partial x} \right|_{j=0}^{k+1} = \frac{u_1^{k+1} - u_0^{k+1}}{h} + O(h); \quad \left. \frac{\partial u}{\partial x} \right|_{j=N}^{k+1} = \frac{u_N^{k+1} - u_{N-1}^{k+1}}{h} + O(h).$$

Тогда в общем случае граничных условий третьего рода (5.7), (5.8) уравнения, связывающие значения искомой функции в двух крайних узлах разностной сетки, выглядят следующим образом:

$$\alpha \frac{u_1^{k+1} - u_0^{k+1}}{h} + \beta u_0^{k+1} = \varphi_0(t^{k+1})O(h),$$

$$\gamma \frac{u_N^{k+1} - u_{N-1}^{k+1}}{h} + \delta u_N^{k+1} = \varphi_l(t^{k+1})O(h).$$

Дополняя полученными уравнениями явную конечно-разностную аппроксимацию во внутренних узлах, получим явную разностную схему для третьей начально-краевой задачи (5.1), (5.4), (5.7), (5.8):

$$\alpha \frac{u_1^{k+1} - u_0^{k+1}}{h} + \beta u_0^{k+1} = \varphi_0(t^{k+1}),$$

$$\frac{u_j^{k+1} - u_j^k}{\tau} = a^2 \frac{u_{j+1}^k - 2u_j^k + u_{j-1}^k}{h^2},$$

$$\gamma \frac{u_N^{k+1} - u_{N-1}^{k+1}}{h} + \delta u_N^{k+1} = \varphi_l(t^{k+1}), \quad j = 1, \dots, N-1, \quad k = 0, 1, \dots, K-1.$$

В результате алгоритм перехода на новый временной слой t^{k+1} с использованием явной схемы можно представить в следующем виде:

$$u_j^{k+1} = \sigma \cdot u_{j+1}^k + (1 - 2\sigma)u_j^k + \sigma \cdot u_{j-1}^k, \quad \sigma = \frac{a^2 \tau}{h^2},$$

$$u_0^{k+1} = \frac{\alpha/h}{\beta - \alpha/h} u_1^{k+1} + \frac{\varphi_0(t^{k+1})}{\beta - \alpha/h},$$

$$u_N^{k+1} = \frac{\gamma/h}{\delta - \gamma/h} u_{N-1}^{k+1} + \frac{\varphi_l(t^{k+1})}{\delta - \gamma/h}, \quad j = 1, \dots, N-1.$$

Здесь сначала рассчитываются значения искомой функции во всех внутренних узлах на новом временном слое, а затем определяются значения на границах.

При использовании неявной конечно-разностной схемы получаем разностный аналог дифференциальной задачи:

$$b_0 u_0^{k+1} + c_0 u_1^{k+1} = d_0,$$

$$a_j u_{j-1}^{k+1} + b_j u_j^{k+1} + c_j u_{j+1}^{k+1} = d_j,$$

$$a_N u_{N-1}^{k+1} + b_N u_N^{k+1} = d_N, \quad j = 1, \dots, N-1,$$

где

$$\begin{aligned} b_0 &= \beta - \frac{\alpha}{h}, \quad c_0 = \frac{\alpha}{h}, \quad d_0 = \varphi_0(t^{k+1}), \\ a_N &= -\frac{\gamma}{h}, \quad b_N = \delta + \frac{\gamma}{h}, \quad d_N = \varphi_l(t^{k+1}), \\ a_j &= \sigma, \quad b_j = -(1+2\sigma), \quad c_j = \sigma, \quad d_j = -u_j^k, \quad \sigma = \frac{a^2\tau}{h^2}. \end{aligned}$$

В результате для получения решения на новом временном слое t^{k+1} решается система линейных алгебраических уравнений с трехдиагональной матрицей. Аналогичная картина имеет место и при использовании неявно-явной схемы с весами.

Принципиальной особенностью рассмотренного подхода является первый порядок аппроксимации граничных условий. Это означает, что порядок аппроксимации в граничных узлах ниже порядка аппроксимации во внутренних узлах расчетной области.

- При этом глобальный порядок аппроксимации во всей расчетной области равен *наименьшему* относительно всех узлов сетки порядку аппроксимации.

Одним из способов повышения порядка аппроксимации граничных условий является использование формул численного дифференцирования второго порядка:

$$\begin{aligned} \frac{\partial u}{\partial x}(0, t^{k+1}) &= \frac{-3u_0^{k+1} + 4u_1^{k+1} - u_2^{k+1}}{2h} + O(h^2); \\ \frac{\partial u}{\partial x}(l, t^{k+1}) &= \frac{-u_{N-2}^{k+1} + 4u_{N-1}^{k+1} - 3u_N^{k+1}}{2h} + O(h^2). \end{aligned}$$

В случае явной схемы алгоритм вычисления решения на новом временном слое при такой аппроксимации граничных условий не приобретает изменений. Если же используется неявная схема, то получающаяся при этом система линейных алгебраических уравнений теряет трехдиагональный вид (первое и последнее уравнение содержат три неизвестных). Этот недостаток легко устраним, так как путем линейной комбинации первого уравнения со вторым (последнего с предпоследним) можно добиться исключения третьего неизвестного из соответствующего уравнения. Однако при этом возможно нарушение диагонального преобладания матрицы и, следовательно, нарушение условий применимости метода прогонки.

Более эффективным является подход, позволяющий повысить порядок аппроксимации граничных условий без увеличения числа узлов в аппроксимационных соотношениях. Для иллюстрации этого подхода рассмотрим следующий пример.

Пример 5.1. Решить третью начально-краевую задачу для параболического уравнения, содержащего как конвективные члены (пропорциональные производной du/dx), так и источниковые члены, содержащие искомую функцию $u(x, t)$:

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} + b \frac{\partial u}{\partial x} + gu, \quad 0 < x < l, \quad t > 0; \quad (5.21)$$

$$\left\{ \begin{array}{l} \alpha \frac{\partial u(0, t)}{\partial x} + \beta u(0, t) = \varphi_0(t), \\ \gamma \frac{\partial u(l, t)}{\partial x} + \delta u(l, t) = \varphi_l(t), \end{array} \right. \quad t > 0; \quad (5.22)$$

$$\left\{ \begin{array}{l} u(x, 0) = \psi(x), \end{array} \right. \quad 0 \leq x \leq l; \quad (5.23)$$

$$\left\{ \begin{array}{l} u(x, 0) = \psi(x), \end{array} \right. \quad 0 \leq x \leq l; \quad (5.24)$$

Решение. Во внутренних узлах конечно-разностной сетки неявная конечно-разностная схема для уравнения (5.21) имеет вид

$$\begin{aligned} \frac{u_j^{k+1} - u_j^k}{\tau} = & \frac{a^2}{h^2} (u_{j+1}^{k+1} - 2u_j^{k+1} + u_{j-1}^{k+1}) + \\ & + \frac{b}{2h} (u_{j+1}^{k+1} - u_{j-1}^{k+1}) + gu_j^{k+1} + O(\tau + h^2), \quad j = 1, \dots, N-1. \end{aligned} \quad (5.25)$$

Если производные первого порядка в граничных условиях (5.22) и (5.23) аппроксимировать по схеме

$$\left. \frac{\partial u}{\partial x} \right|_{j=0}^{k+1} = \frac{u_1^{k+1} - u_0^{k+1}}{h} + O(h); \quad \left. \frac{\partial u}{\partial x} \right|_{j=N}^{k+1} = \frac{u_N^{k+1} - u_{N-1}^{k+1}}{h} + O(h),$$

то граничные условия аппроксимируются с первым порядком, и глобальный порядок будет равен первому порядку, несмотря на то, что во всех остальных узлах порядок аппроксимации по пространственным переменным равен двум. Для сохранения порядка аппроксимации, равного двум, в граничных узлах разложим на точном решении значение u_1^{k+1} в окрестности точки $x = 0$ в ряд Тейлора по переменной x до третьей производной включительно, а u_{N-1}^{k+1} — в аналогичный ряд в окрестности точки $x = l$.

Тогда в предположении, что функция $u(x, t)$ в граничных узлах имеет первые производные по времени и вторые — по x , получим

$$u_1^{k+1} = u(0 + h, t^{k+1}) = u_0^{k+1} + \left. \frac{\partial u}{\partial x} \right|_0^{k+1} h + \left. \frac{\partial^2 u}{\partial x^2} \right|_0^{k+1} \frac{h^2}{2} + O(h^3), \quad (5.26)$$

$$u_{N-1}^{k+1} = u(l - h, t^{k+1}) = u_N^{k+1} - \left. \frac{\partial u}{\partial x} \right|_N^{k+1} h + \left. \frac{\partial^2 u}{\partial x^2} \right|_N^{k+1} \frac{h^2}{2} + O(h^3). \quad (5.27)$$

Подставим сюда значения второй производной в граничных узлах, полученные из дифференциального уравнения (5.21):

$$\left. \frac{\partial^2 u}{\partial x^2} \right|_{j=0,N}^{k+1} = \left(\frac{1}{a^2} \cdot \frac{\partial u}{\partial t} - \frac{b}{a^2} \cdot \frac{\partial u}{\partial x} - \frac{g}{a^2} \cdot u \right)_{j=0,N}^{k+1},$$

и найдем из полученных выражений (5.26), (5.27) значения первой производной $\left. \frac{\partial u}{\partial x} \right|_{j=0,N}^{k+1}$ в граничных узлах с порядком $O(\tau + h^2)$:

$$\begin{aligned} \left. \frac{\partial u}{\partial x} \right|_0^{k+1} &= \frac{2a^2}{h(2a^2 - bh)} \cdot (u_1^{k+1} - u_0^{k+1}) - \frac{h}{2a^2 - bh} \cdot \left. \frac{\partial u}{\partial t} \right|_0^{k+1} + \frac{gh}{2a^2 - bh} \cdot u_0^{k+1} + \\ &+ O(h^2), \\ \left. \frac{\partial u}{\partial x} \right|_N^{k+1} &= \frac{2a^2}{h(2a^2 + bh)} \cdot (u_N^{k+1} - u_{N-1}^{k+1}) + \frac{h}{2a^2 + bh} \cdot \left. \frac{\partial u}{\partial t} \right|_N^{k+1} - \frac{gh}{2a^2 - bh} \cdot u_N^{k+1} + \\ &+ O(h^2). \end{aligned}$$

Подставляя $\left. \frac{\partial u}{\partial x} \right|_0^{k+1}$ в (5.22), а $\left. \frac{\partial u}{\partial x} \right|_N^{k+1}$ в (5.23) и аппроксимируя полу-

ченные соотношения в соответствующих граничных узлах, отмечая при этом, что

$$\left. \frac{\partial u}{\partial t} \right|_0^{k+1} = (u_0^{k+1} - u_0^k) / \tau + O(\tau), \quad \left. \frac{\partial u}{\partial t} \right|_N^{k+1} = (u_N^{k+1} - u_N^k) / \tau + O(\tau),$$

получим алгебраические уравнения для граничных узлов, в каждом из которых два неизвестных:

$$\begin{aligned} b_0 \cdot u_0^{k+1} + c_0 \cdot u_1^{k+1} &= d_0, \\ a_0 = 0, \quad b_0 = \frac{2a^2}{h} + \frac{h}{\tau} - gh - \frac{\beta}{\alpha}(2a^2 - bh), \quad c_0 = -\frac{2a^2}{h}, \\ d_0 = \frac{h}{\tau} u_0^k - \Phi_0(t^{k+1}) \frac{2a^2 - bh}{a}, \end{aligned} \quad (5.28)$$

$$\begin{aligned} a_N \cdot u_{N-1}^{k+1} + b_N \cdot u_N^{k+1} &= d_N, \\ a_N = -\frac{2a^2}{h}, \quad b_N = \frac{2a^2}{h} + \frac{h}{\tau} - gh + \frac{\delta}{\gamma}(2a^2 + bh), \end{aligned} \quad (5.29)$$

$$c_N = 0, \quad d_N = \frac{h}{\tau} u_N^k + \Phi_l(t^{k+1}) \frac{2a^2 + bh}{\gamma}.$$

Таким образом, (5.28) — конечно-разностная аппроксимация граничного условия третьего рода (5.22) на левой границе $x = 0$, а (5.29) — конечно-разностная аппроксимация граничного условия третьего рода (5.23) на правой границе $x = l$, сохраняющие тот же порядок аппроксимации, что и аппроксимация (5.25) дифференциального уравнения (5.21).

Приписывая к граничным конечно-разностным уравнениям (5.28), (5.29), каждое из которых содержит два значения сеточной функции, алгебраические уравнения (5.25), записанные в виде

$$\begin{aligned} a_j u_{j+1}^{k+1} + b_j u_j^{k+1} + c_j u_{j-1}^{k+1} &= d_j, \quad j = 1, \dots, N-1; \\ a_j &= -\left(\frac{a^2}{h^2} - \frac{b}{2h}\right); \quad b_j = \frac{2a^2}{h^2} + \frac{1}{\tau} - g; \\ c_j &= -\left(\frac{a^2}{h^2} + \frac{b}{2h}\right); \quad d_j = \frac{1}{\tau} u_j^k, \end{aligned} \quad (5.30)$$

получим систему линейных алгебраических уравнений с трехдиагональной матрицей, решаемую методом прогонки ($a_0 = 0$, $c_N = 0$):

$$A_j = -\frac{c_j}{b_j + a_j A_{j-1}}, \quad B_j = \frac{d_j - a_j B_{j-1}}{b_j + a_j A_{j-1}}, \quad j = 0, \dots, N; \quad (5.31)$$

$$\begin{aligned} u_N^{k+1} &= B_N, \\ u_j^{k+1} &= A_j u_{j+1}^{k+1} + B_j, \quad j = N-1, \dots, 0. \end{aligned} \quad (5.32)$$

Изложенный метод аппроксимации краевых условий, содержащих производные по пространственным переменным, не только повышает порядок аппроксимации, но и сохраняет консервативность конечно-разностной схемы. Это означает, что в конечно-разностной аппроксимации соблюдаются законы сохранения, на основе которых выведены дифференциальные соотношения задачи (5.21)–(5.24). Аналогичный подход можно осуществить в краевых задачах для дифференциальных уравнений любых типов.

5.2. Метод конечных разностей для решения уравнений гиперболического типа

5.2.1. Постановка задач для уравнений гиперболического типа

Классическим примером уравнения гиперболического типа является волновое уравнение, которое в области $0 < x < l$, $t > 0$ имеет вид

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}.$$

Данное уравнение описывает, в частности, процесс малых поперечных колебаний струны. В этом случае $u(x, t)$ — поперечные перемещения струны, a — скорость распространения малых возмущений в материале, из которого изготовлена струна.

Если концы струны движутся по заданным законам, т.е. на концах заданы перемещения или значения искомой функции, то *первая начально-краевая задача для волнового уравнения* имеет вид

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < l, \quad t > 0, \quad (5.33)$$

$$\left. \begin{array}{l} u(0, t) = \varphi_0(t), \\ u(l, t) = \varphi_l(t), \end{array} \right\} t > 0; \quad (5.34)$$

$$\left. \begin{array}{l} u(x, 0) = \psi_1(x), \\ \frac{\partial u(x, 0)}{\partial t} = \psi_2(x), \end{array} \right\} 0 \leq x \leq l, \quad (5.35)$$

$$\left. \begin{array}{l} u(x, 0) = \psi_1(x), \\ \frac{\partial u(x, 0)}{\partial t} = \psi_2(x), \end{array} \right\} 0 \leq x \leq l, \quad (5.36)$$

$$\left. \begin{array}{l} u(x, 0) = \psi_1(x), \\ \frac{\partial u(x, 0)}{\partial t} = \psi_2(x), \end{array} \right\} 0 \leq x \leq l, \quad (5.37)$$

причем если концы струны жестко закреплены, то $\varphi_0(t) = \varphi_l(t) = 0$. Как видно, в задачах для волнового уравнения, кроме начального распределения искомой функции, задается еще распределение начальной скорости перемещения.

Если на концах струны заданы значения силы, которая пропорциональна значениям производной перемещения по пространственной переменной (т.е. на концах заданы значения первых производных по переменной x), то ставится *вторая начально-краевая задача для волнового уравнения*:

$$\left. \begin{array}{l} \frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < l, \quad t > 0; \\ \frac{\partial u(0, t)}{\partial x} = \varphi_0(t), \quad t > 0; \\ \frac{\partial u(l, t)}{\partial x} = \varphi_l(t), \quad t > 0; \\ u(x, 0) = \psi_1(x), \quad 0 \leq x \leq l, \\ \frac{\partial u(x, 0)}{\partial t} = \psi_2(x), \quad 0 \leq x \leq l. \end{array} \right\}$$

В условиях, когда концы струны свободны, $\varphi_0(t) = \varphi_l(t) = 0$.

В условиях, когда концы закреплены *упруго*, т.е. на концевые заделки действуют силы, пропорциональные перемещениям, ставится *третья начально-краевая задача* для *волнового уравнения*:

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, & 0 < x < l, \quad t > 0; \\ \alpha \frac{\partial u(0,t)}{\partial x} + \beta u(0,t) = \varphi_0(t), & t > 0; \\ \gamma \frac{\partial u(l,t)}{\partial x} + \delta u(l,t) = \varphi_l(t), & t > 0; \\ u(x,0) = \psi_1(x), & 0 \leq x \leq l, \\ \frac{\partial u(x,0)}{\partial t} = \psi_2(x), & 0 \leq x \leq l. \end{cases}$$

Аналогично ставятся двумерные и трехмерные начально-краевые задачи для двумерного и трехмерного волновых уравнений.

5.2.2. Конечно-разностная аппроксимация уравнений гиперболического типа

Рассмотрим первую начально-краевую задачу для волнового уравнения (5.33)–(5.37). На пространственно-временной сетке (5.12) будем аппроксимировать дифференциальное уравнение (5.33) одной из следующих конечно-разностных схем:

$$\frac{u_j^{k+1} - 2u_j^k + u_j^{k-1}}{\tau^2} = a^2 \frac{u_{j+1}^k - 2u_j^k + u_{j-1}^k}{h^2} + O(\tau + h^2), \quad (5.38)$$

$$j = 1, \dots, N-1; \quad k = 1, 2, \dots$$

с шаблоном, приведенным на рис. 5.4, а) и

$$\frac{u_j^{k+1} - 2u_j^k + u_j^{k-1}}{\tau^2} = a^2 \frac{u_{j+1}^{k+1} - 2u_j^{k+1} + u_{j-1}^{k+1}}{h^2} + O(\tau + h^2), \quad (5.39)$$

$$j = 1, \dots, N-1; \quad k = 1, 2, \dots$$

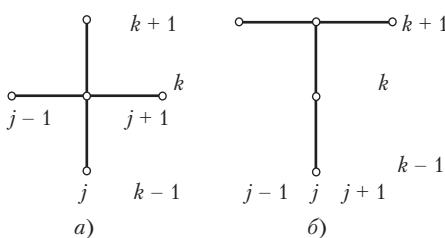


Рис. 5.4

с шаблоном на рис. 5.4, б.

Схема (5.38) является *явной*. С ее помощью, поскольку значения сеточных функций u_j^{k+1} , u_j^k и на нижних временных слоях должны быть известны, решение u_j^{k+1} , $j=1, \dots, N-l$, $k=1, 2, \dots$, определяется сразу. В соответствии с шаблоном для этой схемы порядок аппроксимации равен двум, как по пространственной, так и по временной переменной. При этом явная конечно-разностная схема (5.38) для волнового уравнения условно устойчива с условием $\sigma \frac{a^2\tau^2}{h^2} < 1$, накладываемым на сеточные характеристики τ и h .

Схема (5.39) является *неявной* и обладает абсолютной устойчивостью. Ее можно свести к системе линейных алгебраических уравнений с трехдиагональной матрицей, решаемой методом прогонки.

В обеих схемах необходимо знать значения u_j^{k+1} , u_j^k , $j = 1, \dots, N-1$, $k = 1, 2, \dots$, на нижних временных слоях. Для $k=1$ это делается следующим образом:

$$u_j^0 = \psi_1(x_j), \quad j = 0, 1, \dots, N, \quad (5.40)$$

где $\psi_1(x)$ — функция из начального условия (5.36). Для определения u_j^1 можно воспользоваться простейшей аппроксимацией второго начального условия (5.37):

$$\frac{u_j^1 - u_j^0}{\tau} = \psi_2(x_j).$$

Отсюда для искомых значений u_j^1 получаем выражение

$$u_j^1 = \psi_1(x_j) + \psi_2(x_j)\tau.$$

Недостатком такого подхода является первый порядок аппроксимации второго начального условия. Для повышения порядка аппроксимации воспользуемся следующей процедурой. Разложим u_j^1 в ряд Тейлора на точном решении по времени в окрестности $t = 0$:

$$u_j^1 = u(x_j, 0 + \tau) = u_j^0 + \left. \frac{\partial u}{\partial t} \right|_j \tau + \left. \frac{\partial^2 u}{\partial t^2} \right|_j \frac{\tau^2}{2} + O(\tau^3). \quad (5.41)$$

Для определения второй производной в выражении (5.41) воспользуемся исходным дифференциальным уравнением

$$\left. \frac{\partial^2 u}{\partial t^2} \right|_j^0 = a^2 \left. \frac{\partial^2 u}{\partial x^2} \right|_j^0 = a^2 \psi''_1(x_j).$$

В результате получим искомую сеточную функцию u_j^1 со вторым порядком точности:

$$u_j^1 = \psi_1(x_j) + \psi_2(x_j)\tau + a^2 \psi''_1(x_j) \frac{\tau^2}{2}.$$

После определения из начальных условий значений сеточных функций u_j^0 , u_j^1 на двух первых временных слоях вычислительный процесс продолжается согласно схемам (5.38) или (5.39). При этом аппроксимация краевых условий (5.34) и (5.35) производится аналогично тому, как это описывалось выше для уравнений параболического типа. Для иллюстрации этого этапа рассмотрим следующий пример.

Пример 5.2. Выписать явную конечно-разностную схему для третьей начально-краевой задачи:

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + b \frac{\partial u}{\partial x} + cu + f(x, t), & 0 < x < l, \quad t > 0; \\ \alpha \frac{\partial u(0, t)}{\partial x} + \beta u(0, t) = \varphi_0(t), & t > 0; \\ \gamma \frac{\partial u(l, t)}{\partial x} + \delta u(l, t) = \varphi_l(t), & t > 0; \\ u(x, 0) = \psi_1(x), & 0 \leq x \leq l; \\ \frac{\partial u(x, 0)}{\partial t} = \psi_2(x), & 0 \leq x \leq l. \end{cases}$$

Решение. Аппроксимируем дифференциальное уравнение по шаблону рис. 5.4, б:

$$\begin{aligned} \frac{u_j^{k+1} - 2u_j^k + u_j^{k-1}}{\tau^2} &= a^2 \frac{u_{j+1}^k - 2u_j^k + u_{j-1}^k}{h^2} + b \frac{u_{j+1}^k - u_{j-1}^k}{2h} + cu_j^k + f_j^k, \\ f_j^k &= f(x_j, t^k), \quad j = 1, \dots, N-1; \quad k = 1, 2, \dots. \end{aligned}$$

Границные условия аппроксимируем с первым порядком точности:

$$\alpha \frac{u_1^{k+1} - u_0^{k+1}}{h} + \beta u_0^{k+1} = \varphi_0(t^{k+1}), \quad \gamma \frac{u_N^{k+1} - u_{N-1}^{k+1}}{h} + \delta u_N^{k+1} = \varphi_l(t^{k+1}).$$

В результате переход на новый временной слой представляется следующим алгоритмом:

$$\begin{aligned} u_j^{k+1} &= u_{j+1}^k \left(\frac{a^2 \tau^2}{h^2} + b \frac{\tau^2}{2h} \right) + u_j^k \left(-2 \frac{a^2 \tau^2}{h^2} + 2 + c\tau^2 \right) + \\ &\quad + u_{j-1}^k \left(\frac{a^2 \tau^2}{h^2} - b \frac{\tau^2}{2h} \right) - u_{j-1}^k + \tau^2 f_j^k, \\ u_0^{k+1} &= \frac{\alpha/h}{\beta - \alpha/h} u_1^{k+1} + \frac{\varphi_0(t^{k+1})}{\beta - \alpha/h}, \\ u_N^{k+1} &= \frac{\gamma/h}{\delta - \gamma/h} u_{N-1}^{k+1} + \frac{\varphi_l(t^{k+1})}{\delta - \gamma/h}. \end{aligned}$$

Таким образом, сначала рассчитываются значения искомой функции во внутренних узлах на новом временном слое, после чего из аппроксимации граничных условий находятся значения функции u в крайних узлах.

Для окончательного замыкания вычислительного процесса определим, исходя из начальных условий, значения искомой функции на двух первых временных слоях t^0, t^1 . В начальный момент времени значения u_j определяются точно:

$$u_j^0 = \psi_1(x_j).$$

Если воспользоваться аппроксимацией первого порядка по времени, то, как было показано, получим

$$u_j^1 = \psi_1(x_j) + \psi_2(x_j)\tau.$$

Для повышения порядка аппроксимации разложим u_j^1 в ряд Тейлора на точном решении по времени в окрестности $t = 0$:

$$u_j^1 = u(x_j, 0 + \tau) = u_j^0 + \frac{\partial u}{\partial t} \Big|_j^\tau \tau + \frac{\partial^2 u}{\partial t^2} \Big|_j^\tau \frac{\tau^2}{2} + O(\tau^3),$$

где, согласно исходному уравнению,

$$\frac{\partial^2 u}{\partial t^2} \Big|_j^\tau = a^2 \frac{\partial^2 u}{\partial x^2} \Big|_j^\tau + b \frac{\partial u}{\partial x} \Big|_j^\tau + c u_j^k + f_j^k = a^2 \psi_1''(x_j) + b \psi_1'(x_j) + c \psi_1(x_j) + f_j^k.$$

Окончательно получаем

$$u_j^1 = \psi_1(x_j) + \psi_2(x_j)\tau + (a^2\psi_1''(x_j) + b\psi_1'(x_j) + c\psi_1(x_j) + f_j^k)\frac{\tau^2}{2}.$$

5.3. Метод конечных разностей для решения уравнений эллиптического типа

5.3.1. Постановка задач для уравнений эллиптического типа

Классическими примерами уравнений эллиптического типа являются *уравнение Пуассона*

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y)$$

и уравнение Лапласа

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

Здесь функция $u(x, y)$ имеет различный физический смысл. Это может быть стационарное, т.е. не зависящее от времени, распределение температуры, скорость потенциального (безвихревого) течения идеальной (без трения и теплопроводности) жидкости, распределение напряженностей электрического и магнитного полей, потенциала в силовом поле тяготения и т.п.

Если на границе $\Gamma = \Omega + \Gamma$ задана исходная функция, то соответствующая *первая краевая задача* для *уравнения Лапласа* или *Пуассона* называется *задачей Дирихле*:

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y), & (x, y) \in \Omega; \\ u(x, y)|_{\Gamma} = \varphi(x, y), & (x, y) \in \Gamma. \end{cases} \quad (5.42)$$

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y), & (x, y) \in \Omega; \\ \frac{\partial u(x, y)}{\partial n}|_{\Gamma} = \varphi(x, y), & (x, y) \in \Gamma. \end{cases} \quad (5.43)$$

Если на границе Γ задается нормальная производная искомой функции, то соответствующая *вторая краевая задача* называется *задачей Неймана* для *уравнения Лапласа* или *Пуассона*

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y), & (x, y) \in \Omega; \\ \frac{\partial u(x, y)}{\partial n}|_{\Gamma} = \varphi(x, y), & (x, y) \in \Gamma. \end{cases} \quad (5.44)$$

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y), & (x, y) \in \Omega; \\ \frac{\partial u(x, y)}{\partial n}|_{\Gamma} = \varphi(x, y), & (x, y) \in \Gamma. \end{cases} \quad (5.45)$$

Здесь n — направление внешней к границе Γ нормали.

Более приемлемой является координатная форма краевого условия (5.45):

$$\frac{\partial u}{\partial x} \cos(\mathbf{n}, \mathbf{i}) + \frac{\partial u}{\partial y} \cos(\mathbf{n}, \mathbf{j}) = \varphi(x, y),$$

где $\cos(\mathbf{n}, \mathbf{i})$, $\cos(\mathbf{n}, \mathbf{j})$ — направляющие косинусы внешнего вектора единичной нормали к границе Γ , \mathbf{i} и \mathbf{j} — орты базисных векторов.

Наконец, *третья краевая задача для уравнения Пуассона (Лапласа)* имеет вид

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y), & (x, y) \in \Omega; \\ \left. \frac{\partial u(x, y)}{\partial n} \right|_{\Gamma} + \alpha u|_{\Gamma} = \varphi(x, y), & (x, y) \in \Gamma. \end{cases}$$

5.3.2. Конечно-разностная аппроксимация задач для уравнений эллиптического типа

Рассмотрим краевую задачу для уравнений Пуассона или Лапласа (5.42), (5.43) в прямоугольнике $x \in [0, l_1]$, $y \in [0, l_2]$, на который наложим сетку

$$\omega_{h_1, h_2} = \{x_i = ih_1, \quad i = 0, 1, \dots, N_1; \quad y_j = jh_2, \quad j = 0, 1, \dots, N_2\}. \quad (5.46)$$

Введем сеточную функцию $u_{ij} = 0, 1, \dots, N_1, j = 0, 1, \dots, N_2$ и на сетке (5.46) аппроксимируем дифференциальную задачу во внутренних узлах с помощью отношения конечных разностей по следующей схеме:

$$\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h_1^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{h_2^2} + O(h_1^2 + h_2^2) = f(x_i, y_j), \quad (5.47)$$

$$i = 1, \dots, N_1 - 1, \quad j = 1, \dots, N_2 - 1.$$

Схема (5.47) на шаблоне, показанном на рис. 5.5, имеет второй порядок по переменным x и y , поскольку шаблон центрально симметричен.

Система линейных алгебраических уравнений (5.47) имеет пятидиагональный вид: каждое уравнение содержит пять неизвестных и при соответствующей нумерации переменных матрица имеет ленточную структуру. Решать ее можно различными методами линейной алгебры, например итерационными методами, методом матричной прогонки и т.п.

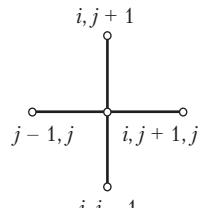


Рис. 5.5

Рассмотрим разностно-итерационный метод Либмана численного решения задачи Дирихле (5.42), (5.43). Для простоты изложения этого метода примем $h_1 = h_2 = h$, тогда из схемы (5.47) получим

$$\begin{aligned} u_{i,j}^{(k+1)} &= \frac{1}{4} \left[u_{i+1,j}^{(k)} + u_{i-1,j}^{(k)} + u_{i,j+1}^{(k)} + u_{i,j-1}^{(k)} - h^2 \cdot f_{i,j} \right], \\ f_{i,j} &= f(x_i, y_j), \quad i = 1, \dots, N_1 - 1, \quad j = 1, \dots, N_2 - 1, \end{aligned} \quad (5.48)$$

где k — номер итерации.

На каждой координатной линии, например $y_j = \text{const}$, $j = 1, \dots, N_2 - 1$, с помощью линейной интерполяции граничных значений $\varphi(x, y)$, иллюстрируемой рис. 5.6, определим значения $u_{i,j}^{(0)}$ на нулевой итерации. Эти значения подставим в (5.48) и получим распределение $u_{i,j}^{(1)}$ на первой итерации:

$$\begin{aligned} u_{i,j}^{(1)} &= \frac{1}{4} \left[u_{i+1,j}^{(0)} + u_{i-1,j}^{(0)} + u_{i,j+1}^{(0)} + u_{i,j-1}^{(0)} - h^2 \cdot f_{i,j} \right], \\ i &= 1, \dots, N_1 - 1, \quad j = 1, \dots, N_2 - 1. \end{aligned}$$

Это распределение $u_{i,j}^{(1)}$ снова подставим в (5.48), чтобы получить распределение $u_{i,j}^{(2)}$ и т.д. Процесс Либмана прекращается, когда

$$\|u^{(k+1)} - u^{(k)}\| \leq \varepsilon, \|u^{(k)}\| = \max_{i,j} |u_{i,j}^{(k)}|,$$

где ε — наперед заданная погрешность.

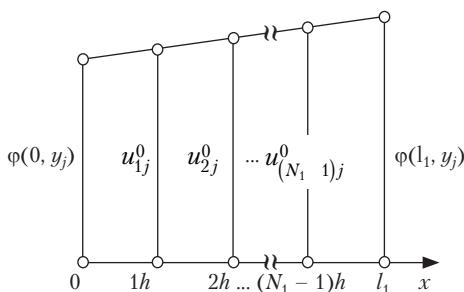


Рис. 5.6

При решении задач с граничными условиями второго и третьего рода наряду с аппроксимацией дифференциального уравнения производится также аппроксимация граничных условий. Этот вопрос подробно рассмотрен в разделе 5.1.3. Здесь в качестве примера приведем разностную схему, аппроксимирующую третью краевую задачу для уравнения Пуассона в прямоугольнике.

Пример 5.3. Решить краевую задачу для уравнения Пуассона

$$\begin{cases} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y), & x \in (0, l_1), \quad y \in (0, l_2); \\ \alpha_1 \frac{\partial u}{\partial x}(0, y) + \beta_1 u(0, y) = \varphi_1(y), \\ \alpha_2 \frac{\partial u}{\partial x}(l_1, y) + \beta_2 u(l_1, y) = \varphi_2(y), \\ \alpha_3 \frac{\partial u}{\partial x}(x, 0) + \beta_3 u(x, 0) = \varphi_3(y), \\ \alpha_4 \frac{\partial u}{\partial x}(x, l_2) + \beta_4 u(x, l_2) = \varphi_4(y). \end{cases}$$

Решение. Как и ранее, в прямоугольнике $x \in [0, l_1], y \in [0, l_2]$ построим сетку

$$\omega_{h_1, h_2} = \{x_i = ih_1, \quad i = 0, 1, \dots, N_1; \quad y_j = jh_2, \quad j = 0, 1, \dots, N_2\}.$$

На этой сетке аппроксимируем дифференциальную задачу во внутренних узлах по рассмотренной центрально-разностной схеме:

$$\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{h_1^2} + \frac{u_{i,j+1} - 2u_{i,j} + u_{i,j-1}}{h_2^2} = f(x_i, y_j),$$

$$i = 1, \dots, N_1 - 1, \quad j = 1, \dots, N_2 - 1.$$

Граничные условия аппроксимируем с первым порядком с помощью направленных разностей:

$$\begin{cases} \alpha_1 \frac{u_{1j} - u_{0j}}{h_1} + \beta_1 u_{0j} = \varphi_1(y_j), & j = 1, \dots, N_2 - 1, \\ \alpha_2 \frac{u_{N_1 j} - u_{N_1 - 1, j}}{h_1} + \beta_2 u_{N_1 j} = \varphi_2(y_j), & j = 1, \dots, N_2 - 1, \\ \alpha_3 \frac{u_{i1} - u_{i0}}{h_2} + \beta_3 u_{i0} = \varphi_3(x_i), & i = 1, \dots, N_1 - 1, \\ \alpha_4 \frac{u_{iN_2} - u_{iN_2 - 1}}{h_2} + \beta_4 u_{iN_2} = \varphi_4(x_i), & i = 1, \dots, N_1 - 1. \end{cases}$$

В результате получена система линейных алгебраических уравнений, содержащая $[(N_1 + 1)(N_2 + 1) - 4]$ уравнений относительно неизвестных $u_{ij} = 0, \dots, N_1, j = 0, \dots, N_2$, при этом угловые узлы с координатами $(0, 0), (0, N_2), (N_1, 0), (N_1, N_2)$ в вычислениях не участвуют. Как и в случае граничных условий первого рода, она имеет пятидиагональный вид и может быть решена, например, итерационным методом Либмана.

Замечание. Метод простых итераций для решения систем линейных алгебраических уравнений, возникающих при аппроксимации уравнения Пуассона (Лапласа), отличается довольно медленной сходимостью. Этот недостаток может стать существенным при использовании мелких сеток, когда число уравнений в системе становится большим. С более эффективными методами решения таких систем можно познакомиться, например, в монографии [8].

5.4. Метод конечных разностей решения многомерных задач математической физики. Методы расщепления

При численном решении многомерных задач математической физики исключительно важным является вопрос об экономичности используемых методов.

- Конечно-разностную схему называют экономичной, если число выполняемых операций (операций типа умножения) пропорционально числу узлов сетки.

В настоящее время известно значительное количество экономичных разностных схем численного решения многомерных задач математической физики, основанных на *расщеплении* пространственных дифференциальных операторов по координатным направлениям и использовании метода скалярной прогонки вдоль этих направлений.

Из экономичных конечно-разностных схем, получивших наибольшее распространение, в данном разделе рассматриваются схемы методов *переменных направлений* и *дробных шагов*. Будем называть их общим термином — *методы расщепления*.

Рассмотрим эти методы на примере задачи для двумерного уравнения параболического типа в прямоугольнике со сторонами l_1, l_2 и граничными условиями первого рода.

Для пространственно-временной области $\bar{G}_T = \bar{G} \times [0, T]$, $t \in [0, T]$, $\bar{G} = G + \Gamma$, $G = l_1 \times l_2$ рассмотрим следующую задачу:

$$\frac{\partial u}{\partial t} = a \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) + f(x, y, t), \quad x \in (0, l_1), \quad y \in (0, l_2), \quad t > 0; \quad (5.49)$$

$$u(x, 0, t) = \varphi_1(x, t), \quad x \in [0, l_1], \quad t > 0; \quad (5.50)$$

$$u(x, l_2, t) = \varphi_2(x, t), \quad x \in [0, l_1], \quad t > 0; \quad (5.51)$$

$$u(0, y, t) = \varphi_3(y, t), \quad y \in [0, l_2], \quad t > 0; \quad (5.52)$$

$$u(l_1, y, t) = \varphi_4(y, t), \quad y \in [0, l_2], \quad t > 0; \quad (5.53)$$

$$u(x, y, t) = \psi(x, y), \quad x \in [0, l_1], \quad y \in [0, l_2]. \quad (5.54)$$

Введем пространственно-временную сетку с шагами h_1, h_2, τ соответственно по переменным x, y, t :

$$\begin{aligned} \omega_{h_1, h_2}^{\tau} = \{ & x_i = ih_1, \quad i = 0, 1, \dots, I; \quad y_j = jh_2, \quad j = 0, 1, \dots, J; \\ & t^k = k\tau, \quad k = 0, 1, 2, \dots \}, \end{aligned} \quad (5.55)$$

и на этой сетке будем аппроксимировать дифференциальную задачу (5.49)–(5.54) методом конечных разностей.

5.4.1. Метод переменных направлений

В схеме метода переменных направлений (МПН), как и во всех методах расщепления, шаг по времени разбивается на несколько шагов, число которых равно числу независимых пространственных переменных (в двумерном случае – на два). На каждом дробном временном слое один из пространственных дифференциальных операторов аппроксимируется неявно (по соответствующему координатному направлению осуществляются скалярные прогонки), а остальные – явно. На следующем дробном шаге другой дифференциальный оператор аппроксимируется неявно, а остальные – явно и т.д. В двумерном случае схема метода переменных направлений для задачи (5.49)–(5.54) имеет вид

$$\begin{aligned} \frac{u_{ij}^{k+1/2} - u_{ij}^k}{\tau / 2} = & \frac{a}{h_1^2} \left(u_{i+1j}^{k+1/2} - 2u_{ij}^{k+1/2} + u_{i-1j}^{k+1/2} \right) + \\ & + \frac{a}{h_2^2} \left(u_{ij+1}^k - 2u_{ij}^k + u_{ij-1}^k \right) + f_{ij}^{k+1/2}, \end{aligned} \quad (5.56)$$

$$\begin{aligned} \frac{u_{ij}^{k+1} - u_{ij}^{k+1/2}}{\tau / 2} = & \frac{a}{h_1^2} \left(u_{i+1j}^{k+1/2} - 2u_{ij}^{k+1/2} + u_{i-1j}^{k+1/2} \right) + \\ & + \frac{a}{h_2^2} \left(u_{ij+1}^{k+1} - 2u_{ij}^{k+1} + u_{ij-1}^{k+1} \right) + f_{ij}^{k+1/2}. \end{aligned} \quad (5.57)$$

В подсхеме (5.56) на первом дробном шаге оператор ad_2/dx_2 аппроксимируется неявно, а оператор ad_2/dy_2 — явно (в результате весь конечно-разностный оператор по переменной y переходит в правые части, поскольку u_{ij}^k известно). С помощью скалярных прогонок в количестве, равном числу $(J - 1)$, в направлении переменной x получаем распределение сеточной функции $u_{ij}^{k+1/2}$, $i = 1, \dots, I - 1, j = 1, \dots, J - 1$ на первом временном полуслое $t^k + 1/2 = t^k + \tau/2$.

В подсхеме (5.57) оператор $a\partial^2/\partial y^2$ аппроксимируется неявно на верхнем временном слое $t^{k+1} = (k + 1)\tau$, а оператор $a\partial^2/\partial x^2$ — явно в момент времени $t^{k+1/2} = t^k + \tau/2$. С помощью скалярных прогонок в направлении переменной y в количестве, равном числу $(I - 1)$, получаем распределение сеточной функции $u_{ij}^{k+1/2}$, $i = 1, \dots, I - 1, j = 1, \dots, J - 1$ на втором полуслое. Шаблон схемы МПН представлен на рис. 5.7.

Можно показать, что в двумерном случае схема МПН абсолютно устойчива. К достоинствам метода переменных направлений можно отнести высокую точность, поскольку метод имеет второй порядок точности по времени. К недостаткам можно отнести условную устойчивость при числе пространственных

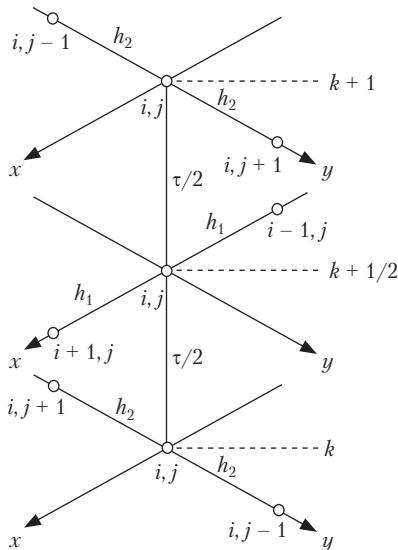


Рис. 5.7

переменных больше двух. Кроме этого, МПН условно устойчив в задачах со смешанными производными уже в двумерном случае.

5.4.2. Метод дробных шагов

В отличие от МПН метод дробных шагов (МДШ) использует только неявные конечно-разностные операторы, что делает его абсолютно устойчивым в задачах, не содержащих смешанных производных. Он обладает довольно значительным запасом устойчивости и в задачах со смешанными производными.

Для задачи (5.49)–(5.54) схема МДШ имеет вид

$$\frac{u_{ij}^{k+1/2} - u_{ij}^k}{\tau} = \frac{a}{h_1^2} \left(u_{i+1,j}^{k+1/2} - 2u_{ij}^{k+1/2} + u_{i-1,j}^{k+1/2} \right) + \frac{f_{ij}^k}{2}, \quad (5.58)$$

$$\frac{u_{ij}^{k+1} - u_{ij}^{k+1/2}}{\tau} = \frac{a}{h_2^2} \left(u_{ij+1}^{k+1} - 2u_{ij}^{k+1} + u_{ij-1}^{k+1} \right) + \frac{f_{ij}^{k+1}}{2}. \quad (5.59)$$

С помощью чисто неявной подсхемы (5.58) осуществляются скалярные прогонки в направлении оси x в количестве, равном $(J - 1)$, в результате чего получается сеточная функция $u_{ij}^{k+1/2}$.

На втором дробном шаге по времени с помощью подсхемы (5.59) осуществляются скалярные прогонки в направлении оси y в количестве, равном $(I - 1)$, в результате чего получается сеточная функция u_{ij}^{k+1} . Шаблон схемы МДШ приведен на рис. 5.8.

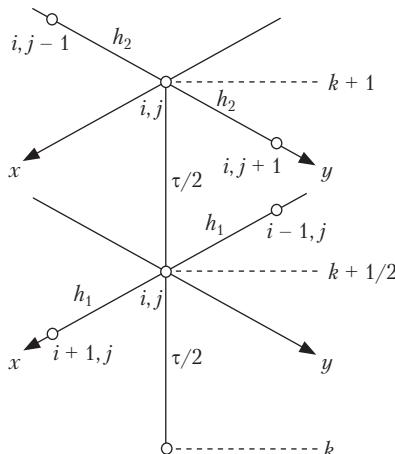


Рис. 5.8

Схема МДШ имеет порядок сходимости $O(\tau + h^2)$, т.е. первый порядок по времени и второй — по переменным x и y .

В литературе МДШ называют также методом *покоординатного расщепления* и *локально-одномерным* методом.

К достоинствам схемы МДШ можно отнести простоту в алгоритмизации и программировании и абсолютную устойчивость с большим запасом устойчивости даже для задач, содержащих смешанные производные. В то же время недостатки МДШ следующие: на каждом дробном шаге достигается частичная аппроксимация (полная аппроксимация достигается на последнем дробном шаге, т.е. имеет место суммарная аппроксимация), схема имеет первый порядок точности по времени.

5.4.3. Методы расщепления численного решения эллиптических задач

Для стационарных многомерных задач математической физики искомая функция в задаче (5.49)–(5.54) не зависит от времени и, следовательно, уравнение (5.49) становится уравнением эллиптического типа (нет производной du/dt), т.е. уравнением Лапласа или Пуассона, а поскольку и начальное условие (5.54) отсутствует, рассмотренные выше методы приходится несколько видоизменить. Однако если при решении задач для уравнений Лапласа или Пуассона используется метод установления, то стационарное уравнение трансформируется в нестационарное уравнение (5.49), являющееся уже уравнением параболического типа, с введением однородного начального условия (5.54), т.е. $\psi(x, y) = 0$. В этом случае все рассмотренные выше методы применяются без изменения.

Методы расщепления напрямую можно применять также и к решению стационарных задач, заменив номер временного слоя номером итерации. В качестве начального приближения на нулевой итерации можно использовать линейную интерполяцию краевых условий так, как это делалось в разностно-итерационном методе Либмана (см. раздел 5.3).

5.5. Основные понятия, связанные с конечно-разностной аппроксимацией дифференциальных задач

К основным понятиям, связанным с методом конечных разностей, относятся аппроксимация, порядок аппроксимации, устойчивость, сходимость, порядок сходимости. Определим каждое из этих понятий.

5.5.1. Апроксимация и порядок аппроксимации

Запишем дифференциальное уравнение в операторной форме

$$LU = f,$$

где L — дифференциальный оператор. В качестве L могут выбираться, например, такие операторы, как

диффузионный	волевой	лапласиан
$\frac{\partial}{\partial t} - a^2 \frac{\partial^2}{\partial x^2},$	$\frac{\partial^2}{\partial t^2} - a^2 \frac{\partial^2}{\partial x^2},$	$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2},$

U — искомая функция, удовлетворяющая дифференциальной задаче; f — входные данные, т.е. начальные и краевые условия, правые части и т.п. Операторная форма $(LU)_h = f_h$ описывает дифференциальную задачу в узлах сетки, а операторная форма $L_h U_h = f_h$ описывает конечно-разностную схему на точном решении U , т.е. если в конечно-разностной схеме вместо сеточных значений сеточной функции подставлены точные (неизвестные) значения искомой функции. Для перечисленных дифференциальных операторов L конечно-разностные операторы L_h имеют вид

$$\frac{\Delta}{\tau} = a^2 \frac{\Delta^2}{h^2}, \quad \frac{\Delta^2}{\tau^2} = a^2 \frac{\Delta^2}{h^2}, \quad \frac{\Delta^2}{h_1^2} + \frac{\Delta^2}{h_2^2},$$

где Δ — конечно-разностный оператор, например,

$$\begin{aligned}\Delta u_i &= u_{i+1} - u_i, \\ \Delta^2 u_i &= u_{i+1} - 2u_i + u_{i-1}.\end{aligned}$$

Наконец, операторная форма конечно-разностной схемы имеет вид

$$L_h u_h = f_h. \quad (5.60)$$

Отметим, что помимо аппроксимации основного дифференциального уравнения конечно-разностная схема содержит также аппроксимации начальных и краевых условий.

Введем норму сеточной функции, например, с помощью выражения

$$\|u^{(k)}\| = \max_j |u_j^{(k)}|, \quad k = 0, 1, 2, \dots. \quad (5.61)$$

Конечно-разностная схема (5.60) аппроксимирует дифференциальную задачу на точном решении, если какая-либо норма разности $(LU)_h$ и $L_h U_h$ стремится к нулю при $\tau \rightarrow 0, h \rightarrow 0$:

$$\|(LU)_h - L_h U_h\| \xrightarrow{\tau, h \rightarrow 0} 0. \quad (5.62)$$

Конечно-разностная схема (5.60) аппроксимирует дифференциальную задачу на точном решении с порядком p по времени и порядком q по пространственной переменной, если какая-либо норма разности $\|(LU)_h - L_h U_h\|$ удовлетворяет равенству

$$\|(LU)_h - L_h U_h\| = O(\tau^p + h^q). \quad (5.63)$$

Таким образом, если конечно-разностная схема аппроксимирует дифференциальную задачу, то речь идет о *близости* дифференциального и конечно-разностного операторов в узлах сетки.

5.5.2. Устойчивость

Пусть в конечно-разностной схеме (5.60) входные данные f_h получили возмущения и стали равными \tilde{f}_h . Тогда сеточная функция u_h также получит возмущение и станет равной \tilde{u}_h .

- Конечно-разностная схема (5.60) устойчива по входным данным, если найдется ограниченная константа K , не зависящая от сеточных характеристик τ, h и входных данных f_h , такая, что выполняется неравенство

$$\|u_h - \tilde{u}_h\| \leq K \|f_h - \tilde{f}_h\|. \quad (5.64)$$

Таким образом, понятие устойчивости интерпретируется следующим образом: конечно-разностная схема устойчива, если для малых возмущений входных данных (начально-краевых условий и правых частей) конечно-разностная схема обеспечивает малые возмущения сеточной функции u_h , т.е. решение с помощью конечно-разностной схемы находится под контролем входных данных.

Если входные данные f_h содержат только начальные условия, или только краевые условия, или только правые части, то говорят об *устойчивости* соответственно по *начальным условиям*, по *краевым условиям* или по *правым частям*.

- Конечно-разностная схема (5.60) *абсолютно (безусловно) устойчива*, если неравенство (5.64) выполняется при любых значениях сеточных характеристик τ и h , т.е. на шаги сетки не накладывается никаких ограничений.

- Конечно-разностная схема (5.60) *условно устойчива*, если неравенство (5.64) выполняется для сеточных характеристик τ и h , на которые накладываются определенные ограничения.

5.5.3. Сходимость и порядок сходимости

- Решение u_h , полученное с помощью конечно-разностной схемы (5.60), сходится к точному решению U , если какая-либо норма разности $\|U - u_h\|$ стремится к нулю при стремлении к нулю сеточных характеристик τ, h

$$\|U - u_h\| \xrightarrow[\tau, h \rightarrow 0]{} 0. \quad (5.65)$$

- Конечно-разностная схема (5.60) имеет p -й порядок сходимости (порядок точности) по времени и q -й порядок сходимости по пространственной переменной, если какая-либо норма разности $\|U - u_h\|$ удовлетворяет равенству

$$\|U - u_h\| = O(\tau^p + h^q). \quad (5.66)$$

Таким образом, порядок сходимости (порядок точности) характеризует близость конечно-разностного и точного (неизвестного) решения.

При численном решении задач математической физики в общем случае необходимо исследовать и аппроксимацию, и устойчивость, и сходимость. Однако следующая теорема утверждает, что достаточно исследовать аппроксимацию и устойчивость, и, в случае положительного ответа, сходимость будет обеспечена.

Теорема. Если конечно-разностная схема (5.60) аппроксирует на точном решении дифференциальную задачу с p -м порядком по времени и q -м порядком по пространственной переменной и эта схема устойчива, то решение с помощью этой конечно-разностной схемы сходится к решению дифференциальной задачи с p -м порядком по времени и q -м порядком по пространственной переменной.

Задачи

Решение предлагаемых задач связано с достаточно большим объемом вычислительной работы и требует использования компьютерной техники. Предполагается, что читатель владеет навыками программирования и будет разрабатывать свое собственное программное обеспечение. Ряд задач содержит параметры a, b, c , которые наряду с сеточными параметрами следует задавать во входном потоке. Такой подход дает возможность проведения параметрического анализа. Для всех задач приводятся точные аналитические решения, что, с одной стороны, является эффективным средством тестирования разработанного читателем программного обеспе-

чения, а с другой — позволяет оценить и проанализировать погрешность численного решения.

5.1. Используя явную и неявную конечно-разностные схемы, а также схему Кранка — Николсона, решить начально-краевую задачу для дифференциального уравнения параболического типа. В различные моменты времени вычислить погрешность численного решения путем сравнения результатов с приведенным в задании аналитическим решением $U(x, t)$. Исследовать зависимость погрешности от сеточных параметров τ, h .

$$1. \frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2}, \quad a > 0,$$

$$u(0, t) = 0,$$

$$u(1, t) = 0,$$

$$u(x, 0) = \sin(2\pi x).$$

Аналитическое решение: $U(x, t) = \exp(-4\pi^2 at) \sin(\pi x)$.

$$2. \frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2}, \quad a > 0,$$

$$u(0, t) = 0,$$

$$u(1, t) = 0,$$

$$u(x, 0) = x + \sin(2\pi x).$$

Аналитическое решение: $U(x, t) = x + \exp(-4\pi^2 at) \sin(\pi x)$.

$$3. \frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2}, \quad a > 0,$$

$$u(0, t) = \exp(-at),$$

$$u(\pi, t) = -\exp(-at),$$

$$u(x, 0) = \cos x.$$

Аналитическое решение: $U(x, t) = \exp(-at) \cos x$.

$$4. \frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2}, \quad a > 0,$$

$$u_x(0, t) = \exp(-at),$$

$$u_x(\pi, t) = -\exp(-at),$$

$$u(x, 0) = \sin x.$$

Аналитическое решение: $U(x, t) = \exp(-at) \sin x$.

5. $\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + \sin(\pi x),$

$$u(0, t) = 0,$$

$$u(1, t) = 0,$$

$$u(x, 0) = 0.$$

Аналитическое решение: $U(x, t) = \frac{1}{\pi^2} (1 - \exp(-\pi^2 t)) \sin(\pi x).$

6. $\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + \cos x (\cos t + \sin t),$

$$u(0, t) = \sin t,$$

$$u_x(\pi/2, t) = -\sin t,$$

$$u(x, 0) = 0.$$

Аналитическое решение: $U(x, t) = \sin t \cos x.$

7. $\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + 0.5 \exp(-0.5t) \cos x,$

$$u(0, t) = \exp(-0.5t),$$

$$u_x(\pi, t) = -\exp(-0.5t),$$

$$u(x, 0) = \sin x.$$

Аналитическое решение: $U(x, t) = \exp(-0.5t) \sin x.$

8. $\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + cu, \quad a > 0, \quad c < 0,$

$$u_x(0, t) = \exp((c-a)t),$$

$$u(\pi/2, t) = \exp((c-a)t),$$

$$u(x, 0) = \sin x.$$

Аналитическое решение: $U(x, t) = \exp((c-a)t) \sin(\pi x).$

9. $\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial u}{\partial x}, \quad a > 0, \quad b > 0,$

$$u_x(0, t) - u(0, t) = \exp(-at)(\cos(bt) + \sin(bt)),$$

$$u_x(\pi, t) - u(\pi, t) = \exp(-at)(\cos(bt) + \sin(bt)),$$

$$u(x, 0) = \cos x.$$

Аналитическое решение: $U(x,t) = \exp(-at) \cos(x+bt)$.

$$10. \frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial u}{\partial x} + cu, \quad a > 0, \quad b > 0, \quad c < 0.$$

$$u_x(0,t) + u(0,t) = \exp((c-a)t)(\cos(bt) + \sin(bt)),$$

$$u_x(\pi,t) + u(\pi,t) = -\exp((c-a)t)(\cos(bt) + \sin(bt)),$$

$$u(x,0) = \sin x.$$

Аналитическое решение: $U(x,t) = \exp((c-a)t) \sin(x+bt)$.

5.2. Используя явную схему *крест* и неявную схему, решить начально-краевую задачу для дифференциального уравнения гиперболического типа. В различные моменты времени вычислить погрешность численного решения путем сравнения результатов с приведенным в задании аналитическим решением $U(x, t)$. Исследовать зависимость погрешности от сеточных параметров τ, h .

$$1. \frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2},$$

$$u(0,t) = -\sin(at),$$

$$u(\pi,t) = \sin(at),$$

$$u(x,0) = \sin x,$$

$$u_t(x,0) = -a \cos x.$$

Аналитическое решение: $U(x,t) = \sin(x-at)$.

$$2. \frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2},$$

$$u_x(0,t) - u(0,t) = 0,$$

$$u_x(\pi,t) - u(\pi,t) = 0,$$

$$u(x,0) = \sin x + \cos x,$$

$$u_t(x,0) = -a(\sin x + \cos x).$$

Аналитическое решение: $U(x,t) = \sin(x-at) + \cos(x+at)$.

$$3. \frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} - 3u,$$

$$u(0,t) = \sin(2t),$$

$$u(\pi,t) = -\sin(2t),$$

$$\begin{aligned} u(x, 0) &= 0, \\ u_t(x, 0) &= 2 \cos x. \end{aligned}$$

Аналитическое решение: $U(x, t) = \cos x \sin(2t)$.

$$\begin{aligned} 4. \quad \frac{\partial^2 u}{\partial t^2} &= \frac{\partial^2 u}{\partial x^2} - 5u, \\ u_x(0, t) - 2u(0, t) &= 0, \\ u_x(1, t) - 2u(1, t) &= 0, \\ u(x, 0) &= \exp(2x), \\ u_t(x, 0) &= 0. \end{aligned}$$

Аналитическое решение: $U(x, t) = \exp(2x) \cos t$.

$$\begin{aligned} 5. \quad \frac{\partial^2 u}{\partial t^2} &= 2 \frac{\partial^2 u}{\partial x^2} + 4 \frac{\partial u}{\partial x}, \\ u(0, t) &= 0, \\ u(\pi, t) &= 0, \\ u(x, 0) &= 0, \\ u_t(x, 0) &= \exp(-x) \sin x. \end{aligned}$$

Аналитическое решение: $U(x, t) = 0.5 \exp(-x) \sin x \sin(2t)$.

$$\begin{aligned} 6. \quad \frac{\partial^2 u}{\partial t^2} &= a^2 \frac{\partial^2 u}{\partial x^2} + 2 \frac{\partial u}{\partial x} - 2u, \\ u(0, t) &= \cos(2t), \\ u(\pi/2, t) &= 0, \\ u(x, 0) &= \exp(-x) \cos x, \\ u_t(x, 0) &= 0. \end{aligned}$$

Аналитическое решение: $U(x, t) = 0.5 \exp(-x) \cos x \cos(2t)$.

$$\begin{aligned} 7. \quad \frac{\partial^2 u}{\partial t^2} + 2 \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} + 2 \frac{\partial u}{\partial x} - 3u, \\ u(0, t) &= \exp(-t) \cos(2t), \\ u(\pi/2, t) &= 0, \\ u(x, 0) &= \exp(-x) \cos x, \\ u_t(x, 0) &= -\exp(-x) \cos x. \end{aligned}$$

Аналитическое решение: $U(x,t) = \exp(-t-x) \cos x \cos(2t)$.

8. $\frac{\partial^2 u}{\partial t^2} + 2 \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + 2 \frac{\partial u}{\partial x} - 3u,$

$$u(0,t) = 0,$$

$$u(\pi,t) = 0,$$

$$u(x,0) = 0,$$

$$u_t(x,0) = 2 \exp(-x) \sin x.$$

Аналитическое решение: $U(x,t) = \exp(-t-x) \sin x \sin(2t)$.

9. $\frac{\partial^2 u}{\partial t^2} + 3 \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial u}{\partial x} - u + \sin x \exp(-t),$

$$u(0,t) = \exp(-t),$$

$$u(\pi,t) = -\exp(-t),$$

$$u(x,0) = \cos x,$$

$$u_t(x,0) = -\cos x.$$

Аналитическое решение: $U(x,t) = \exp(-t) \cos x$.

10. $\frac{\partial^2 u}{\partial t^2} + 3 \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial u}{\partial x} - u - \cos x \exp(-t),$

$$u_x(0,t) = \exp(-t),$$

$$u_x(\pi,t) = -\exp(-t),$$

$$u(x,0) = \sin x,$$

$$u_t(x,0) = -\sin x.$$

Аналитическое решение: $U(x,t) = \exp(-t) \sin x$.

5.3. Используя центрально-разностную схему и метод Либмана, решить краевую задачу для дифференциального уравнения эллиптического типа. Вычислить погрешность численного решения путем сравнения результатов с приведенным в задании аналитическим решением $U(x, y)$. Исследовать зависимость погрешности от сеточных параметров h_x, h_y .

1. $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$

$$u(0,y) = 0,$$

$$u(1,y) = 1 + y,$$

$$\begin{aligned} u(x, 0) &= x, \\ u(x, 1) &= 1 + x. \end{aligned}$$

Аналитическое решение: $U(x, y) = x + y$.

2. $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$

$$\begin{aligned} u(0, y) &= 0, \\ u(1, y) &= 1 - y^2, \\ u_y(x, 0) &= 0, \\ u(x, 1) &= x^2 - 1. \end{aligned}$$

Аналитическое решение: $U(x, y) = x^2 - y^2$.

3. $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$

$$\begin{aligned} u(0, y) &= \cos y, \\ u(1, y) &= e \cos y, \\ u_y(x, 0) &= 0, \\ u_y(x, \pi/2) &= -\exp(x). \end{aligned}$$

Аналитическое решение: $U(x, y) = \exp(x) \cos y$.

4. $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$

$$\begin{aligned} u_x(0, y) &= \exp(y), \\ u_x(\pi, y) &= -\exp(y), \\ u(x, 0) &= \sin x, \\ u(x, 1) &= e \sin x. \end{aligned}$$

Аналитическое решение: $U(x, y) = \sin x \exp(y)$.

5. $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -u,$

$$\begin{aligned} u_x(0, y) &= \cos y, \\ u_x(1, y) - u(1, y) &= 0, \end{aligned}$$

$$\begin{aligned} u(x, 0) &= x, \\ u(x, \pi/2) &= 0. \end{aligned}$$

Аналитическое решение: $U(x, y) = x \cos y$.

$$6. \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -u,$$

$$\begin{aligned} u(0, y) &= 0, \\ u(\pi/2, y) &= y, \\ u_y(x, 0) &= \sin x, \\ u_y(x, 1) - u(x, 1) &= 0. \end{aligned}$$

Аналитическое решение: $U(x, y) = y \sin x$.

$$7. \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -2u,$$

$$\begin{aligned} u(0, y) &= \cos y, \\ u(\pi/2, y) &= 0, \\ u(x, 0) &= \cos x, \\ u(x, \pi/2) &= 0. \end{aligned}$$

Аналитическое решение: $U(x, y) = \cos x \cos y$.

$$8. \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -2 \frac{\partial u}{\partial x} - 3u,$$

$$\begin{aligned} u(0, y) &= \cos y, \\ u(\pi/2, y) &= 0, \\ u(x, 0) &= \exp(-x) \cos x, \\ u(x, \pi/2) &= 0. \end{aligned}$$

Аналитическое решение: $U(x, y) = \exp(-x) \cos x \cos y$.

$$9. \quad \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -2 \frac{\partial u}{\partial x} - 3u,$$

$$\begin{aligned} u(0, y) &= \exp(-y) \cos y, \\ u(\pi/2, y) &= 0, \\ u(x, 0) &= \cos x, \\ u(x, \pi/2) &= 0. \end{aligned}$$

Аналитическое решение: $U(x, y) = \exp(-y) \cos x \cos y$.

$$10. \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -2 \frac{\partial u}{\partial x} - 2 \frac{\partial u}{\partial y},$$

$$u(0, y) = \exp(-y) \cos y,$$

$$u(\pi/2, y) = 0,$$

$$u(x, 0) = \exp(-x) \cos x,$$

$$u(x, \pi/2) = 0.$$

Аналитическое решение: $U(x, y) = \exp(-x - y) \cos x \cos y$.

5.4. Используя схемы переменных направлений и дробных шагов, решить двумерную начально-краевую задачу для дифференциального уравнения параболического типа. В различные моменты времени вычислить погрешность численного решения путем сравнения результатов с приведенным в задании аналитическим решением $U(x, t)$. Исследовать зависимость погрешности от сеточных параметров τ, h_x, h_y .

$$1. \frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + a \frac{\partial^2 u}{\partial y^2}, \quad a > 0,$$

$$u(0, y, t) = \cos(\mu_2 y) \exp[-(\mu_1^2 + \mu_2^2)at],$$

$$u(\pi, y, t) = (-1) \cos(\mu_2 y) \exp[-(\mu_1^2 + \mu_2^2)at],$$

$$u(x, 0, t) = \cos(\mu_1 x) \exp[-(\mu_1^2 + \mu_2^2)at],$$

$$u(x, \pi, t) = (-1)^{\mu_1} \cos(\mu_1 x) \exp[-(\mu_1^2 + \mu_2^2)at],$$

$$u(x, y, 0) = \cos(\mu_1 x) \cos(\mu_2 y).$$

$$\text{а) } \mu_1 = 1, \mu_2 = 1; \quad \text{б) } \mu_1 = 2, \mu_2 = 1; \quad \text{в) } \mu_1 = 1, \mu_2 = 2.$$

Аналитическое решение: $U(x, y, t) = \cos(\mu_1 x) \cos(\mu_2 y) \exp[-(\mu_1^2 + \mu_2^2)at]$.

$$2. \frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + a \frac{\partial^2 u}{\partial y^2}, \quad a > 0,$$

$$u(0, y, t) = \cos(\mu_2 y) \exp[-(\mu_1^2 + \mu_2^2)at],$$

$$u((\pi/2)\mu_1, y, t) = 0,$$

$$u(x, 0, t) = \cos(\mu_1 x) \exp[-(\mu_1^2 + \mu_2^2)at],$$

$$u(x, (\pi/2)\mu_2, t) = 0,$$

$$u(x, y, 0) = \cos(\mu_1 x) \cos(\mu_2 y).$$

а) $\mu_1 = 1, \mu_2 = 1$; б) $\mu_1 = 2, \mu_2 = 1$; в) $\mu_1 = 1, \mu_2 = 2$.

Аналитическое решение: $U(x, y, t) = \cos(\mu_1 x) \cos(\mu_2 y) \exp[-(\mu_1^2 + \mu_2^2) at]$.

3. $\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + a \frac{\partial^2 u}{\partial y^2}, \quad a > 0,$

$$u(0, y, t) = \operatorname{ch}(y) \exp(-3at),$$

$$u(\pi/4, y, t) = 0,$$

$$u(x, 0, t) = \cos(2x) \exp(-3at),$$

$$u(x, \ln 2, t) = (5/4) \cos(2x) \exp(-3at),$$

$$u(x, y, 0) = \cos(2x) \operatorname{ch}(y).$$

Аналитическое решение: $U(x, y, t) = \cos(2x) \operatorname{ch}(y) \exp(-3at)$.

4. $\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + a \frac{\partial^2 u}{\partial y^2}, \quad a > 0,$

$$u(0, y, t) = \operatorname{ch}(y) \exp(-3at),$$

$$u(\pi/2, y, t) = 0,$$

$$u(x, 0, t) = \cos(2x) \exp(-3at),$$

$$u_y(x, \ln 2, t) = (3/4) \cos(2x) \exp(-3at),$$

$$u(x, y, 0) = \cos(2x) \operatorname{ch}(y).$$

Аналитическое решение: $U(x, y, t) = \cos(2x) \operatorname{ch}(y) \exp(-3at)$.

5. $\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + a \frac{\partial^2 u}{\partial y^2}, \quad a > 0,$

$$u(0, y, t) = \operatorname{ch}(y) \exp(-3at),$$

$$u(\pi/2, y, t) = -\operatorname{sh}(y) \exp(-3at),$$

$$u_y(x, 0, t) = \cos(2x) \exp(-3at),$$

$$u(x, \ln 2, t) = (3/4) \cos(2x) \exp(-3at),$$

$$u(x, y, 0) = \cos(2x) \operatorname{sh}(y).$$

Аналитическое решение: $U(x, y, t) = \cos(2x) \operatorname{sh}(y) \exp(-3at)$.

6. $\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + a \frac{\partial^2 u}{\partial y^2}, \quad a > 0,$

$$u(0, y, t) = \operatorname{sh}(y) \exp(-3at),$$

$$u_x(\pi/4, y, t) = -2\operatorname{sh}(y) \exp(-3at),$$

$$u_y(x, 0, t) = \cos(2x) \exp(-3at),$$

$$u(x, \ln 2, t) = (3/4)\cos(2x) \exp(-3at),$$

$$u(x, y, 0) = \cos(2x) \operatorname{sh}(y).$$

Аналитическое решение: $U(x, y, t) = \cos(2x) \operatorname{sh}(y) \exp(-3at).$

7. $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - xy \sin t,$

$$u(0, y, t) = 0,$$

$$u(1, y, t) = y \cos t,$$

$$u(x, 0, t) = 0,$$

$$u(x, 1, t) = x \cos t,$$

$$u(x, y, 0) = xy.$$

Аналитическое решение: $U(x, y, t) = xy \cos t.$

8. $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - xy \sin t,$

$$u(0, y, t) = 0,$$

$$u(1, y, t) - u_x(1, y, t) = 0,$$

$$u(x, 0, t) = 0,$$

$$u(x, 1, t) - u_y(x, 1, t) = 0,$$

$$u(x, y, 0) = xy.$$

Аналитическое решение: $U(x, y, t) = xy \cos t.$

9. $\frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial^2 u}{\partial y^2} + \sin x \sin y [\mu \cos \mu t + (a+b) \sin \mu t],$

$$u(0, y, t) = 0,$$

$$u(\pi/2, y, t) = \sin y \sin(\mu t),$$

$$u(x, 0, t) = 0,$$

$$u(x, \pi, t) = -\sin y \sin(\mu t),$$

$$u(x, y, 0) = 0.$$

$$\text{а) } a=1, \ b=1, \ \mu=1; \quad \text{б) } a=2, \ b=1, \ \mu=1;$$

$$\text{в) } a=1, \ b=2, \ \mu=1; \quad \text{г) } a=1, \ b=1, \ \mu=2.$$

Аналитическое решение: $U(x,y,t) = \sin x \sin y \sin(\mu t)$.

$$10. \frac{\partial u}{\partial t} = a \frac{\partial^2 u}{\partial x^2} + b \frac{\partial^2 u}{\partial y^2} + \sin x \sin y [\mu \cos \mu t + (a+b) \sin \mu t],$$

$$u(0, y, t) = 0,$$

$$u(\pi, y, t) = -\sin y \sin(\mu t),$$

$$u(x, 0, t) = 0,$$

$$u(x, \pi, t) = -\sin y \sin(\mu t),$$

$$u(x, y, 0) = 0.$$

$$\text{а) } a=1, \ b=1, \ \mu=1; \quad \text{б) } a=2, \ b=1, \ \mu=1;$$

$$\text{в) } a=1, \ b=2, \ \mu=1; \quad \text{г) } a=1, \ b=1, \ \mu=2.$$

Аналитическое решение: $U(x,y,t) = \sin x \sin y \sin(\mu t)$.

Литература

1. *Вержбицкий, В. М.* Численные методы (линейная алгебра и нелинейные уравнения) : учеб. пособие для вузов. — М. : Высшая школа, 2000. — 266 с.
2. *Демидович, Б. П.* Основы вычислительной математики / Б. П. Демидович, И. Ф. Марон. — М. : Наука, 1970. — 664 с.
3. *Деммель, Дж.* Вычислительная линейная алгебра. Теория и приложения. — М. : Мир, 2001. — 430 с.
4. *Калиткин, Н. Н.* Численные методы. — М. : Наука, 1976. — 512 с.
5. *Каханер, Д., Моулер, К., Нэш, С.* Численные методы и программное обеспечение / Д. Каханер, К. Моулер, С. Нэш. — М. : Мир, 1998. — 575 с.
6. *Косарев, В. И.* 12 лекций по вычислительной математике (вводный курс) : учеб. пособие для вузов. — М. : Изд-во МФТИ, 2000. — 224 с.
7. *Пиругов, У. Г.* Численные методы : учеб. пособие. — М. : МАИ, 1998.
8. *Пиругов, У. Г.* Численные методы : учеб. пособие для студ. втузов. — 3-е изд., испр. — М. : Дрофа, 2004. — 224 с.
9. Сборник задач по методам вычислений : учеб. пособие для вузов / под ред. П. И. Монастырного. — М. : Физматлит, 1994. — 320 с.

Оглавление

Предисловие	3
Введение.....	4
Глава 1. Численные методы алгебры	7
1.1. Основные понятия линейной алгебры.....	8
1.2. Основные трудности решения систем линейных уравнений. Классификация методов решения.....	13
1.3. Метод исключения Гаусса	16
1.4. Вычисление определителя и обратной матрицы методом исключения.....	18
1.5. Метод прогонки для решения систем линейных уравнений с трехдиагональной матрицей.....	19
1.6. Итерационные методы решения линейных уравнений	23
1.7. Решение нелинейных уравнений. Два этапа отыскания корня.....	30
1.8. Метод половинного деления.....	33
1.9. Метод простой итерации нахождения корней нелинейных уравнений	34
1.10. Метод Ньютона и некоторые его модификации.....	38
1.11. Метод секущих	41
1.12. Метод парабол	42
1.13. Методы нахождения корней систем нелинейных уравнений. Ускорение сходимости по Эйткену.....	43
1.14. Введение в проблему собственных значений	48
1.15. Итерационный метод вращения для нахождения собственных значений.....	50
1.16. Метод Данилевского для построения характеристического многочлена матрицы	53
1.17. Метод интерполяции для построения характеристического многочлена.....	55
1.18. Численные методы оптимизации	56
1.19. Численные методы отыскания безусловного экстремума функции одной переменной.....	57

1.20. Численные методы отыскания безусловного экстремума функции многих переменных.....	59
1.21. Численные методы отыскания условного экстремума	65
1.22. Линейное программирование	69
Глава 2. Приближение функций.	
Численное интегрирование и дифференцирование.....	71
2.1. Постановка задачи об аппроксимации функций.....	71
2.2. Интерполяция.....	73
2.3. Интерполяционный многочлен в форме Лагранжа	74
2.4. Интерполяционный многочлен в форме Ньютона.....	77
2.5. Погрешность и сходимость интерполяции.....	80
2.6. Интерполяция сплайнами	84
2.7. Приближение методом наименьших квадратов.....	87
2.8. Постановка задачи численного дифференцирования	90
2.9. Дифференцирование интерполяционного многочлена Ньютона	91
2.10. Безразностные формулы численного дифференцирования для равноотстоящих узлов.....	94
2.11. Применение ряда Тейлора для численного дифференцирования.....	96
2.12. Постановка задачи численного интегрирования.....	97
2.13. Формулы численного интегрирования	98
2.14. Метод Рунге – Ромберга – Ричардсона повышения порядков точности.....	103
2.15. Метод статистических испытаний	104
Глава 3. Численные методы решения обыкновенных дифференциальных уравнений.....	125
3.1. Основы теории обыкновенных дифференциальных уравнений	125
3.2. Понятие о методе конечных разностей. Порядок точности разностной схемы.....	130
3.3. Метод Эйлера. Метод Эйлера с пересчетом.....	133
3.4. Метод Рунге – Кутты	136
3.5. Многошаговый метод Адамса	138
3.6. Неявные схемы. Понятие о жестких системах	140
3.7. Постановка краевой задачи для обыкновенного дифференциального уравнения.....	144
3.8. Метод стрельбы	144
3.9. Конечно-разностный метод решения краевых задач	146
3.10. Интегральные уравнения	147

Глава 4. Численные методы решения уравнений в частных производных	151
4.1. Некоторые сведения из теории уравнений в частных производных.....	151
4.2. Основные понятия метода сеток.	
Задача Дирихле для уравнения Лапласа.....	159
4.3. Явные и неявные разностные схемы.....	165
4.4. Аппроксимация, устойчивость, сходимость разностных схем. Основные понятия.....	167
4.5. Примеры неустойчивых разностных схем.....	169
4.6. Практические правила исследования устойчивости	173
4.7. Спектральный признак устойчивости.....	174
4.8. Принцип максимума.....	175
4.9. Метод гармоник Фурье исследования устойчивости разностных схем.....	175
4.10. Применение метода гармоник Фурье для исследования устойчивости разностных схем в уравнениях переноса.....	177
4.11. Применение метода гармоник Фурье в задачах исследования устойчивости разностных схем для уравнения теплопроводности.....	180
4.12. Метод переменных направлений	181
4.13. Математические и физические основы метода установления	187
4.14. Разностные схемы метода установления	191
4.15. Методы сквозного счета	196
4.16. Метод прямых.....	201
4.17. Метод характеристик	204
4.18. Метод конечных элементов	207
Литература	216
Именной указатель	217

СБОРНИК ЗАДАЧ

Предисловие	221
Глава 1. Численные методы линейной алгебры	223
1.1. Численные методы решения систем линейных алгебраических уравнений	223
1.1.1. Метод Гаусса	223
1.1.2. Метод прогонки.....	233
1.1.3. Нормы векторов и матриц.....	236
1.1.4. Итерационные методы решения СЛАУ	240

1.2. Численные методы решения задач на собственные значения и собственные векторы матриц	246
1.2.1. Основные определения и спектральные свойства матриц.....	246
1.2.2. Метод вращений Якоби численного решения задач на собственные значения и собственные векторы матриц.....	248
1.2.3. Частичная проблема собственных значений и собственных векторов матрицы. Степенной метод.....	253
1.2.4. QR-алгоритм нахождения собственных значений матриц.....	255
<i>Задачи</i>	262
<i>Ответы</i>	271
Глава 2. Нелинейные уравнения и системы нелинейных уравнений	280
2.1. Решение нелинейных уравнений	280
2.2. Решение систем нелинейных уравнений.....	286
<i>Задачи</i>	296
<i>Ответы</i>	297
Глава 3. Приближение функций.	
Численное дифференцирование и интегрирование.....	299
3.1. Интерполяция.....	299
3.2. Метод наименьших квадратов	306
3.3. Численное дифференцирование	309
3.4. Численное интегрирование.....	312
<i>Задачи</i>	316
<i>Ответы</i>	331
Глава 4. Численные методы решения обыкновенных дифференциальных уравнений	335
4.1. Численные методы решения задачи Коши	335
4.1.1. Задача Коши для одного обыкновенного дифференциального уравнения.....	335
4.1.2. Одношаговые методы.....	336
4.1.3. Решение задачи Коши для системы обыкновенных дифференциальных уравнений	341
4.1.4. Решение задачи Коши для ОДУ второго и более высоких порядков.....	343
4.1.5. Решение дифференциальных уравнений с запаздывающим аргументом	350
4.1.6. Многошаговые методы. Метод Адамса.....	352
4.2. Численные методы решения краевой задачи для обыкновенных дифференциальных уравнений.....	355

4.2.1. Метод стрельбы.....	356
4.2.2. Конечно-разностный метод решения краевой задачи	358
Задачи	361
Глава 5. Численное решение дифференциальных уравнений с частными производными	376
5.1. Численное решение уравнений параболического типа. Понятие о методе конечных разностей. Основные определения и конечно-разностные схемы.....	376
5.1.1. Постановка задач для уравнений параболического типа	376
5.1.2. Понятие о методе конечных разностей. Применение метода конечных разностей к решению уравнений параболического типа	378
5.1.3. Аппроксимация граничных условий, содержащих производные	382
5.2. Метод конечных разностей для решения уравнений гиперболического типа	387
5.2.1. Постановка задач для уравнений гиперболического типа	387
5.2.2. Конечно-разностная аппроксимация уравнений гиперболического типа	389
5.3. Метод конечных разностей для решения уравнений эллиптического типа	393
5.3.1. Постановка задач для уравнений эллиптического типа	393
5.3.2. Конечно-разностная аппроксимация задач для уравнений эллиптического типа	394
5.4. Метод конечных разностей решения многомерных задач математической физики. Методы расщепления	397
5.4.1. Метод переменных направлений	398
5.4.2. Метод дробных шагов	400
5.4.3. Методы расщепления численного решения эллиптических задач.....	401
5.5. Основные понятия, связанные с конечно-разностной аппроксимацией дифференциальных задач.....	401
5.5.1. Аппроксимация и порядок аппроксимации.....	402
5.5.2. Устойчивость	403
5.5.3. Сходимость и порядок сходимости.....	404
Задачи	404
Литература	416
Оглавление	417

Покупайте наши книги:

В офисе издательства «ЮРАЙТ»:

111123, г. Москва, ул. Плеханова, д. 4а,
тел.: (495) 744-00-12, e-mail: sales@urait.ru, www.urait.ru

В логистическом центре «ЮРАЙТ»:

140053, Московская область, г. Котельники, мкр. Ковровый, д. 37,
тел.: (495) 744-00-12, e-mail: sales@urait.ru, www.urait.ru

В интернет-магазине «ЮРАЙТ»: www.urait-book.ru,
e-mail: order@urait-book.ru, тел.: (495) 742-72-12

Для закупок у Единого поставщика в соответствии
с Федеральным законом от 21.07.2005 № 94-ФЗ обращаться
по тел.: (495) 744-00-12, e-mail: sales@urait.ru, vuz@urait.ru

**Новые издания и дополнительные материалы доступны
в электронной библиотечной системе «Юрайт»
biblio-online.ru**

Учебное издание

ЧИСЛЕННЫЕ МЕТОДЫ

Учебник и практикум для академического бакалавриата

Под редакцией *У. Г. Пирумова*

Формат 60×90¹/₁₆.
Гарнитура «Petersburg». Печать цифровая.
Усл. печ. л. 26,31. Тираж 100 экз. Заказ №

ООО «Издательство Юрайт»

111123, г. Москва, ул. Плеханова, д. 4а.
Тел.: (495) 744-00-12. E-mail: izdat@urait.ru, www.urait.ru