

Министерство образования и науки Российской Федерации
Федеральное агентство по образованию
Ярославский государственный университет им. П.Г. Демидова

В.Н. Матвеев

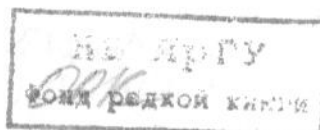
Методы вычислений

Учебное пособие

*Рекомендовано
Научно-методическим советом университета
для студентов, обучающихся по специальности
и направлению Математика*

Ярославль 2007

275480



УДК 519.6
ББК В 192.1я73
М33

*Рекомендовано
Редакционно-издательским советом университета
в качестве учебного издания. План 2007 года*

Рецензенты:

канд. физ.-мат. наук П.А. Корнилов; кафедра прикладной математики
и вычислительной техники Ярославского государственного технического университе-
та.

М33 **Матвеев, В.Н.** Методы вычислений: учеб. пособие / В.Н.Матвеев; Яросл. гос.
ун-т. — Ярославль: ЯрГУ 2007. — 156 с.
ISBN 978-5-8397-0523-4

В учебном пособии представлены основные темы курса "Методы вычислений", ко-
торый автор течение ряда лет читает в Ярославском государственном университете.
Большинство тем сопровождаются лабораторными работами, в которых предлагается
реализовать изучаемые в лекционном курсе вычислительные методы и алгоритмы.

Предназначено для студентов, обучающихся по специальности 010100 Математика
(дисциплина "Методы вычислений", блок ОПД) и направлению 510100 Математика (дис-
циплина "Методы вычислений", блок ОПД), очной формы обучения.

Библиогр.: 13 назв.

УДК 519.6
ББК В 192.1я73

ISBN 978-5-8397-0523-4

© Ярославский государственный
университет, 2007
© В.Н. Матвеев, 2007

БИБЛИОТЕКА ЯрГУ
ОСНОВНОЙ ФОНД

Оглавление

Основные понятия	5
§ 1. Точные методы решения линейных систем. Понятие обусловленности	9
§ 2. Частичная проблема собственных значений. Итерационные методы	13
§ 3. Итерационные методы решения линейных систем	18
§ 4. Итерационные методы решения нелинейных уравнений и систем	28
§ 5. Интерполяционный многочлен Лагранжа	35
§ 6. Минимизация остаточного члена. Многочлены Чебышева	38
§ 7. Разделенные разности. Интерполяционный многочлен в форме Ньютона	43
§ 8. Интерполирование с кратными узлами.	47
§ 9. Ортогональные многочлены и их свойства	51
§ 10. Численное дифференцирование	58
§ 11. Численное интегрирование. Квадратурные формулы Ньютона — Котеса	60
§ 12. Квадратурные формулы Гаусса	67
§ 13. Наилучшее приближение в гильбертовом пространстве	72
§ 14. Наилучшее равномерное приближение	77
§ 15. Численные методы решения задачи Коши для обыкновенных дифференциальных уравнений и систем	82
§ 16. Разностные методы решения краевых задач для обыкновенных дифференциальных уравнений 2-го порядка. Понятие аппроксимации, устойчивости и сходимости разностных схем	89
§ 17. Методы решения краевых задач 2-го порядка. Методы прогонки, метод стрельбы	92

§ 18. Принцип максимума для разностных уравнений 2-го порядка. Оценка решений первой краевой задачи ...	97
§ 19. Разностные уравнения	101
§ 20. Собственные значения простейшего разностного оператора	107
§ 21. Разностные методы решения краевых задач для уравнения теплопроводности	111
§ 22. Аппроксимация, устойчивость и сходимость разностной схемы с весами для уравнения теплопроводности	116
§ 23. Разностные схемы для уравнения колебаний струны	123
§ 24. Разностные схемы для уравнения Пуассона	125
Лабораторная работа №1	129
Лабораторная работа №2	131
Лабораторная работа №3	135
Лабораторная работа №4	138
Лабораторная работа №5	141
Лабораторная работа №6	144
Лабораторная работа №7	147
Лабораторная работа №8	151
Литература	155

Основные понятия

Понятие о погрешности. При решении тех или иных прикладных задач важно не только построить алгоритм решения поставленной задачи, но и оценить погрешность полученного решения. Погрешность принято разделять на неустранимую, вычислительную и погрешность метода.

Неустраняемая погрешность связана с тем, что при составлении математической модели, которая описывает исследуемый процесс, приходится обычно прибегать к некоторым упрощающим предположениям, пренебрегая величинами, влияние которых предполагается несущественным. Далее, при реальном изучении поставленной задачи входные параметры задачи являются результатом измерений или оценок, выполненных с определенной погрешностью. Поэтому результат решения поставленной задачи будет получен с некоторой погрешностью, которую и называют неустранимой.

Вычислительная погрешность связана с тем, что вычисления выполняются с конечным числом разрядов. При арифметических операциях на компьютере число записывается в форме с плавающей точкой, т.е. в виде AEB , где A – мантисса, а B – порядок. Например, число 103,54 записывается в виде $0,10354E+03$. Для записи мантиссы отводится конечное число разрядов, определяемое типом компьютера и операционной системы. Но при умножении двух чисел, имеющих мантиссу длины n , в результате получается число с мантиссой длины больше n . Если для записи мантиссы числа используются k разрядов и длина мантиссы больше k , то число округляется в соответствии с правилами округления: если значение $k+1$ разряда больше 5, то значение k -го разряда увеличивается на 1, в противном случае значения больших разрядов отбрасываются. Таким образом, в арифметических операциях значения младшего разряда есть результат округления. Но нетрудно заметить, что погрешность в k -м разряде сомножителей приводит к тому, что в произведении с погрешностью может определяться уже $k-1$ разряд. Следовательно, при большом числе вычислений погрешность, связанная с округлением в младшем разряде, может постепенно переходить в старшие разряды, искажая тем самым результат вычислений. Таким образом, вычислительная погрешность возникает как следствие арифметических операций с конечной разрядной сеткой.

Наконец, погрешность метода возникает в результате применения алгоритма приближенного решения поставленной задачи. Изучению таких алгоритмов и оценке погрешностей и посвящен этот курс.

Напомним основные понятия, используемые в дальнейшем изложении.

Пространство R будем называть линейным, если в нем определена операция сложения, ассоциативная и коммутативная, операция умножения на число и существует нулевой элемент \emptyset , такой, что $x + \emptyset = \emptyset + x = x$. Элементы x_1, x_2, \dots, x_n называются линейно независимыми, если равенство

$$c_1x_1 + c_2x_2 + \dots + c_nx_n = \emptyset$$

возможно тогда и только тогда, когда $c_1 = c_2 = \dots = c_n = 0$.

Пространство называется нормированным, если каждому элементу x ставится в соответствие действительное число $\|x\|$, такое, что

- 1) $\|x\| \geq 0$, причем $\|x\| = 0$ тогда и только тогда, когда $x = \emptyset$,
- 2) $\|kx\| = |k|\|x\|$ для любого комплексного k ,
- 3) $\|x + y\| \leq \|x\| + \|y\|$.

Последнее свойство называется неравенством треугольника. Пространство называется строго нормированным, если неравенство треугольника превращается в равенство тогда и только тогда, когда $y=kx$, $k \geq 0$.

Пространство называется метрическим, если для любых x и y , принадлежащих пространству, определено число, называемое расстоянием $\rho(x, y)$, такое, что $\rho(x, y) = \rho(y, x)$, причем если $\rho(x, y) = 0$, то $x=y$. Кроме того, имеет место неравенство треугольника

$$\rho(x, y) \leq \rho(x, z) + \rho(z, y).$$

В нормированном пространстве $\rho(x, y) = \|x - y\|$.

Говорят, что в пространстве H определено скалярное произведение, если для любых x и y из H поставлено в соответствие комплексное число (x, y) , такое, что

- 1) $(x, y) = \overline{(y, x)}$,
- 2) $(c_1x + c_2y, z) = c_1(x, z) + c_2(y, z)$,
- 3) $(x, x) \geq 0$. Если $(x, x) = 0$, то $x = \emptyset$.

Пространство, в котором определено скалярное произведение, называют гильбертовым. Если $(x, y) = (y, x)$, то его называют вещественным. В гильбертовом пространстве норма определяется естественным образом: $\|x\| = \sqrt{(x, x)}$.

Норма вектора. Норма матрицы

В R^n для вектора $x = (x_1, x_2, \dots, x_n)^T$ используются следующие нормы.

Первая векторная норма $\|x\|_1 = \max_i |x_i|$.

Первую векторную норму часто обозначают $\|x\|_c$.

Вторая векторная норма $\|x\|_2 = \sum_{i=1}^n |x_i|$.

Третья, или евклидова, норма $\|x\|_3 = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{(x, x)}$.

Согласно определению, норма оператора A , действующего из пространства H_1 в пространство H_2 , есть наименьшая из констант C , для которой $\|Ax\|_2 \leq C\|x\|_1$, где $\|\cdot\|_1$ — норма в пространстве H_1 , а $\|\cdot\|_2$ — норма в пространстве H_2 . Из определения очевидно следует неравенство

$$\|Ax\|_2 \leq \|A\| \|x\|_1.$$

Таким образом,

$$\|A\| = \sup_x \frac{\|Ax\|_2}{\|x\|_1}.$$

Поэтому норма матрицы определяется в зависимости от того, какая норма выбрана в пространстве R^n . Покажем, что в случае первой векторной нормы

$$\|A\|_1 = \max_i \sum_{j=1}^n |a_{i,j}|.$$

В самом деле, пусть первая норма вектора $x = (x_1, x_2, \dots, x_n)_T$ достигается на компоненте x_p : $\|x\|_1 = \max_i |x_i| = |x_p|$. Тогда для вектора $y = Ax$ имеем:

$$\|y\|_1 \leq \max_i \sum_{j=1}^n |a_{i,j}| |x_j| \leq (\max_i \sum_{j=1}^n |a_{i,j}|) |x_p|.$$

Если введена вторая векторная норма, то

$$\|A\|_2 = \max_j \sum_{i=1}^n |a_{i,j}|.$$

Действительно, для вектора $y = Ax = (y_1, y_2, \dots, y_n)$ имеем $\|y\|_2 = |y_1| + \dots + |y_n|$. Следовательно,

$$\|y\|_2 = \sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \sum_{i=1}^n \sum_{j=1}^n |a_{ij}| |x_j| = \sum_{j=1}^n \left(\sum_{i=1}^n |a_{ij}| \right) |x_j|.$$

Пусть $Q = \max_j \sum_{i=1}^n |a_{ij}|$. Тогда $\|y\|_2 \leq Q \sum_{j=1}^n |x_j|$.

В случае третьей векторной нормы

$$\|A\|_3 = \sqrt{\max_i |\lambda_i(A^*A)|}.$$

Здесь $\lambda_i(A^*A)$ — собственные значения матрицы A^*A .

В самом деле,

$$\|Ax\|_3^2 = (Ax, Ax) = (A^*Ax, x).$$

Собственные вектора матрицы A^*A ортогональны. Пусть они еще и ортонормированны: $(e_i, e_j) = \delta_{i,j}$. Здесь δ_{ij} — символы Кронекера. Разложим вектор x по собственным векторам e_k матрицы A^*A .

$$x = c_1 e_1 + c_2 e_2 + \dots + c_n e_n.$$

Тогда

$$\|x\|_3^2 = c_1^2 + c_2^2 + \dots + c_n^2.$$

Но $A^*Ax = \lambda_1 c_1 e_1 + \lambda_2 c_2 e_2 + \dots + \lambda_n c_n e_n$.

Поэтому

$$\|Ax\|_3^2 = (Ax, Ax) = (A^*Ax, x) = \left(\sum_{i=1}^n \lambda_i c_i e_i, \sum_{i=1}^n c_i e_i \right) = \sum_{i=1}^n \lambda_i c_i^2 \leq \max_i |\lambda_i| \|x\|^2.$$

§ 1. Точные методы решения линейных систем.

Понятие обусловленности

Метод Гаусса. Рассмотрим линейную систему из n уравнений относительно n неизвестных

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \dots + a_{3n}x_n &= b_3 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nn}x_n &= b_n \end{aligned} \tag{1.1}$$

Пусть a_{11} отлично от нуля. В противном случае перенумеруем переменные так, чтобы a_{11} не равнялось нулю. Умножим 1-е уравнение на $\frac{a_{21}}{a_{11}}$ и вычтем из 2-го уравнения. Продолжим этот процесс, умножая 1-е уравнение на $\frac{a_{31}}{a_{11}}$ и вычитая из 3-го уравнения и т.д.

В результате получим систему, в которой, кроме 1-го уравнения, отсутствует переменная x_1

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n &= b_1 \\ a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n &= b_2^1 \\ a_{32}x_2 + a_{33}x_3 + \dots + a_{3n}x_n &= b_3^1 \\ &\vdots \\ a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nn}x_n &= b_n^1. \end{aligned}$$

Продолжим этот процесс, умножая 2-е уравнение на $\frac{a_{12}^1}{a_{22}^1}$ и вычитая из i -го уравнения. В случае, если $a_{22}^1 = 0$, перенумеруем переменные x_i начиная с x_2 . Это дает возможность понизить порядок системы еще на единицу и т.д. В результате приходим к треугольной системе:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1$$

$$a_{22}^1 x_2 + a_{23}^1 x_3 + \dots + a_{2n}^1 x_n = b_2^1$$

$$a_{33}^2 x_3 + \dots + a_{3n}^2 x_n = b_3^2$$

• • • • •

$$a_{n-1n-1}^{n-2}x_{n-1} + a_{nn}^{n-2}x_n = b_n^{n-2}$$

$$a_{nn}^{n-1} x_n = b_n^{n-1}.$$

Определяя из последнего уравнения x_n , затем из предпоследнего x_{n-1} и т.д., обратным ходом определим все x_i . Наш процесс может остановиться, если на k -м шаге не удастся найти ненулевой элемент в k строке даже в результате перенумерации переменных от x_k до x_n . Это будет означать, что все коэффициенты этой строки равны нулю. Если b_k не равно нулю, то система не имеет решения, если же и $b_k = 0$, то решение не единственно. Этот алгоритм требует порядка n^2 операций и в силу накопления вычислительной погрешности неприменим к системам большой размерности. Существуют модельные примеры систем размерности >20 с заранее известным решением, такие, что применение этого алгоритма приводило к результату, далекому от решения. Особенно этот эффект проявляется в случае плохо обусловленных систем.

Понятие обусловленности систем. Мера обусловленности

Рассмотрим простой пример

$$4x + 2y = 6$$

$$11.7x + 6.1y = 17.8.$$

Решение этой системы $x=y=1$. Изменим правую часть 2-го уравнения

$$4x + 2y = 6$$

$$11.7x + 6.1y = 18$$

Решение системы $x=0.6$, $y=1.8$. Замечаем, что небольшое изменение правой части системы (примерно на процент) приводит к существенному изменению решения. Продолжим изучение примера и при той же правой части немного изменим в исходном примере коэффициенты системы.

$$4x + 2y = 6$$

$$12x + 6.1y = 17.8$$

Здесь решение $x=2.5$, $y=-2$. Обычно считается, что небольшое изменение коэффициентов приводит к небольшому изменению решения. Здесь это не так. Ситуация крайне неприятная, поскольку в практических расчетах неизвестно, с какой точностью определены коэффициенты системы. Изучим этот пример подробнее и найдем обратную матрицу для исходной и измененной матрицы. Для матриц A и B

$$A = \begin{pmatrix} 4 & 2 \\ 11.7 & 6.1 \end{pmatrix}$$

$$B = \begin{pmatrix} 4 & 2 \\ 12 & 6.1 \end{pmatrix}$$

вычислим обратные матрицы

$$A^{-1} = \begin{pmatrix} 6.1 & -2 \\ -11.7 & 4 \end{pmatrix}$$

$$B^{-1} = \begin{pmatrix} 15.25 & -5 \\ -29.25 & 10 \end{pmatrix}.$$

Таким образом, замечаем, что небольшое изменение коэффициентов исходной матрицы приводит к непропорционально сильному изменению обратной матрицы. Если вычислить нормы, то получим $\|A^{-1}\|_1 = 15.7$, а $\|B^{-1}\|_1 = 39.25$. Исторически матрицы, у которых небольшое изменение коэффициентов приводит к сильному изменению обратной матрицы, называют плохо обусловленными матрицами, а системы с такими матрицами — плохо обусловленными системами. Для числовой характеристики этого эффекта вводится понятие меры обусловленности.

Пусть заданы системы

$$Ax = b \quad \text{и} \quad Ay = b + \eta.$$

Определение. Пусть $r=y-x$. Мерой обусловленности системы назовем величину

$$\tau = \frac{\|b\|}{\|x\|} \sup_{\eta} \frac{\|r\|}{\|\eta\|}.$$

Поскольку $Ar = \eta$, т. е. $r = A^{-1}\eta$, то, согласно определению нормы оператора,

$$\sup_{\eta} \frac{\|r\|}{\|\eta\|} = \sup_{\eta} \frac{\|A^{-1}\eta\|}{\|\eta\|} = \|A^{-1}\|.$$

Так как $b=Ax$, то

$$\frac{\|b\|}{\|x\|} = \frac{\|Ax\|}{\|x\|} \leq \|A\|.$$

Поэтому

$$\tau \leq \|A\| \|A^{-1}\|.$$

Величину

$$\gamma = \|A\| \|A^{-1}\|$$

называют мерой обусловленности матрицы. Эта величина характеризует свойства системы только через свойства матрицы. В практике для характеристики меры обусловленности матрицы используют величину

$$\bar{\gamma} = \frac{\max |\lambda(A)|}{\min |\lambda(A)|}.$$

Легко видеть, что $\bar{\gamma} \leq \gamma$, а в случае симметричных матриц и использования 3-й векторной нормы $\bar{\gamma} = \gamma$. В самом деле, если матрица A симметричная, то ее 3-я норма совпадает с ее максимальным по модулю собственным значением, а $\max |\lambda(A)| = |\min \lambda(A^{-1})|$. В остальных случаях неравенство $\bar{\gamma} \leq \gamma$ следует из того, что любая норма больше (либо равна) максимального по модулю собственного значения. Последнее утверждение с очевидностью следует из неравенства $\|A\| \|x\| \geq \|Ax\|$, если вектор x есть собственный вектор, соответствующий максимальному по модулю собственному значению. В разобранный выше примере $\lambda_1 = 10$, $\lambda_2 = 0.1$ и $\bar{\gamma} = 100$.

§ 2. Частичная проблема собственных значений Итерационные методы

Под частичной проблемой собственных значений понимается задача отыскания максимального по модулю собственного значения. В силу ее важности в приложениях, исследования в этой области имеют большую историю как отечественных, так и зарубежных математиков. Здесь следует упомянуть в первую очередь А. Н. Крылова, Д. К. Фадеева, Ф. М. Данилевского.

Напомним, что собственными значениями матрицы A называются корни ее характеристического многочлена, т. е. корни уравнения

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

Ниже будет рассмотрен алгоритм приближенного нахождения наибольшего по модулю собственного значения матрицы A . Далее мы будем предполагать, что собственные значения матрицы A пронумерованы в порядке убывания их модулей:

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

Рассмотрим итерационный алгоритм

$$X^{(k+1)} = AX^{(k)}, \quad (2.1)$$

выбрав в качестве вектора $X^{(0)}$ любой ненулевой вектор. Пусть разложение вектора $X^{(0)}$ по собственным векторам e_k матрицы A имеет вид

$$X^{(0)} = c_1 e_1 + c_2 e_2 + \dots + c_n e_n. \quad (2.2)$$

Тогда

$$X^{(k)} = \lambda_1^k c_1 e_1 + \lambda_2^k c_2 e_2 + \dots + \lambda_n^k c_n e_n. \quad (2.3)$$

Для любой фиксированной компоненты $x_*^{(0)}$ вектора $X^{(0)}$ имеем разложение

$$x_*^{(0)} = c_1 e_{1*} + c_2 e_{2*} + \dots + c_n e_{n*}.$$

Здесь e_{i*} — соответствующая фиксированная компонента собственного вектора e_i . Пусть $d_i = c_i e_{i*}$, т. е.

$$x_*^{(0)} = d_1 + d_2 + \dots + d_n. \quad (2.2')$$

Для этой же компоненты $x_*^{(k)}$ вектора $X^{(k)}$

$$x_*^{(k)} = \lambda_1^k c_1 e_{1*} + \lambda_2^k c_2 e_{2*} + \dots + \lambda_n^k c_n e_{n*}.$$

Или

$$x_*^{(k)} = \lambda_1^k d_1 + \lambda_2^k d_2 + \dots + \lambda_n^k d_n. \quad (2.3')$$

Предположим

1. Максимальное по модулю собственное значение вещественно и единственно.

Пусть

$$\lambda_1^{(k)} = \frac{(X^{(k+1)}, X^{(k)})}{(X^{(k)}, X^{(k)})}. \quad (2.4)$$

Покажем, что $\lambda_1^{(k)} \rightarrow \lambda_1$ со скоростью порядка $O((\frac{\lambda_2}{\lambda_1})^k)$.

В самом деле,

$$X^{(k+1)} = \lambda_1^{k+1} c_1 e_1 + \sum_{i=2}^n \lambda_i^{k+1} c_i e_i = \lambda_1^{k+1} [c_1 e_1 + O((\frac{\lambda_2}{\lambda_1})^{k+1})]$$

$$X^{(k)} = \lambda_1^k c_1 e_1 + \sum_{i=2}^n \lambda_i^k c_i e_i = \lambda_1^k [c_1 e_1 + O((\frac{\lambda_2}{\lambda_1})^k)].$$

Тогда

$$\frac{(X^{(k+1)}, X^{(k)})}{(X^{(k)}, X^{(k)})} = \frac{\lambda_1^{2k+1} [c_1 e_1 + O((\frac{\lambda_2}{\lambda_1})^{k+1})] [c_1 e_1 + O((\frac{\lambda_2}{\lambda_1})^k)]}{\lambda_1^{2k} [c_1 e_1 + O((\frac{\lambda_2}{\lambda_1})^k)]^2}.$$

Из последнего выражения видно, что $\lambda_1^{(k)}$ стремится к λ_1 .

Максимальное по модулю собственное значение можно вычислить и по более простой формуле

$$\lambda_1 \approx \frac{x_*^{(k+1)}}{x_*^{(k)}}. \quad (2.4')$$

Напомним, что $x_*^{(k)}$ — любая фиксированная компонента вектора $X^{(k)}$. В самом деле, из (2.3')

$$x_*^{(k+1)} = \lambda_1^{k+1} (d_1 + O((\frac{\lambda_2}{\lambda_1})^{k+1}))$$

и

$$\bar{\lambda}_1^n = \frac{\lambda_1^{k+1} (d_1 + O((\frac{\lambda_2}{\lambda_1})^{k+1}))}{\lambda_1^k (d_1 + O((\frac{\lambda_2}{\lambda_1})^k))} \rightarrow \lambda_1.$$

При практических вычислениях отмечено, что итерационный процесс не даст возможности определить λ_1 , если в качестве начального вектора $X^{(0)}$ выбран вектор, ортогональный собственному вектору, соответствующему λ_1 . Кроме того, процесс может остановиться, если λ_1 и λ_2 близки в пределах вычислительной погрешности.

2. Два собственных значения равны по модулю и различны по знаку.

Тогда с помощью итерационного процесса (2.1) λ_1^2 приближенно можно вычислить с помощью выражения

$$\lambda_1^2 \approx \frac{(X^{(2k+2)}, X^{(2k)})}{(X^{(2k)}, X^{(2k)})} \quad (2.5)$$

или выражения

$$\lambda_1^2 \approx \frac{x_*^{(2k+2)}}{x_*^{(2k)}}. \quad (2.5')$$

Напомним, что x_*^k — любая фиксированная компонента вектора $X^{(k)}$.

Докажем формулу (2.5').

С учетом разложения (2.3') при $\lambda_2 = -\lambda_1$

$$\frac{x_*^{(2k+2)}}{x_*^{(2k)}} = \frac{\lambda_1^{4k+2}[d_1 + d_2 + O((\frac{\lambda_3}{\lambda_1})^{2k+2})][d_1 + d_2 + O((\frac{\lambda_3}{\lambda_1})^{2k})]}{\lambda_1^{4k}[d_1 + d_2 + O((\frac{\lambda_3}{\lambda_1})^{2k})]^2}.$$

Отсюда замечаем, что при увеличении k это выражение стремится к λ_1^2 . Формула (2.5) доказывается аналогично.

3. Наибольшие по модулю собственные значения образуют простую комплексную пару. В этом случае в разложении (2.3') два первых слагаемых $\lambda_1^k d_1 + \lambda_2^k d_2$ комплексно сопряжены. Если $\lambda_1 = re^{i\alpha}$, $d_1 = Re^{i\beta}$, то $\bar{\lambda}_1 = re^{-i\alpha}$ и $\bar{d}_1 = Re^{-i\beta}$. Поэтому

$$x_*^{(k)} = 2Rr \cos(k\alpha + \beta) + \lambda_3^k c_3 e_{3*} + \lambda_4^k c_4 e_{4*} + \dots$$

Последнее выражение показывает, что наличие наибольших по модулю комплексных корней приводит к тому, что величина $x_*^{(k)}$ при итерациях будет сильно колебаться по величине, меняя знак.

Пусть λ_1 и λ_1 — корни квадратного уравнения

$$\lambda^2 + p\lambda + q = 0.$$

Тогда

$$\lambda_1 + \lambda_2 = -p \quad \text{и} \quad \lambda_1 \lambda_2 = q.$$

Коэффициенты p и q можно приближенно определить из следующих соображений. Пусть k настолько велико, что $x_*^k \approx d_1 \lambda_1^k + d_2 \lambda_2^k$. Тогда

$$x_*^{(k+1)} + p x_*^k + q x_*^{(k-1)} \approx d_1 \lambda_1^{(k-1)} [\lambda_1^2 + p \lambda_1 + q] + d_2 \lambda_2^{(k-1)} [\lambda_2^2 + p \lambda_2 + q] = 0.$$

Следовательно,

$$x_*^{(k+1)} + p x_*^k + q x_*^{(k-1)} \approx 0. \quad (2.6)$$

Такое же приближенное равенство будет справедливо для любой другой компоненты z_*^k вектора $X^{(k)}$.

$$z_*^{(k+1)} + p z_*^k + q z_*^{(k-1)} \approx 0. \quad (2.6')$$

Решая (2.6) и (2.6'), получим

$$p \approx - \frac{x^{(k-1)} z^{(k+1)} - z^{(k-1)} x^{(k+1)}}{x^{(k-1)} z^{(k)} - z^{(k-1)} x^{(k)}} \quad (2.7)$$

и

$$q \approx \frac{x^{(k)} z^{(k+1)} - z^{(k)} x^{(k+1)}}{x^{(k-1)} z^{(k)} - z^{(k-1)} x^{(k)}}. \quad (2.8)$$

4. Наибольшее по модулю собственное значение вещественно и находится в жордановой клетке 2-го порядка.

В этом случае в качестве базиса возьмем базис, составленный из собственных векторов и присоединенного вектора, который соответствует собственному значению λ_1 . Пусть e_1 — присоединенный вектор, а e_2 — собственный вектор, соответствующий собственному значению λ_1 . Тогда

$$Ae_1 = \lambda e_1 + e_2, \quad Ae_2 = \lambda e_2, \quad \dots, \quad Ae_n = \lambda e_n.$$

Поэтому

$$\begin{aligned} A^k e_1 &= \lambda_1^k e_1 + k \lambda_1^{(k-1)} e_2 \\ A^k e_2 &= \lambda_2 e_2 \\ &\dots \dots \dots \\ A^k e_n &= \lambda_n e_n \end{aligned}$$

Пусть $X^{(0)}$ — начальный вектор. Будем предполагать, что проекция этого вектора на корневое подпространство, которое соответствует собственному значению λ_1 , не равна нулю и не является собственным вектором. Пусть это вектор h . Примем его за первый вектор базиса. Тогда

$$X^{(k)} = A^k X^{(0)} = \lambda_1^k h + k \lambda_1^{(k-1)} e_1 + \lambda_3^k c_3 e_3 + \dots + \lambda_n^{(n)} c_n e_n.$$

Для любой фиксированной компоненты вектора $x^{(k)}$ получим

$$x_*^{(k)} = d_1 \lambda_1^k + d_2 k \lambda_1^{k-1} + \dots + d_n \lambda_n^k.$$

Из последнего разложения видно, что отношение $\frac{x_*^{k+1}}{x_*^k}$ по-прежнему стремится к λ_1 , но медленнее, чем любая геометрическая прогрессия из-за наличия множителя k во втором слагаемом. Именно

$$\frac{x_*^{k+1}}{x_*^k} = \lambda_1 (1 + O(\frac{1}{k})).$$

Обычно в этом случае сильно влияет вычислительная погрешность, и в такой ситуации следует поступить так же, как и в случае комплексной пары собственных значений, а именно вычислить коэффициенты p и q квадратного уравнения по формулам (2.6) и (2.7). Если $-p \approx 4q^2$, то это свидетельствует о правильности гипотезы о кратном собственном значении с жордановой клеткой размерности 2.

В заключение заметим, что в случае 1 вектор $X^{(k)}$ стремится к собственному вектору, соответствующему λ_1 , который следует нормировать на единицу. В случае разных по знаку, максимальных по модулю собственных значений вектор $X^{(k)}$ стремится к вектору, который принадлежит линейной комбинации их собственных векторов. В последнем разобранном случае вектор $X^{(k)}$ стремится к вектору, принадлежащему корневому подпространству собственного значения λ_1 . В случае комплексной пары максимальных по модулю собственных значений можно показать, что вектора

$$X^{(k+1)} - \lambda_2 X^{(k)}$$

и

$$X^{(k+1)} - \lambda_1 X^{(k)}$$

при увеличении k стремятся к собственным векторам, соответствующим собственным значениям λ_1 и λ_2 .

Разумеется, описанные алгоритмы не исчерпывают все возможные случаи в задаче отыскания максимального по модулю собственного значения, а лишь позволяют приближенно найти максимальное по модулю собственное значение в наиболее часто встречающихся ситуациях. Простой пример: $\lambda_1 = 1$, $\lambda_{2,3} = \pm i$ показывает, что ни один из рассмотренных итерационных процессов не позволяет найти наибольшее по модулю собственное значение.

§ 3. Итерационные методы решения линейных систем

Итерационные методы решения линейных систем представляют собой некоторые алгоритмы построения вектора $X^{(k)}$, который является приближенным решением системы n линейных уравнений относительно n неизвестных.

$$AX = b. \quad (3.1)$$

В отличие от точных методов, итерационные методы не являются универсальными. Матрица A системы должна удовлетворять некоторым условиям, чтобы последовательность векторов $X^{(k)}$ сходилась к точному решению системы. Однако для систем большой размерности применение итерационных методов предпочтительнее, так как они более устойчивы к вычислительной погрешности.

Метод простой итерации.

Систему (3.1) преобразуем к виду

$$X = BX + g. \quad (3.2)$$

Это можно сделать, переписав систему (3.1) в виде $-AX + b = 0$, и затем добавить и отнять вектор X . Иногда систему (3.1) умножают на некоторую неособенную матрицу H , а затем добавляют и отнимают вектор X .

Зададим произвольный ненулевой начальный вектор $X^{(0)}$ и построим итерационный процесс по формуле:

$$X^{(k+1)} = BX^{(k)} + g. \quad (3.3)$$

Теорема 1. Пусть $\|B\| \leq q < 1$. Тогда существует единственное решение системы (3.2) и итерационный процесс сходится к решению системы с любого начального приближения.

Доказательство. Докажем единственность. Предположим противное: существуют два решения X_1 и X_2 . Тогда их разность $R = X_1 - X_2$ есть решение однородной системы

$$R = BR. \quad (3.4)$$

Отсюда $\|R\| = \|BR\| \leq \|B\|\|R\|$. Следовательно, $\|R\|(1 - \|B\|) \leq 0$, т. е. $\|R\| \leq 0$. Но $\|R\| \geq 0$. Поэтому $\|R\| = 0$ и по свойству нормы $R = 0$.

Достаточность. Пусть $R^{(k+1)} = X^{(k+1)} - X^{(k)}$. Вектор $R^{(k)}$ удовлетворяет однородной системе $R^{(k+1)} = BR^{(k)}$. Поэтому

$$R^{(k+1)} = B^k R^{(1)}$$

и

$$\|R^{(k+1)}\| \leq q^k \|R^{(1)}\|.$$

Пусть $\|R^{(1)}\| = a$. Покажем, что последовательность $X^{(k)}$ фундаментальна. В силу неравенства треугольника

$$\|X^{(k+p)} - X^{(k)}\| \leq \|X^{(k+p)} - X^{(k+p-1)}\| + \|X^{(k+p-1)} - X^{(k+p-2)}\| + \dots + \|X^{(k+1)} - X^{(k)}\|.$$

Поэтому $\|X^{(k+p)} - X^{(k)}\| \leq$

$$\leq (q^{k+p-1} + q^{k+p-2} + \dots + q^k)a = q^k a(1 + q + \dots + q^{p-1}) \leq q^k a \frac{1}{1-q}.$$

Поскольку $q < 1$, последнее выражение может быть меньше любого числа при $k > k_0$. Последовательность $X^{(k)}$ фундаментальна, и, следовательно, существует предел \bar{X} . Пусть теперь $R^{(k)} = X^{(k)} - \bar{X}$. Последовательность $R^{(k)}$ удовлетворяет однородному уравнению $R^{(k+1)} = BR^{(k)}$. Поэтому $\|R^{(k)}\| \leq q^k \|R^{(0)}\|$. Это означает, что последовательность $X^{(k)}$ сходится к \bar{X} со скоростью q^k . Теорема доказана.

Достаточные признаки сходимости метода простой итерации.

Лемма. Пусть собственные значения матрицы B удовлетворяют условию $|\lambda_i| \leq q$, причем $\lambda_i = q$ соответствует жорданова клетка размерности 1. Тогда существует матрица H и неособенная матрица D такая, что $H = D^{-1}BD$, причем $\|H\|_1 \leq q$.

Доказательство. Пусть $p = q - \max_{|\lambda_i| < q} |\lambda_i|$. Рассмотрим матрицу $B_1 = p^{-1}B$. Собственными значениями матрицы B_1 будут $p^{-1}\lambda_i$. Преобразуем матрицу B_1 к нормальной жордановой форме

$$H_1 = D^{-1}B_1D = D^{-1}p^{-1}BD.$$

Но

$$H_1 = \begin{pmatrix} p^{-1}\lambda_1 & \alpha_{12} & 0 & \dots \\ 0 & p^{-1}\lambda_2 & \alpha_{23} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}.$$

Здесь $\alpha_{i,i+1} = 0$, если λ_i - простое собственное значение, или 1, если λ_i соответствует клетка Жордана размерности > 1 . Пусть $H = pH_1$. Тогда

$$H = pD^{-1}p^{-1}BD = \begin{pmatrix} \lambda_1 & p\alpha_{12} & 0 & \dots \\ 0 & \lambda_2 & p\alpha_{23} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}.$$

Вычислим первую норму матрицы H .

Если $|\lambda_i| = q$, то $\alpha_{i,i+1} = 0$ по условию леммы. Если же $|\lambda_i| < q$, то возможно $\alpha_{i,i+1} = 1$. В этом случае $|\lambda_i| + p = |\lambda_i| + q - \max_{|\lambda_i| < q} |\lambda_i| \leq q$.

Поэтому $\|H\|_1 \leq q$. Лемма доказана.

Теорема 2. Пусть собственные значения матрицы B по абсолютной величине меньше единицы. Тогда система (3.2) имеет единственное решение и итерационный процесс (3.3) сходится к решению системы с любого начального приближения.

Доказательство. Докажем сначала единственность. Пусть $X=DY$, где D — некоторая невырожденная матрица. Тогда систему (3.2) можно записать в виде

$$Y = D^{-1}BDY + D^{-1}g. \quad (3.5)$$

Пусть матрица D такова, что матрица $H = D^{-1}BD$ есть нормальная жорданова форма матрицы B . Если $|\lambda_i| \leq q < 1$ $\|H\|_1 < 1$, то, согласно теореме 1, система (3.5) имеет единственное решение, и это можно утверждать и для исходной системы в силу невырожденности матрицы D .

Достаточность. Пусть H — нормальная жорданова форма матрицы B : $B = DHD^{-1}$. Но

$$B^k = DHD^{-1} \dots DHD^{-1} = DH^kD^{-1}.$$

Если \bar{X} — решение системы, то для $R^k = X^{(k)} - \bar{X}$ имеем

$$\|R^{(k+1)}\|_1 \leq \|D\|_1 \|H^k\|_1 \|D^{-1}\|_1 \leq a q^k,$$

где $a = \|D\|_1 \|D^{-1}\|_1$. Таким образом, доказано, что $X^{(k+1)} \rightarrow \bar{X}$ со скоростью q^k .

Сделаем еще несколько замечаний по поводу теоремы 2. Пусть $|\lambda_i| < 1$ $i=2,3,\dots,n$, а $|\lambda_1| > 1$. Тогда, если выбрать начальное приближение так, что $R^{(0)} = e_1$, где e_1 — собственный вектор, соответствующий λ_1 , то очевидно, что итерационный процесс будет расходиться. С другой стороны, если решение системы принадлежит линейной комбинации векторов e_2, e_3, \dots, e_n , а вектор $X^{(0)}$ таков, что $R^{(0)} = c_2 e_2 + c_3 e_3 + \dots c_n e_n$, то итерационный процесс будет сходиться. Поэтому в некоторых учебниках теорема 2' формулируется следующим образом

Теорема 2'. Для того чтобы итерационный процесс (3.3) сходил к любому начальному приближению, необходимо и достаточно, чтобы все собственные значения матрицы B были по модулю меньше 1. Как мы заметили, фраза "с любого начального приближения" является существенной.

$$\begin{aligned} a_{11}x_1^{(k+1)} + a_{12}x_2^{(k)} + a_{13}x_3^{(k)} + \dots + a_{1n}x_n^{(k)} &= b_1 \\ a_{21}x_1^{(k+1)} + a_{22}x_2^{(k+1)} + a_{23}x_3^{(k)} + \dots + a_{2n}x_n^{(k)} &= b_2 \end{aligned} \quad (3.8)$$

Этот итерационный процесс называется методом Зейделя.

Этот же процесс, записанный в виде

$$x_p^{(k+1)} = \sum_{j=1}^{p-1} \frac{a_{pj}}{a_{pp}} x_j^{(k+1)} + \sum_{j=p+1}^n \frac{a_{pj}}{a_{pp}} x_j^{(k)} + \frac{b_p}{a_{pp}},$$

в отечественной литературе иногда называют методом Некрасова.

Введем в рассмотрение матрицы

$$B = \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & & 0 \\ \dots & \dots & \dots & & \dots \\ a_{n1} & a_{n2} & a_{n3} & & a_{nn} \end{pmatrix}$$

С помощью матриц В и С итерационный процесс (3.8) можно записать в виде:

$$X^{(k+1)} = -B^{-1}CX^{(k)} + B^{-1}b.$$

Поэтому метод Зейделя есть метод простой итерации с матрицей, равной $-B^{-1}C$, и правой частью $B^{-1}b$. Как показано выше, для сходимости этого метода с любого начального приближения необходимо и достаточно, чтобы собственные значения матрицы $-B^{-1}C$ были по модулю меньше единицы. Но

$$\det(-B^{-1}C - \lambda E) = \det(-B^{-1})\det(C + \lambda B),$$

и мы приходим к теореме

сужения необходимо и достаточно, чтобы корни уравнения

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

были по модулю меньше 1.

Докажем еще несколько достаточных признаков.

Теорема 5. Пусть матрица A — матрица с диагональным преобладанием, т. е.

$$\sum_{j=1, j \neq i}^n \frac{|a_{ij}|}{|a_{ii}|} \leq q < 1 \quad i = 1, \dots, n \quad (3.9)$$

Тогда метод Зейделя сходится.

Доказательство. Пусть \bar{X} — точное решение $B\bar{X} + C\bar{X} = b$, а $X^{(k)}$ — приближение по методу Зейделя: $BX^{(k+1)} + CX^{(k)} = b$. Для $R^{(k)} = X^{(k)} - \bar{X}$ имеем:

$$BR^{(k+1)} + CR^{(k)} = 0. \quad (3.10)$$

Покажем, что $\|R^{(k+1)}\|_1 \rightarrow 0$.

Пусть первая норма $R^{(k+1)} = (r_1^{(k+1)}, r_2^{(k+1)}, \dots, r_n^{(k+1)})^T$ достигается на компоненте $r_p^{(k+1)}$, т. е. $\|R^{(k+1)}\|_1 = |r_p^{(k+1)}|$.

Запишем p -е уравнение системы (3.10) в виде

$$r_p^{(k+1)} = - \sum_{i=1}^{p-1} \frac{a_{ip}}{a_{pp}} r_i^{(k+1)} - \sum_{i=p+1}^n \frac{a_{ip}}{a_{pp}} r_i^{(k)}.$$

Отсюда

$$|r_p^{(k+1)}| \leq \alpha \|R^{(k+1)}\|_1 + \beta \|R^{(k)}\|_1,$$

где

$$\alpha = \sum_{i=1}^{p-1} \left| \frac{a_{ip}}{a_{pp}} \right|, \quad \beta = \sum_{i=p+1}^n \left| \frac{a_{ip}}{a_{pp}} \right|, \quad \alpha + \beta \leq q < 1.$$

Следовательно, так как $|r_p^{(k+1)}| = \|R^{(k+1)}\|_1$ по предположению, получаем

$$\|R^{(k+1)}\|_1 \leq \alpha \|R^{(k+1)}\|_1 + \beta \|R^{(k)}\|_1.$$

Отсюда

$$\|R^{(k+1)}\|_1 \leq \frac{\beta}{1-\alpha} \|R^{(k)}\|_1.$$

Но $\beta \leq q - \alpha$. Поэтому

$$\frac{\beta}{1-\alpha} = q_1 < 1.$$

Таким образом,

$$\|R^{(k+1)}\|_1 \leq q_1 \|R^{(k)}\|_1 \leq \dots \leq q_1^k \|R^{(0)}\|_1.$$

Теорема доказана.

Следующий достаточный признак мы докажем при изучении метода скорейшего спуска.

Теорема 6. Пусть матрица A симметричная и положительно определенная. Тогда метод Зейделя сходится к решению системы (3.1).

Напомним, что матрица A называется положительно определенной, если для любого ненулевого вектора имеет место неравенство

$$(AX, X) \geq \gamma^2 (X, X).$$

Метод скорейшего спуска. Для вещественной, симметричной и положительно определенной матрицы A рассмотрим функционал

$$F(X) = (AX, X) - 2(b, X). \quad (3.11)$$

Теорема 7. Пусть \bar{X} — решение системы

$$AX = b. \quad (3.12)$$

Тогда вектор \bar{X} сообщает функционалу $F(X)$ минимум. Обратно, вектор \bar{X} , на котором функционал достигает своего минимума, есть решение системы (3.12).

Доказательство. Пусть X — произвольный вещественный вектор, \bar{X} — решение системы (3.12) и $X - \bar{X} = R$. Поскольку $b = A\bar{X}$ и в силу симметричности матрицы A

$$\begin{aligned} F(X) &= (AX, X) - 2(b, X) = (A(\bar{X} + R), (\bar{X} + R)) - 2(b, \bar{X} + R) = \\ &= (A\bar{X}, \bar{X}) + (A\bar{X}, R) + (AR, \bar{X}) + (AR, R) - 2[(A\bar{X}, \bar{X}) + (A\bar{X}, R)] = \\ &= (A\bar{X}, \bar{X}) - 2(b, \bar{X}) + (AR, R) = F(\bar{X}) + (AR, R). \end{aligned}$$

В силу того, что матрица A положительно определенная, $F(X) \geq F(\bar{X})$, что и доказывает, что \bar{X} сообщает функционалу $F(X)$ минимум. Докажем обратное утверждение, что элемент \bar{X} , на котором $F(X)$ достигает своего минимума, есть решение системы (3.12). Функционал $F(X)$ есть функция координат вектора X - переменных x_1, x_1, \dots, x_1 и в точке минимума $\text{grad } F(X)=0$, т. е. $\frac{\partial F}{\partial x_i} = 0 \quad i = 1, 2, \dots, n$. Но

$$F(X) = \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} x_j \right) x_i - 2 \sum_{i=1}^n b_i x_i$$

и

$$\frac{\partial F}{\partial x_p} = \sum_{j=1}^n a_{pj} x_j + \sum_{i=1}^n a_{ip} x_j - 2b_p.$$

Матрица A симметричная: $a_{ij} = a_{ji}$, поэтому

$$\frac{\partial F}{\partial x_p} = 2 \left(\sum_{j=1}^n a_{pj} x_j - b_p \right).$$

Следовательно,

$$\text{grad} F(X) = 2(AX - b) \quad (3.13)$$

и равенство нулю градиента $F(X)$ при некотором $X = \bar{X}$ означает, что компоненты этого вектора есть решение системы (3.12). Теорема доказана.

Таким образом, задача отыскания решения системы (3.12) сведена к задаче отыскания минимума функционала $F(X)$. Рассматриваемый ниже метод скорейшего спуска есть один из алгоритмов отыскания минимума функционала $F(X)$.

Построим следующий итерационный процесс

$$X^{(k+1)} = X^{(k)} - \delta_k \text{grad} F(X^{(k)}). \quad (3.14)$$

Согласно этому алгоритму, для минимизации $F(X)$ от вектора $X^{(k)}$ мы движемся в направлении $-\text{grad} F(X^{(k)})$ — направлении наибольшего убывания функционала — на некоторое расстояние, которое будет определено из условия минимума функции

$$\Psi(\delta) = F(X^{(k)} - \delta_k \text{grad} F(X^{(k)})).$$

Согласно (3.13) $\text{grad} F(X) = 2(AX - b)$. Пусть $R^{(k)} = b - AX^{(k)}$, а $2\delta_k = \lambda_k$. Процесс (3.14) можно записать в виде

$$X^{(k+1)} = X^{(k)} + \lambda_k R^{(k)}.$$

Определим λ_k из условия минимума функции

$$\Psi(\lambda) = F(X^{(k)} + \lambda R^{(k)}),$$

$$\Psi(\lambda) = (A(X^{(k)} + \lambda R^{(k)}), (X^{(k)} + \lambda R^{(k)})) - 2(b, X^{(k)} + \lambda R^{(k)}).$$

Вычислим $\Psi'(\lambda)$.

$$\begin{aligned} \Psi'(\lambda) &= (AR^{(k)}, X^{(k)} + \lambda R^{(k)}) + (A(X^{(k)} + \lambda R^{(k)}), R^{(k)}) - 2(b, R^{(k)}) = \\ &= 2(AX^{(k)}, R^{(k)}) + \lambda(AR^{(k)}, R^{(k)}) - 2(b, R^{(k)}) = \\ &= 2((AX^{(k)} - b, R^{(k)}) + \lambda(AR^{(k)}, R^{(k)})). \end{aligned}$$

Но $AX^{(k)} - b = -R^{(k)}$, и из условия $\Psi'(\lambda) = 0$ получаем

$$\lambda_k = \frac{(R^{(k)}, R^{(k)})}{(AR^{(k)}, R^{(k)})}.$$

Соберем воедино формулы алгоритма метода скорейшего спуска для линейных систем:

$$X^{(k+1)} = X^{(k)} + \lambda_k R^{(k)}, \quad (3.15)$$

где

$$\lambda_k = \frac{(R^{(k)}, R^{(k)})}{(AR^{(k)}, R^{(k)})}. \quad (3.16)$$

В заключение заметим, что метод Зейделя есть метод покоординатной минимизации функционала $F(X)$. В самом деле, пусть заданы начальные приближения $x_2^0, x_3^0, \dots, x_n^0$ вектора X . Приближение к компоненте x_1 - величину x_1^1 выберем так, чтобы $F(X) = F(x_1, x_2^0, \dots, x_n^0)$ было минимально. В этом случае

$$\frac{\partial F(X)}{\partial x_1} = 0.$$

При доказательстве теоремы 7 мы заметили, что

$$\frac{\partial F(X)}{\partial x_j}$$

есть удвоенное j -е уравнение системы (3.12). Следовательно, x_1^1 определяется из уравнения

$$a_{11}x_1^1 + a_{12}x_2^0 + \dots + a_{1n}x_n^0 = b_1.$$

Пусть теперь заданы начальные приближения $x_1^1, x_2^0, \dots, x_n^0$. Минимизация функционала по переменной x_2 приводит к тому, что x_2^1 определяется из уравнения

$$a_{21}x_1^1 + a_{22}x_2^1 + \dots + a_{2n}x_n^0 = b_2.$$

Продолжая этот процесс и считая, что компоненты $x_1^1, x_2^1, \dots, x_{n-1}^1$ найдены, определим x_n^1 из условия минимума $F(X)$ по переменной x_n , что приводит к уравнению

$$a_{n1}x_1^1 + a_{n2}x_2^1 + \dots + a_{nn}x_n^1 = b_n.$$

Таким образом, замечаем, что метод Зейделя есть метод покоординатной минимизации функционала $F(X)$, что и доказывает признак сходимости, сформулированный в теореме 6.

§ 4. Итерационные методы решения нелинейных уравнений и систем

В начале этой темы необходимо отметить, что ситуация с методами решения нелинейных систем существенно отличается от той, которая имеет место при изучении итерационных методов для линейных систем. В предыдущем параграфе мы доказывали различные теоремы о сходимости тех или иных методов с любого начального приближения. При исследовании итерационных процессов для нелинейных уравнений и систем выбор начального приближения играет существенную роль, поскольку неудачный выбор начального приближения к решению может привести к неверному результату или расходимости исследуемого итерационного процесса. При практических вычислениях начальное приближение к решению определяется стандартными методами исследования функций или из иных соображений с учетом специфики задачи.

При изучении этой темы функции, входящие в уравнения и системы, будут предполагаться непрерывными и нужное число раз дифференцируемыми.

Метод простой итерации. Пусть задана система

$$f_1(x_1, x_2, \dots, x_n) = 0$$

$$f_2(x_1, x_2, \dots, x_n) = 0$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$f_n(x_1, x_2, \dots, x_n) = 0,$$

или в векторной форме

$$F(X) = 0. \tag{4.1}$$

Преобразуем систему к виду

$$X = G(X) \tag{4.2}$$

и рассмотрим итерационный процесс

$$X^{(k+1)} = G(X^{(k)}). \tag{4.3}$$

Определение. Пусть в полном метрическом пространстве определено отображение $G(X)$. Отображение $G(X)$ называется сжимающим, если для любых X_1 и X_2 из области определения $G(X)$ имеет место

отображение. Тогда уравнение (4.2) имеет единственное решение и итерационный процесс (4.3) сходится к решению системы (4.2).

Доказательство. Единственность решения доказывается элементарно. Пусть есть два решения X_1 и X_2 системы (4.2). Имеем

$$\rho(X_1, X_2) = \rho(G(X_1), C(X_2)) \leq q\rho(X_1, X_2),$$

т. е. $\rho(X_1, X_2)(1 - q) \leq 0$. Следовательно, $\rho(X_1, X_2) = 0$.

Достаточность. Докажем, что последовательность $X^{(k)}$ фундаментальна. Поскольку

$$\rho(X^{(k+1)}, X^{(k)}) \leq q\rho(X^{(k)}, X^{(k-1)}) \leq q^k a,$$

где $a = \rho(X^{(1)}, X^{(0)})$,

и силу неравенства треугольника

$$\begin{aligned} \rho(X^{(k+p)}, X^{(k)}) &\leq \rho(X^{(k+p)}, X^{(k+p-1)}) + \rho(X^{(k+p-1)}, X^{(k+p-2)}) + \dots \\ &+ \rho(X^{(k+1)}, X^{(k)}) \leq aq^k(1 + q + \dots + q^{p-1}) = aq^k \frac{1 - q^p}{1 - q}. \end{aligned}$$

Следовательно, $\rho(X^{(k+1)}, X^{(k)}) < \varepsilon$ при $k > k_0$. Согласно критерию Коши, последовательность $X^{(k)}$ имеет предел, и при $p \rightarrow \infty$ из последнего неравенства имеем

$$\rho(\bar{X}, X^{(k)}) \leq aq^k \frac{1}{1 - q}.$$

Теорема доказана.

Метод секущих. Этот достаточно простой метод применяется при нахождении корней уравнения

$$f(x) = 0.$$

Пусть на концах некоторого отрезка $[x_1, x_2]$ функция $f(x)$ принимает значения разных знаков. Это означает, что внутри отрезка имеется, по крайней мере, один корень уравнения (4.1). Для его отыскания проведем прямую через точки $M_1(x_1, f(x_1))$ и $M_2(x_2, f(x_2))$. Уравнение этой прямой

$$y - f(x_2) = \frac{x - x_2}{x_2 - x_1}(f(x_2) - f(x_1)).$$

Точка x_3 пересечения этой прямой с осью x равна

$$x_3 = x_2 - f(x_2) \frac{x_2 - x_1}{f(x_2) - f(x_1)}.$$

Вычислим значения функции $f(x)$ в точке x_3 и в качестве следующего отрезка для построения секущей возьмем тот из отрезков $[x_1, x_2]$ и $[x_3, x_2]$, на концах которого функция $f(x)$ принимает значения разных знаков. В результате приходим к алгоритму

$$x_{k+1} = x_k - f(x_k) \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}. \quad (4.4)$$

Сравним знаки функции $f(x)$ в точках x_{k-1} и x_{k+1} и в точках x_k и x_{k+1} . В качестве точек x_k и x_{k-1} следующего шага выбирается та пара точек, в которой значение функции $f(x)$ противоположно по знаку. Сходимость этого алгоритма для непрерывной функции очевидна.

Метод Ньютона. Приведем эвристические соображения, связанные с построением этого метода. Пусть \bar{x} — решение уравнения

$$f(x) = 0 \quad (4.1')$$

и x_0 — приближение к \bar{x} , принадлежащее некоторой области $\Omega : \{|\bar{x} - x| < a\}$, и пусть

$$\bar{x} = h + x_0.$$

Разложим $f(\bar{x}) = f(x_0 + h)$ по степеням h , ограничиваясь лишь линейными членами.

$$f(\bar{x}) = f(x_0 + h) = f(x_0) + hf'(x_0) + O(h^2) = 0.$$

Отсюда

$$h = -\frac{f(x_0)}{f'(x_0)} + O(h^2).$$

Поэтому

$$\bar{x} = x_0 - \frac{f(x_0)}{f'(x_0)} + O(h^2).$$

Следовательно, величина

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

будет отличаться от решения \bar{x} уже на величину $O(h^2)$. Эти соображения приводят к построению итерационного процесса уточнения начального приближения

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}. \quad (4.5)$$

Метод Ньютона иногда называют методом касательных, поскольку следующая точка определяется как точка пересечения касательной к $f(x)$ в точке x_k . В самом деле, уравнение касательной к $f(x)$ в этой точке имеет вид

$$y - f(x_k) = f'(x_k)(x - x_k)$$

и точка x_{k+1} есть точка пересечения касательной к $f(x)$ с осью ox .

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

Аналогичные соображения приводят к построению метода Ньютона для систем.

Пусть задана система уравнений

$$\begin{aligned} f_1(x_1, x_2, \dots, x_n) &= 0 \\ f_2(x_1, x_2, \dots, x_n) &= 0 \\ &\dots \dots \dots \\ f_n(x_1, x_2, \dots, x_n) &= 0. \end{aligned} \tag{4.6}$$

Или в векторной форме

$$F(X) = 0. \tag{4.6'}$$

Пусть вектор $\bar{X} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)^T$ есть решение системы и $\bar{X} = X_0 + h$, то есть

$$\bar{X} = (x_1^0 + h_1, x_2^0 + h_2, \dots, x_n^0 + h_n)^T.$$

Разложим $f(x_1^0 + h_1, x_2^0 + h_2, \dots, x_n^0 + h_n)$ по степеням h_i , ограничиваясь линейными членами. Так же, как и для одного уравнения, получим

$$\begin{aligned} f_1(x_1^0, x_2^0, \dots, x_n^0) + \frac{\partial f_1(x_1^0 \dots x_n^0)}{\partial x_1} h_1 + \frac{\partial f_1(x_1^0 \dots x_n^0)}{\partial x_2} h_2 + \dots + \frac{\partial f_1(x_1^0 \dots x_n^0)}{\partial x_n} h_n + O(h^2) &= 0 \\ f_2(x_1^0, x_2^0, \dots, x_n^0) + \frac{\partial f_2(x_1^0 \dots x_n^0)}{\partial x_1} h_1 + \frac{\partial f_2(x_1^0 \dots x_n^0)}{\partial x_2} h_2 + \dots + \frac{\partial f_2(x_1^0 \dots x_n^0)}{\partial x_n} h_n + O(h^2) &= 0 \\ &\dots \dots \dots \\ f_n(x_1^0, x_2^0, \dots, x_n^0) + \frac{\partial f_n(x_1^0 \dots x_n^0)}{\partial x_1} h_1 + \frac{\partial f_n(x_1^0 \dots x_n^0)}{\partial x_2} h_2 + \dots + \frac{\partial f_n(x_1^0 \dots x_n^0)}{\partial x_n} h_n + O(h^2) &= 0. \end{aligned}$$

Пусть определитель матрицы Якоби

$$W(X) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} \quad (4.7)$$

отличен от нуля. Найдем из этой системы вектор h

$$h = W^{-1}(x_0)F(x_0) + O(h^2).$$

Поэтому вектор

$$X_1 = X_0 - W^{-1}(X_0)F(X_0)$$

будет отличаться от X_0 на величину порядка $O(h^2)$. Эти соображения приводят к построению итерационного процесса

$$X_{k+1} = X_k - W^{-1}(X_k)F(X_k). \quad (4.8)$$

В случае скалярного уравнения легко доказывается теорема о сходимости метода Ньютона.

Теорема 2. Пусть \bar{x} — решение уравнения (4.1') и пусть в некоторой области $\Omega_a : |\bar{x} - x| \leq a$ выполнены условия

1. Первая производная отделена от 0: $|f'(x)|^{-1} < b$.
2. Вторая производная ограничена: $|f''(x)| < d$.

Тогда метод скорейшего спуска сходится в области Ω_a при

$$a < \frac{1}{bd}.$$

Доказательство. Рассмотрим выражение

$$f(\bar{x}) - f(x_k) - f'(x_k)(\bar{x} - x_k).$$

Поскольку $f(\bar{x}) = 0$ и так как в силу (4.5) $f'(x_k)(x_{k+1} - x_k) = -f(x_k)$

$$|f(\bar{x}) - f(x_k) - f'(x_k)(\bar{x} - x_k)| = |f'(x_k)| |x_{k+1} - \bar{x}|. \quad (4.9)$$

С другой стороны, по теореме Лагранжа

$$|f(\bar{x}) - f(x_k) - f'(x_k)(\bar{x} - x_k)| = |(f'(\xi) - f'(x_k))(\bar{x} - x_k)| = |(f''(\eta)(\bar{x} - x_k)(\eta - x_k))|.$$

Здесь $\bar{x} \leq \xi \leq x_k$, $\xi \leq \eta \leq x_k$. Следовательно,

$$|f(\bar{x}) - f(x_k) - f'(x_k)(\bar{x} - x_k)| \leq |(f''(\eta))| |\bar{x} - x_k|^2. \quad (4.10)$$

Сравнивая (4.9) и (4.10), получаем

$$|f'(x_k)| |x_{k+1} - \bar{x}| \leq |(f''(\eta))| |\bar{x} - x_k|^2. \quad (4.11)$$

Пусть $c = bd$. Неравенство (4.11) означает, что

$$|x_{k+1} - \bar{x}| \leq c |\bar{x} - x_k|^2. \quad (4.12)$$

Пусть $\rho_k = c|x_k - \bar{x}|$. Умножая (4.12) на c , получим $\rho_{k+1} \leq \rho_k^2$ и

$$\rho_{k+1} \leq \rho_k^2 \leq \rho_{k-1}^4 \leq \dots \leq \rho_0^{2^k}.$$

Следовательно,

$$|\bar{x} - x_{k+1}| \leq c^{-1} (c |\bar{x} - x_0|)^{2^k}.$$

Для сходимости надо, чтобы $c |\bar{x} - x_0| < 1$, т.е. область Ω_a должна быть такой, что $a < \frac{1}{bd}$. Теорема доказана.

Метод наискорейшего градиентного спуска. Известно достаточно большое число различных градиентных методов. Мы рассмотрим один из простейших вариантов метода наискорейшего градиентного спуска.

Системе (4.6) поставим в соответствие функцию

$$\Psi(x) = f_1^2(x_1, \dots, x_n) + \dots + f_n^2(x_1, \dots, x_n). \quad (4.13)$$

Решение системы (4.6) сообщает этой функции нулевой минимум, и, обратно, значения вектора $X = (x_1, x_2, \dots, x_n)^T$, в котором функция $\Psi(x)$ достигает значения нулевого минимума, является решением системы (4.6).

Построим следующий итерационный процесс

$$X^{(k+1)} = X^{(k)} - \lambda_k V^{(k)},$$

где $V^{(k)} = \text{grad}\Psi(X^{(k)})$.

Таким образом, поскольку $X^{(k+1)} - X^{(k)} = -\lambda_k \text{grad}\Psi(X^{(k)})$, движение от точки $X^{(k)}$ к точке $X^{(k+1)}$ происходит в направлении наибольшего убывания функции $\Psi(X)$ на некоторый шаг λ_k . Величину λ_k следует определить из условия минимума функции $\Psi(X^{(k)} - \lambda V^{(k)})$, что опять приводит к решению нелинейной задачи отыскания корня производной функции $\Psi(\lambda) = \Psi(X^{(k)} - \lambda V^{(k)})$. В простейшем варианте метода для отыскания λ приближенно отыскивается корень уравнения

$$\Psi(\lambda) = 0.$$

Если разложить $\Psi(\lambda)$ по степеням λ , пренебрегая квадратами и высшими степенями λ

$$\Psi(\lambda) \approx \Psi(0) + \lambda \Psi'(0),$$

то

$$\lambda_k = -\frac{\Psi(0)}{\Psi'_\lambda(0)}.$$

Поскольку

$$\Psi(\lambda) = \Psi(x_1 - \lambda v_1, \dots, x_n - \lambda v_n),$$

где

$$v_i = \frac{\partial \Psi}{\partial x_i} \quad \Psi'(\lambda) = \sum_{i=1}^n \frac{\partial \Psi}{\partial x_i} \frac{dx_i}{d\lambda} = \sum_{i=1}^n \left(\frac{\partial \Psi}{\partial x_i} \right)^2 = (\text{grad} \Psi(X^{(k)}), \text{grad} \Psi(X^{(k)})).$$

В итоге приходим к следующему итерационному процессу

$$X^{(k+1)} = X^{(k)} - \lambda_k V^{(k)}, \quad (4.14)$$

где

$$V^{(k)} = \text{grad} \Psi(X^{(k)}), \quad \Psi(X) = f_1^2 + f_2^2 + \dots + f_n^2, \quad \lambda_k = \frac{\Psi(X^{(k)})}{(V^{(k)}, V^{(k)})}. \quad (4.15)$$

Следует еще раз подчеркнуть, что при неудачном выборе начального приближения к решению системы построенный процесс может расходиться или сходиться, но не к решению системы, о чем свидетельствуют допущения при вычислении λ .

Отметим без доказательства, что более точный метод отыскания λ как корень уравнения $\Psi'(\lambda) = 0$ приводит к формуле

$$x^{(k+1)} = x^{(k)} - \lambda_k W'(x^{(k)}) F(x^{(k)}), \quad (4.16)$$

где

$$\lambda_k = \frac{(F(x^{(k)}), W(x^{(k)}) W'(x^{(k)}) F(x^{(k)}))}{(W(x^{(k)}) W'(x^{(k)}) F(x^{(k)}), W(x^{(k)}) W'(x^{(k)}) F(x^{(k)}))}. \quad (4.17)$$

Здесь вектор $F = (f_1, f_2, \dots, f_n)^T$, W — матрица Якоби, определенная в (4.7), а W' — транспонированная матрица Якоби.

§ 5. Интерполяционный многочлен Лагранжа

Интерполяционный многочлен Лагранжа. Пусть на отрезке $[a, b]$ задана функция $f(x)$ и пусть внутри этого отрезка заданы n различных точек x_1, x_2, \dots, x_n . Требуется построить многочлен минимальной степени

$$P_k(x) = a_0x^k + a_1x^{k-1} + \dots + a_{k-1}x + a_k,$$

такой, что

$$P_k(x_j) = f(x_j) \quad j = 1, 2, \dots, n. \quad (5.1)$$

Теорема 1. Существует единственный многочлен степени $n - 1$, такой, что $P_n(x_i) = f(x_i) \quad i = 1, \dots, n$.

Доказательство. Условие (5.1) означает, что

$$a_0x_1^{n-1} + a_1x_1^{n-2} + \dots + a_{n-1} = f(x_1)$$

$$a_0x_2^{n-1} + a_1x_2^{n-2} + \dots + a_{n-1} = f(x_2)$$

$$\dots \dots \dots$$

$$a_0x_n^{n-1} + a_1x_n^{n-2} + \dots + a_{n-1} = f(x_n).$$

Эта система из n уравнений относительно n неизвестных a_0, a_1, \dots, a_{n-1} . Определитель этой системы

$$\begin{vmatrix} x_1^{n-1} & x_1^{n-2} & \dots & 1 \\ x_2^{n-1} & x_2^{n-2} & \dots & 1 \\ \dots & \dots & \dots & \dots \\ x_n^{n-1} & x_n^{n-2} & \dots & 1 \end{vmatrix}$$

есть определитель Вандермонда. Он отличен от нуля при $x_i \neq x_j$, если $i \neq j$. Следовательно, система имеет единственное решение относительно a_0, a_1, \dots, a_{n-1} . Теорема доказана.

Многочлены, удовлетворяющие условию (5.1), называют интерполяционными многочленами. Лагранж предложил простой способ построения таких многочленов.

Рассмотрим

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

Многочлен $\omega_{i,n}(x)$ степени $n-1$ в точках x_j равен 0 при $j \neq i$, а в точке x_i равен 1. Следовательно, многочлен

$$L_n(x) = \sum_{i=1}^n f(x_i) \frac{\prod_{\substack{j=1 \\ j \neq i}}^n (x - x_j)}{\prod_{\substack{j=1 \\ j \neq i}}^n (x_i - x_j)} \quad (5.2)$$

степени $n-1$ удовлетворяет условиям (5.1). Многочлен, определяемый формулой (5.2), называется интерполяционным многочленом Лагранжа. Введем в рассмотрение многочлен

$$\omega_n(x) = (x - x_1)(x - x_2) \dots (x - x_n). \quad (5.3)$$

С его помощью интерполяционный многочлен Лагранжа можно записать следующим образом

$$L_n(x) = \sum_{i=1}^n f(x_i) \frac{\omega_n(x)}{(x - x_i) \omega'_n(x_i)}. \quad (5.3')$$

В самом деле,

$$\omega'_n(x) = \sum_{i=1}^n \prod_{\substack{j=1 \\ j \neq i}}^n (x - x_j),$$

и при $x = x_i$ в сумме все слагаемые равны нулю, кроме одного, где отсутствует множитель $x - x_i$.

Величину $\varphi(x) = L_N(x) - f(x)$ называют остаточным членом интерполяционной формулы

$$L_N(x) \approx f(x).$$

Следующая теорема является важным инструментом для оценки величины $\varphi(x)$.

Теорема 2. Для любой точки $z \neq x_i$ найдется такая точка ξ , что

$$f(z) - L_n(z) = \frac{f^{(n)}(\xi)}{n!} \omega_n(z). \quad (5.4)$$

Доказательство. Рассмотрим функцию

$$\Psi(x) = f(x) - L_n(x) - A \omega_n(x),$$

где A — некоторая константа. Эта функция обращается в нуль в точках x_i . Выберем константу A так, чтобы эта функция была равна нулю еще в одной точке z . Тогда

$$A = \frac{f(z) - L_n(z)}{\omega_n(z)}.$$

С другой стороны, функция обращается в нуль в n точках x_i и точке z , т. е. в $n+1$ точке. По теореме Ролля между двумя последовательными нулями функции есть, по крайней мере, одна точка, где производная этой функции обращается в нуль. Следовательно, для $\Psi'(x)$ имеется не менее n точек $\xi_1, \xi_2, \dots, \xi_n$, где $\Psi'(\xi) = 0$ обращается в нуль. Продолжая эти рассуждения, замечаем, что существует хотя бы одна точка ξ , в которой $\Psi^{(n)}(\xi) = 0$. Но $L_n(x)$ — многочлен степени $n-1$, а $\omega_n(x)$ — многочлен степени n со старшим коэффициентом 1. Поэтому

$$\Psi^{(n)}(x) = f^{(n)}(x) - An!.$$

Равенство нулю этого выражения в точке ξ приводит к тому, что

$$A = \frac{f^{(n)}(\xi)}{n!}.$$

Сравнение двух полученных выражений для A и доказывает теорему.

§ 6. Минимизация остаточного члена. Многочлены Чебышева

Займемся задачей минимизации остаточного члена интерполяционной формулы. Далее, если не оговорено иное, под нормой функции понимается

$$\|f\| = \max_{a \leq x \leq b} |f(x)|.$$

Из (5.4) следует

$$\|f - L_n\| \leq \frac{\|f^{(n)}\|}{n!} \|\omega_n\|. \quad (5.4')$$

Поэтому задача минимизации остаточного члена для широкого класса функций связана с нахождением такого многочлена

$$\omega_n(x) = (x - x_1)(x - x_2) \dots (x - x_n),$$

норма которого минимальна. Иными словами, требуется так подобрать узлы интерполяции x_i , чтобы минимизировать $\|\omega_n\|$. Эта задача решается с помощью многочленов Чебышева.

Многочлены Чебышева. При $-1 \leq x \leq 1$ рассмотрим функцию

$$T_n(x) = \cos n \arccos x. \quad (6.1)$$

Заметим, что $T_0(x) = 1$, $T_1(x) = x$.

Теорема 1. $T_n(x)$ есть многочлен степени n со старшим коэффициентом 2^{n-1} , четная функция при четном n и нечетная при нечетном n и удовлетворяющий рекуррентному соотношению

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x). \quad (6.2)$$

Доказательство. Докажем рекуррентное соотношение (6.2), из которого следуют остальные свойства.

Пусть $\alpha = \arccos x$. Тогда

$$T_{n+1}(x) + T_{n-1}(x) = \cos(n+1)\alpha + \cos(n-1)\alpha = 2 \cos n\alpha \cos \alpha = 2xT_n(x).$$

Остальные утверждения легко доказываются по индукции с использованием рекуррентного соотношения (6.2).

В самом деле, докажем, что старший коэффициент $T_n(x)$ равен 2^{n-1} . Старший коэффициент $T_1(x)$ равен 1. Пусть старший коэффициент T_n равен 2^{n-1} . Тогда старший коэффициент T_{n+1} определяется старшим коэффициентом выражения $2xT_n(x)$, который равен 2^n .

Так же по индукции доказывается свойство четности — нечетности. Заметим, что $T_0(x) = 1$, т. е. четная функция, $T_1(x) = x$ — нечетная функция. Предположим, например, что $T_{n-1}(x)$ — нечетная функция, а $T_n(x)$ — четная функция. Из (6.2) получаем

$$T_{n+1}(-x) = -2xT_n(-x) - T_{n-1}(x) = -2xT_n(x) + T_{n-1}(x) = -T_{n+1}(x).$$

Определение. Многочлен

$$T_n(x) = \cos n \arccos x, \quad (6.3)$$

определенный на отрезке $[-1, 1]$, называется многочленом Чебышева.

Следует заметить, что в некоторых учебниках определение многочлена Чебышева дается с использованием рекуррентного соотношения (6.2). Именно пусть $T_0(x) = 1$, $T_1(x) = x$, и при любом n $T_n(x)$ определим из (6.2). Эти два определения эквивалентны. В одну сторону этот факт мы установили. При изучении разностных уравнений будет показано (пример 2 § 19), что решением разностного уравнения (6.2) на отрезке $[-1, 1]$ является функция $\cos n \arccos x$.

Свойства многочлена Чебышева.

$$\begin{aligned} 1. \quad T_n(x) &= \frac{(x - \sqrt{x^2 - 1})^n + (x + \sqrt{x^2 - 1})^n}{2}. \\ 2. \quad \int_{-1}^1 \frac{T_m(x)T_n(x)dx}{\sqrt{1-x^2}} &= \begin{cases} 0 & n \neq m \\ \frac{\pi}{2} & n = m \neq 0 \\ \pi & n = m = 0 \end{cases} \end{aligned}$$

Первое свойство доказывается с использованием формулы $\cos n\alpha = \frac{e^{ian} + e^{-ian}}{2}$, второе свойство — непосредственным интегрированием. В качестве упражнения предлагается доказать эти свойства самостоятельно.

Найдем корни многочлена Чебышева. Из уравнения

$$T_n(x) = \cos n \arccos x = 0$$

получаем

$$n \arccos x = \frac{\pi}{2} + k\pi \quad x_k = \cos \frac{(2k+1)\pi}{2n} \quad k = 0, 1, \dots, n-1. \quad (6.4)$$

Поскольку $T_n(x) = \cos n \arccos x$, то $|T_n(x)| \leq 1$.

Найдем точки, где $|T_n(x)| = 1$. Имеем

$$n \arccos x_m = m\pi \quad x_m = \cos \frac{m\pi}{n} \quad m = 0, 1, \dots, n. \quad (6.5)$$

Подчеркнем, что в (6.4) получено n точек, т. е. найдены все корни $T_n(x)$. Точек же максимума модуля — $n + 1$, что будет использовано в дальнейшем.

Оптимальный выбор узлов. Многочлены, наименее уклоняющиеся от нуля

Определение. Рассмотрим класс многочленов фиксированной степени n и с фиксированным старшим коэффициентом. Многочлен $Q_n(x)$ называется многочленом, наименее уклоняющимся от нуля на отрезке $[a, b]$, если на этом отрезке он имеет наименьший максимум модуля среди всех многочленов данного класса.

Таким образом, многочлен $Q_n(x)$ наименее уклоняется от нуля на $[a, b]$, если

$$\max_{a \leq x \leq b} |Q_n(x)| \leq \max_{a \leq x \leq b} |P_n(x)|,$$

где $P_n(x)$ — всевозможные многочлены той же степени и с тем же старшим коэффициентом.

Теорема 2. Многочлен $\bar{T}_n(x) = 2^{1-n}T_n(x)$ является многочленом, наименее уклоняющимся от нуля на отрезке $[-1, 1]$, среди многочленов со старшим коэффициентом, равным единице.

Доказательство. Предположим противное, что существует многочлен $P_n(x)$, такой, что

$$\max_{a \leq x \leq b} |P_n(x)| \leq \max_{a \leq x \leq b} |\bar{T}_n(x)|.$$

Рассмотрим многочлен

$$Q_{n-1}(x) = \bar{T}_n(x) - P_n(x).$$

Поскольку старшие коэффициенты у $P_n(x)$ и $\bar{T}_n(x)$ одинаковы, степень многочлена $Q_{n-1}(x) = n-1$. Рассмотрим значение $Q_{n-1}(x)$ в точках x_m , т. е. в точках, определяемых равенством (6.5). В этих точках $\bar{T}_n(x)$ имеет максимум модуля и знак многочлена $Q_{n-1}(x)$ будет определяться знаком многочлена $\bar{T}_n(x)$. Но в этих точках $\bar{T}_n(x_k) = \pm 2^{1-n}$. Следовательно, многочлен $Q_{n-1}(x)$ имеет $n + 1$ переменну знака и по теореме Больцано —

Коши имеет не менее n корней. Но это невозможно, если $Q_{n-1}(x)$ не есть тождественный нуль. Теорема доказана.

Таким образом, для отрезка $[-1, 1]$ можно утверждать, что если в качестве узлов интерполяции выбрать корни многочлена Чебышева, определяемые формулой (6.4), то норма остаточного члена интерполяционной формулы будет минимальна и имеет место оценка

$$\|f(x) - L_n\| \leq \frac{\|f^n\|}{2^{n-1}n!}.$$

Займемся задачей минимизации остаточного члена для произвольного отрезка $[a, b]$. Рассмотрим на этом промежутке многочлен $P_n(x)$. Замена

$$x = \frac{b-a}{2}t + \frac{b+a}{2}$$

переводит отрезок $[-1, 1]$ в отрезок $[a, b]$. С помощью обратной замены

$$t = \frac{2x}{b-a} - \frac{b+a}{b-a}$$

получим многочлен

$$P_n(t) = P_n\left(\frac{2x}{b-a} - \frac{b+a}{b-a}\right),$$

аргумент которого меняется на $[-1, 1]$. Но у многочлена

$$\bar{T}_n\left(\frac{2x}{b-a} - \frac{b+a}{b-a}\right),$$

при такой замене изменится старший коэффициент. Для того чтобы он был равен 1, многочлен надо умножить на величину $\frac{(b-a)^n}{2^n}$. В итоге получаем многочлен

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

старший коэффициент которого равен 1 и который наименее уклоняется от нуля на $[a, b]$. Если в качестве корней многочлена выбрать

$$x_k = \frac{b-a}{2} \cos \frac{(2k+1)\pi}{2n} + \frac{b+a}{2} \quad k = 0, 1, \dots, n-1,$$

получим многочлен $\bar{T}_{n,(a,b)}(x)$

$$\bar{T}_{n,(a,b)}(x) = \frac{(b-a)^n}{2^{2n-1}} T_n\left(\frac{2x}{b-a} - \frac{b+a}{b-a}\right),$$

наименее уклоняющийся от нуля на $[a, b]$.

Итогом этих рассуждений является

Теорема 3. На промежутке $[a, b]$ многочленом, наименее уклоняющимся от нуля, является многочлен

$$\bar{T}_{n,(a,b)}(x) = \frac{(b-a)^n}{2^{2n-1}} T_n\left(\frac{2x}{b-a} - \frac{b+a}{b-a}\right),$$

и при оптимальном выборе узлов интерполяции

$$\|f - L_n\| \leq \frac{(b-a)^n}{2^{2n-1}n!} \|f^{(n)}\|.$$

§ 7. Разделенные разности.

Интерполяционный многочлен в форме Ньютона

Вопросы, изучаемые в этом параграфе, необходимы для решения задачи построения интерполяционного многочлена, который совпадает с функцией и ее производными до некоторого порядка в заданных точках. Эта задача решается с помощью интерполяционного многочлена в форме Ньютона.

Определение. Будем называть разделенной разностью нулевого порядка значение функции $f(x_i)$ в точке x_i . Разделенную разность первого порядка определим равенством

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

Разделенную разность k -го порядка определим с помощью разностей $k-1$ -го порядка следующим образом:

$$f(x_1; x_2; \dots; x_{k+1}) = \frac{f(x_2; \dots; x_{k+1}) - f(x_1; \dots; x_k)}{x_{k+1} - x_1}.$$

Теорема 1. Для разделенных разностей справедливо равенство

$$f(x_1; x_2; \dots; x_k) = \sum_{i=1}^n \frac{f(x_i)}{\prod_{\substack{j=1 \\ j \neq i}}^n (x_i - x_j)}. \quad (7.1)$$

Доказательство будем проводить по индукции. Для разделенной разности первого порядка имеем:

$$f(x_1; x_2) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} = \frac{f(x_1)}{x_1 - x_2} + \frac{f(x_2)}{x_2 - x_1}.$$

Предположим, что формула верна для разделенной разности $k-1$ -го порядка. Тогда для разделенной разности k -го порядка имеем:

$$f(x_1; x_2; \dots; x_{k+1}) = \frac{f(x_2; x_3; \dots; x_{k+1}) - f(x_1; x_2; \dots; x_k)}{x_{k+1} - x_1}.$$

В соответствии с индуктивным предположением

$$\begin{aligned}
f(x_1; x_2; \dots : x_{k+1}) &= \frac{1}{(x_{k+1} - x_1)} \left(\sum_{i=2}^{k+1} \frac{f(x_i)}{\prod_{\substack{j=2 \\ j \neq i}}^{k+1} (x_i - x_j)} - \sum_{i=1}^k \frac{f(x_i)}{\prod_{\substack{j=1 \\ j \neq i}}^k (x_i - x_j)} \right) = \\
&= \frac{1}{(x_{k+1} - x_1)} \left(\frac{f(x_{k+1})}{\prod_{j=2}^k (x_{k+1} - x_j)} - \frac{f(x_1)}{\prod_{j=2}^k (x_1 - x_j)} + \sum_{i=2}^k \frac{f(x_i)}{\prod_{\substack{j=2 \\ j \neq i}}^{k+1} (x_i - x_j)} - \right. \\
&\quad \left. - \sum_{i=2}^k \frac{f(x_i)}{\prod_{\substack{j=1 \\ j \neq i}}^k (x_i - x_j)} \right).
\end{aligned}$$

Рассмотрим

$$\sum_{i=2}^k \frac{f(x_i)}{\prod_{\substack{j=2 \\ j \neq i}}^{k+1} (x_i - x_j)} - \sum_{i=2}^k \frac{f(x_i)}{\prod_{\substack{j=1 \\ j \neq i}}^k (x_i - x_j)} = \sum_{i=2}^k \left(\frac{f(x_i)}{\prod_{\substack{j=2 \\ j \neq i}}^{k+1} (x_i - x_j)} - \frac{f(x_i)}{\prod_{\substack{j=1 \\ j \neq i}}^k (x_i - x_j)} \right)$$

и рассмотрим множители при $f(x_i)$

$$\left(\frac{1}{\prod_{\substack{j=2 \\ j \neq i}}^{k+1} (x_i - x_j)} - \frac{1}{\prod_{\substack{j=1 \\ j \neq i}}^k (x_i - x_j)} \right) = \frac{x_i - x_1 - x_i + x_{k+1}}{\prod_{\substack{j=1 \\ j \neq i}}^{k+1} (x_i - x_j)} = \frac{x_{k+1} - x_1}{\prod_{\substack{j=1 \\ j \neq i}}^{k+1} (x_i - x_j)}.$$

Поэтому

$$\begin{aligned}
&f(x_1; x_2; \dots : x_{k+1}) = \\
&= \frac{1}{(x_{k+1} - x_1)} \frac{f(x_{k+1})}{\prod_{j=2}^k (x_{k+1} - x_j)} + \frac{1}{(x_1 - x_{k+1})} \frac{f(x_1)}{\prod_{j=2}^k (x_1 - x_j)} + \sum_{i=2}^k \frac{f(x_i)}{\prod_{\substack{j=1 \\ j \neq i}}^{k+1} (x_i - x_j)}.
\end{aligned}$$

В итоге получаем

$$\left(\prod_{n=1}^j (x - x_j) \right) \left(\sum_{n=1}^i \frac{\prod_{n=1}^{j \neq i} (x - x_j)}{f(x_i)} (x - x_i) + \frac{\prod_{n=1}^j (x - x_j)}{f(x)} \right) =$$

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

В самом деле,

$$(7.3) \quad f(x) - L_n(x) = f(x, x_1, \dots, x_n) \omega_n(x).$$

Показательство. Заметим, что

$$(7.2) \quad L_n(x) = f(x_1) + f(x_1, x_2) + \dots + f(x_1, x_2, \dots, x_n) + \dots + f(x_1, x_2, \dots, x_n)(x - x_1)(x - x_1) \dots (x - x_{n-1}).$$

Теорема 2. Интерполяционный многочлен представим в форме Ньютона

позволит решить задачу интерполяции с кратными узлами.

С помощью аппарата разложенных разностей интерполяционный многочлен

$$(\alpha_1 f_1 + \alpha_2 f_2)(x_1, \dots, x_n) = \alpha_1 f_1(x_1, \dots, x_n) + \alpha_2 f_2(x_1, \dots, x_n).$$

Следствие 2. Разделенная разность есть линейный оператор от функции

$$f(x_1, \dots, x_i, \dots, x_j, \dots, x_k) = f(x_1, \dots, x_j, \dots, x_k, \dots, x_i).$$

аргументов.

Следствие 1. Разделенная разность не меняется при перестановке своих

Из формулы (7.1) вытекает ряд следствий.

Теорема доказана.

$$f(x_1, x_2, \dots, x_k) = \sum_{n=1}^i \frac{\prod_{n=1}^{j \neq i} (x_i - x_j)}{f(x_i)}.$$

С учетом того, что $\omega_n(x) = (x - x_1)(x - x_1)\dots(x - x_n)$, равенство (7.3) доказано.

С другой стороны,

$$L_n(x) = (L_n(x) - L_{n-1}(x)) + (L_{n-1}(x)) + \dots + (L_2(x) - L_1(x)) + L_1(x) \quad (7.4)$$

Но разность $L_k(x) - L_{k-1}(x)$ есть многочлен степени $k-1$, обращающийся в нуль в точках x_i $i = 1, 2, \dots, k-1$. Поэтому

$$L_k(x) - L_{k-1}(x) = A_{k-1}\omega_{k-1}(x), \quad (7.5)$$

где $\omega_{k-1}(x) = (x - x_1)(x - x_2)\dots(x - x_{k-1})$.

Рассмотрим $L_k(x) - L_{k-1}(x)$ в точке x_k .

$$f(x_k) - L_{k-1}(x_k) = A_{k-1}\omega_{k-1}(x_k). \quad (7.5')$$

Отсюда

$$A_{k-1} = \frac{f(x_k) - L_{k-1}(x_k)}{\omega_{k-1}(x_k)}.$$

В силу (7.3)

$$A_{k-1} = f(x_k; x_1; \dots; x_{k-1}). \quad (7.6)$$

Поэтому (7.4) с учетом (7.5) и (7.6) и того, что $L_1(x) = f(x_1)$, можно записать следующим образом

$$L_n(x) = f(x_1) + f(x_1; x_2)\omega_1(x) + f(x_1; x_2; x_3)\omega_2(x) + \dots + f(x_1; x_2; \dots; x_n)\omega_{n-1}(x).$$

Теорема доказана.

Следствие. Для разделенной разности $f(x_1; x_2 : \dots; x_k)$ справедливо представление

$$f(x_1; x_2 : \dots; x_k) = \frac{f(x_k) - L_{k-1}(x_k)}{\omega_{k-1}(x_k)}. \quad (7.7)$$

Практическое построение интерполяционного многочлена в форме Ньютона можно проводить с помощью формулы (7.1), вычислив соответствующие разделенные разности. Существует схема Эйткина для построения разделенных разностей и интерполяционного многочлена в форме Ньютона, изложенная, например, в [1].

§ 8. Интерполирование с кратными узлами

В этом параграфе будет рассмотрена задача построения интерполяционного многочлена, совпадающего с заданной функцией и ее производными до некоторого порядка в точках x_i .

Итак, требуется построить многочлен $P_s(x)$ степени $s - 1$, удовлетворяющий условиям

$$P_s(x_1) = f(x_1), \quad P'_s(x_1) = f'(x_1), \quad \dots, \quad P_s^{(l_1-1)}(x_1) = f^{(l_1-1)}(x_1)$$

$$P_s(x_2) = f(x_2), \quad P'_s(x_2) = f'(x_2), \quad \dots, \quad P_s^{(l_2-1)}(x_2) = f^{(l_2-1)}(x_2)$$

$$\dots \dots \dots$$

$$P_s(x_m) = f(x_m), \quad P'_s(x_m) = f'(x_m), \quad \dots, \quad P_s^{(l_m-1)}(x_m) = f^{(l_m-1)}(x_m).$$

Здесь $s = l_1 + l_2 + \dots + l_m$.

Рассмотрим пример. Пусть требуется построить многочлен, который совпадает с заданной функцией в точке x_1 , а в точке x_2 совпадает с функцией и ее производными до второго порядка.

$$P_3(x_1) = f(x_1), \quad P_3(x_2) = f(x_2), \quad P'_3(x_2) = f'(x_2), \quad P''_3(x_2) = f''(x_2).$$

В ϵ -окрестность точки x_2 добавим еще 2 точки x_2^1 и x_2^2 и построим по точкам x_1, x_2, x_2^1, x_2^2 интерполяционный многочлен $P_3(x)$

$$P_3(x) = f(x_1) + f(x_1; x_2)(x - x_1) + \\ + f(x_1; x_2; x_2^1)(x - x_1)(x - x_2) + f(x_1; x_2; x_2^1; x_2^2)(x - x_1)(x - x_2)(x - x_2^1).$$

Устремим теперь ϵ к нулю. При этом $x_2^1 \rightarrow x_2, \quad x_2^2 \rightarrow x_2$.

$$P_3(x) = f(x_1) + f(x_1; x_2)(x - x_1) + \\ + f(x_1; x_2; x_2)(x - x_1)(x - x_2) + f(x_1; x_2; x_2; x_2)(x - x_1)(x - x_2)^2.$$

Выясним смысл выражений

$$f(x_1; x_2; x_2), \quad f(x_1; x_2; x_2; x_2).$$

По определению

$$f(x_1; x_2; x_2) = \frac{f(x_2; x_2) - f(x_1; x_2)}{x_2 - x_1}.$$

Но, сравнивая (7.7) и (5.4), замечаем, что

$$f(x_1; x_1) = f'(x_1).$$

Более того, если точка перечисляется m раз, то

$$f(\underbrace{x_i; x_i; \dots; x_i}_{m \text{ раз}}) = \frac{f^{(m-1)}(x_i)}{(m-1)!}. \quad (8.1)$$

Поэтому

$$f(x_1; x_2; x_2) = \frac{f'(x_2) - f(x_1; x_2)}{x_2 - x_1}.$$

Что касается $f(x_1; x_2; x_2; x_2)$, то

$$f(x_1; x_2; x_2; x_2) = \frac{1}{x_2 - x_1} (f(x_2; x_2; x_2) - f(x_1; x_2; x_2)).$$

Первое слагаемое определено формулой (8.1), а $f(x_1; x_2; x_2)$ только что определено.

В общем случае процесс построения многочлена $P_s(x)$ степени $s-1$ заключается в следующем. В ε -окрестность кратной точки x_j помещаются l_{j-1} вспомогательных точек $x_j^1, x_j^2, \dots, x_j^{l_{j-1}}$. При $\varepsilon \rightarrow 0$ получим многочлен вида

$$\begin{aligned} P_s(x) = & A_1^0 f(x_1) + A_1^1 (x - x_1) + \dots + A_1^{l_1-1} (x - x_1)^{l_1-1} + \\ & + A_2^0 (x - x_1)^{l_1} + A_2^1 (x - x_1)^{l_1} (x - x_2) + \dots + A_2^{l_2-1} (x - x_1)^{l_1-1} (x - x_2)^{l_2-1} + \dots \\ & + A_m^0 \prod_{i=1}^{m-1} (x - x_i)^{l_i} + A_m^1 \prod_{i=1}^{m-1} (x - x_i)^{l_i} (x - x_m) + A_m^2 \prod_{i=1}^{m-1} (x - x_i)^{l_i} (x - x_m)^2 + \\ & \dots + A_m^{l_m} \prod_{i=1}^{m-1} (x - x_i)^{l_i} (x - x_m)^{l_m-1}. \end{aligned} \quad (8.2)$$

Здесь $A_1^0 = 1$. Далее $A_1^k = f(\underbrace{x_1; x_1; \dots; x_1}_{k+1 \text{ раз}})$ и, согласно (8.1),

$$A_1^k = \frac{f^{(k)}(x_1)}{k!} \quad k = 1 \dots l_1 - 1.$$

Коэффициент $A_2^0 = f(\underbrace{x_1; x_1; \dots; x_1}_{l_1 \text{ раз}}; x_2)$. Покажем, что это выражение имеет смысл. Доказательство будем проводить по индукции. Заметим, что $f(x_1; x_2)$

имеет смысл. Это обычная разделенная разность 1-го порядка. Пусть нами определена разность $f(\underbrace{x_1; x_1; \dots; x_1; x_2}_{k \text{ раз}})$. Но

$$f(\underbrace{x_1; x_1; \dots; x_1; x_2}_{k+1 \text{ раз}}) = \frac{1}{x_2 - x_1} (f(\underbrace{x_1; x_1; \dots; x_1; x_2}_{k \text{ раз}}) - f(\underbrace{x_1; x_1; \dots; x_1}_{k \text{ раз}})).$$

Первое слагаемое в скобках определено по индуктивному предположению. Второе слагаемое определено формулой (8.1). Таким образом, коэффициент A_2^0 определен. Продолжая аналогичные рассуждения для других коэффициентов формулы (8.2), приходим к заключению, что все коэффициенты в (8.2) имеют смысл.

Коротко остановимся на оценке остаточного члена интерполяционного многочлена в форме Ньютона. Если построить по точкам x_i и точкам x_i^l интерполяционный многочлен, то, согласно (5.4), получим

$$f(x) - P_s(x) = \frac{f^{(s)}(x)}{s!} \omega_s(x).$$

При $\varepsilon \rightarrow 0$ многочлен $\omega_s(x)$ имеет вид

$$\omega_s = \prod_{i=1}^m (x - x_i)^{l_i}$$

и оценка (5.4) в этом случае имеет вид

$$\|f - P_s\| = \frac{\|f^{(s)}\|}{s!} \|\omega_s\|. \quad (8.3)$$

В заключение этой темы заметим, что задача интерполяции функции многих переменных значительно сложнее и в настоящее время еще недостаточно разработана. Для иллюстрации трудностей при интерполяции функции многих переменных рассмотрим 2 простых примера.

Пусть требуется построить интерполяционный многочлен, совпадающий с функцией двух переменных $f(x, y)$ в трех точках. Ясно, что многочлен 1-го порядка — уравнение плоскости, проходящей через эти три точки, — и будет этим интерполяционным многочленом. Но если эти три точки лежат на одной прямой, то этот многочлен не единственен, их бесконечно много.

Пусть теперь требуется построить интерполяционный многочлен, совпадающий с функцией $f(x, y)$ в четырех точках

$M_1(x_1, y_1)$, $M_2(x_1, y_2)$, $M_3(x_2, y_1)$, $M_4(x_2, y_2)$. Ясно, что это должен быть многочлен двух переменных, степени выше первой, поскольку для построения многочлена 1-й степени $P_1(x, y) = a_0 + a_1x + b_1y$ надо определить три коэффициента a_0, a_1, b_1 . Но функция должна совпадать с многочленом в четырех точках, т. е. получаем систему из четырех уравнений для определения 3-х неизвестных, которая в общем случае неразрешима. Что касается многочлена 2-й степени $P_2(x, y) = a_0 + a_1x + b_1y + 2c_{12}xy + a_2x^2 + b_2y^2$, то для определения шести коэффициентов получаем систему из четырех уравнений, и ясно, что без дополнительных условий решение этой системы, вообще говоря, не единственно. Таким образом, уже на начальном этапе изучения мы сталкиваемся с трудностями, связанными с теоремами существования и единственности интерполяционного многочлена для функции многих переменных. Изучению проблематики теории приближений посвящен ряд работ отечественных и зарубежных математиков. Здесь следует назвать, в первую очередь, работы П.Л. Чебышева, А.Н. Колмогорова, П.С. Урысона. С отдельными результатами по этой теме можно ознакомиться, например, в [1].

§ 9. Ортогональные многочлены и их свойства

Рассмотрим пространство функций таких, что

$$\int_a^b p(x) f^2(x) dx < \infty,$$

где $p(x) > 0$. Зададим в этом пространстве скалярное произведение по формуле

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

Как обычно, если в гильбертовом пространстве для пары элементов имеет место $(f, g) = 0$, то говорят, что эти элементы ортогональны.

Рассмотрим пространство многочленов степени n и в этом пространстве рассмотрим линейно независимую систему элементарных одночленов $1, x, x^2, \dots, x^n$.

Теорема 1. Существует линейно независимая система многочленов со старшим коэффициентом, равным единице,

$$P_k(x) = p_{k,k} + p_{k,k-1}x + \dots + p_{k,1}x^{k-1} + x^k \quad k = 0, 1, \dots, n, \quad (9.1)$$

таких, что

$$(P_i, P_j) = 0 \quad \text{при} \quad i \neq j. \quad (9.2)$$

Здесь индекс k введен для того, чтобы идентифицировать принадлежность коэффициента p_i к многочлену заданной степени.

Доказательство. Пусть $P_0(x) = 1$. Предположим, что система таких многочленов построена при некотором k и пусть

$$P_{k+1}(x) = x^{k+1} - \sum_{i=1}^k a_{ki} P_i(x). \quad (9.3)$$

Коэффициенты a_{ki} выберем из условия $(P_{k+1}, P_i) = 0 \quad i = 1, 2, \dots, k$. Отсюда

$$a_{ki} = \frac{(x^{k+1}, P_i)}{(P_i, P_i)}. \quad (9.4)$$

Следовательно, построенный многочлен будет ортогонален всем многочленам $P_i(x) \quad i=1, 2, \dots, k$. Далее, представление многочлена (9.1) означает, что

многочлен $P_k(x)$ может быть представлен в виде скалярного произведения двух векторов $(p_{k,k}, p_{k,k-1}, \dots, p_{k,1}, 1)$ и $(1, x, \dots, x^{k-1}, x^k)$, и переход от системы элементарных одночленов к системе ортогональных многочленов осуществляется преобразованием

$$\begin{pmatrix} P_0 \\ P_1 \\ P_2 \\ \dots \\ P_k \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ p_{1,1} & 1 & 0 & \dots & 0 \\ p_{2,2} & p_{2,1} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ p_{k,k} & p_{k,k-1} & p_{k,k-3} & \dots & 1 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ x^2 \\ \dots \\ x^k \end{pmatrix}$$

с матрицей B

$$B = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ p_{1,1} & 1 & 0 & \dots & 0 \\ p_{2,2} & p_{2,1} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ p_{k,k} & p_{k,k-1} & p_{k,k-3} & \dots & 1 \end{pmatrix} \quad (9.5)$$

Таким образом, переход от линейно независимой системы элементарных одночленов к системе ортогональных многочленов осуществляется невырожденным преобразованием, поэтому и система ортогональных многочленов линейно независима.

Теорема доказана.

Определение. Систему многочленов, удовлетворяющих условию (9.1), называют ортогональной системой.

Определение. Матрица B называется матрицей ортогонализации. Обратный переход осуществляется также невырожденной треугольной матрицей A

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ a_{1,1} & 1 & 0 & \dots & 0 \\ a_{2,2} & a_{2,1} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{k,k} & a_{k,k-1} & a_{k,k-3} & \dots & 1 \end{pmatrix}$$

Следствие 1. Ортогональный многочлен $P_k(x)$ ортогонален любым многочленам меньшей степени. Доказательство легко провести по индукции, что предлагается сделать самостоятельно в качестве легкого упражнения.

Следствие 2. Система ортогональных многочленов образует базис в пространстве многочленов.

Следующая теорема является основной в теории ортогональных функций.

Теорема 2. Для ортогональных многочленов справедливо равенство

$$P_{k+1}(x) + d_{k+1,k}P_k(x) + d_{k+1,k-1}P_{k-1}(x) - xP_k(x) = 0, \quad (9.6)$$

причем

$$d_{k+1,k-1} > 0$$

и

$$d_{k+1,k} = \frac{\int_a^b xp(x)P_k^2(x)dx}{\int_a^b p(x)P_k^2(x)dx}, \quad (9.7)$$

$$d_{k+1,k-1} = \frac{\int_a^b p(x)P_k^2(x)dx}{\int_a^b p(x)P_{k-1}^2(x)dx}. \quad (9.8)$$

Доказательство. Пусть построена система ортогональных многочленов $P_0(x), P_1(x), \dots, P_{k+1}(x)$. Рассмотрим многочлен $Q_{k+1}(x) = xP_k(x)$, где $P_k(x)$ — ортогональный многочлен степени k . $Q_{k+1}(x)$ уже не относится к системе ортогональных многочленов. Разложим $Q_{k+1}(x) = xP_k(x)$ по системе ортогональных многочленов.

$$xP_k(x) = d_{k+1,k+1}P_{k+1} + d_{k+1,k}P_k(x) + d_{k+1,k-1}P_{k-1}(x) + \dots + d_{k+1,0}P_0. \quad (9.9)$$

Поскольку у $xP_k(x)$ и $P_{k+1}(x)$ старшие коэффициенты одинаковы и равны единице, $d_{k+1,k+1} = 1$.

Коэффициенты $d_{k+1,m}$ этого разложения определяются по формулам

$$d_{k+1,m} = \frac{(xP_k, P_m)}{(P_m, P_m)} = \frac{\int_a^b xp(x)P_k(x)P_m(x)dx}{\int_a^b p(x)P_m^2(x)dx}. \quad (9.10)$$

С другой стороны, согласно следствию 1, ортогональные многочлены ортогональны всем многочленам меньшей степени. Поэтому

$$(xP_k, P_m) = (P_k, xP_m) = 0 \quad \text{при } k > m + 1.$$

Согласно следствию 1, $d_{k+1,m} = 0$ при $k > m + 1$ и в разложении (9.9) равны нулю коэффициенты $d_{k+1,k-2}, d_{k+1,k-3}, \dots, d_{k+1,0}$. Разложение

(9.9) при этом имеет вид (9.6). Докажем теперь, что $d_{k+1,k-1} > 0$. Согласно (9.10)

$$d_{k+1,k-1} = \frac{(xP_k, P_{k-1})}{(P_{k-1}, P_{k-1})}.$$

Знаменатель этой дроби > 0 . Покажем, что и числитель > 0 . Разлагая $xP_{k-1}(x)$ по системе ортогональных многочленов

$$xP_{k-1}(x) = P_k(x) + b_{k-1}P_{k-1}(x) + \dots + b_0P_0,$$

получим

$$(xP_k, P_{k-1}) = (P_k, xP_{k-1}) = (P_k, P_k) > 0.$$

Теорема доказана.

Теорема 3. Все корни ортогонального многочлена $P_k(x)$ различны и принадлежат открытому интервалу (a, b) .

Доказательство. Предположим противное, на (a, b) находятся m корней x_1, x_2, \dots, x_m многочлена $P_k(x)$ и $m < k$. Предположим сначала, что все корни различны, и рассмотрим многочлен

$$Q_m(x) = \prod_{i=1}^m (x - x_i).$$

Поскольку $p(x) > 0$, подынтегральное выражение в

$$\int_a^b p(x)P_k(x)Q_m(x)dx$$

не меняет знак на (a, b) и поэтому значение интеграла > 0 . С другой стороны, ортогональный многочлен ортогонален всем многочленам меньшей степени (следствие 1) и поэтому значение интеграла $= 0$. Получено противоречие. В случае кратного корня x_0 множитель $(x - x_0)$ в многочлене $Q(x)$ можно исключить, если корень четной кратности, и оставить его в первой степени, если он нечетной кратности.

Теорема доказана.

Теорема 4. Корни многочленов $P_{k+1}(x)$ и $P_k(x)$ разделяют друг друга. Иначе, пусть $x_1^k < x_2^k < \dots < x_k^k$ — корни многочлена $P_k(x)$ и $a < x_1^{k+1} < x_2^{k+1} < \dots < x_k^{k+1} < x_{k+1}^{k+1} < b$ — корни многочлена $P_{k+1}(x)$.

Тогда

$$a < x_1^{k+1} < x_1^k < x_2^{k+1} < x_2^k < \dots < x_k^k < x_{k+1}^{k+1} < b. \quad (9.11)$$

Доказательство. Докажем это утверждение по индукции. Докажем сначала, что для корней многочленов $P_2(x)$ и $P_1(x)$ имеет место цепочка неравенств $a < x_1^2 < x_1^1 < x_2^2 < b$. Рассмотрим рекуррентное соотношение (9.6) для многочленов $P_0(x) = 1$, $P_1(x)$ и $P_2(x)$ в точке, где $P_1(x_1^1) = 0$.

$$P_2(x_1^1) + d_{2,0}P_0(x_1^1) = 0.$$

Поскольку $d_{2,0} > 0$, то $P_2(x_1^1) < 0$. Но $P_2(x)$ — многочлен второй степени со старшим коэффициентом 1, и вне корней он принимает положительные значения. Поэтому должны существовать точка x_1^2 : $a < x_1^2 < x_1^1$ такая, что $P_2(x_1^2) = 0$, и точка x_2^2 : $x_1^1 < x_2^2 < b$, что $P_2(x_2^2) = 0$. Следовательно, для многочленов $P_1(x)$ и $P_2(x)$ утверждение теоремы имеет место.

Предположим теперь, что утверждение теоремы имеет место для корней многочленов $P_{k-1}(x)$ и $P_k(x)$.

$$a < x_1^k < x_1^{k-1} < x_2^k < x_2^{k-1} < \dots < x_{k-1}^{k-1} < x_k^k < b.$$

Докажем, что оно имеет место для корней многочленов $P_k(x)$ и $P_{k+1}(x)$.

Заметим, что многочлены $P_{k-1}(x)$ и $P_{k+1}(x)$ имеют одинаковый знак вне корней, так как старший член у этих многочленов равен соответственно x^{k-1} и x^{k+1} , т. е. x одновременно или в четной, или в нечетной степени. Рассмотрим (9.6) в точках $P_k(x_j^k) = 0$ $j = 1, 2, \dots, k$. В этих точках значения $P_{k+1}(x_j^k)$ и $P_{k-1}(x_j^k)$ имеют разные знаки. Согласно индуктивному предположению, между двумя последовательными корнями $P_k(x)$ есть один корень $P_{k-1}(x)$, и, следовательно, в точках x_j^k и x_{j+1}^k значения $P_{k-1}(x)$ имеют разные знаки. Но тогда и $P_{k+1}(x)$ будет иметь разные знаки в точках x_j^k и x_{j+1}^k . Тем самым мы заметили, что между двумя последовательными корнями многочлена $P_k(x)$ есть корень многочлена $P_{k+1}(x)$. Итак, на $k-1$ интервале (x_j^k, x_{j+1}^k) $j = 1, 2, \dots, k$ есть, по крайней мере, один корень $P_{k+1}(x)$. Рассмотрим теперь интервал $[x_k^k, b]$. В точке x_k^k значения P_{k-1} и $P_{k+1}(x)$ имеют разные знаки, поэтому $P_{k+1}(x)$ на (x_k^k, b) должен сменить знак, чтобы при $x > b$ у многочленов $P_{k-1}(x)$ и $P_{k+1}(x)$ были одинаковые знаки. Значит, на этом интервале также должен быть корень $P_{k+1}(x)$. Такая же ситуация на интервале (a, x_1^k) . Итак, найден $k+1$ интервал, где есть корни $P_{k+1}(x)$. Одновременно замечаем, что там может находиться только один корень $P_{k+1}(x)$ и корни многочленов $P_k(x)$ и $P_{k+1}(x)$ удовлетворяют условию (9.11). Теорема доказана.

Теорема 5. Пусть весовая функция $p(x)$ является четной функцией относительно середины отрезка. Тогда многочлены $P_k(x)$ являются четными функциями относительно середины отрезка при четном k и нечетными функциями относительно середины отрезка при нечетном k .

Доказательство. Замена

$$t = \frac{2x - (b + a)}{b - a}$$

переводит отрезок $[a, b]$ в отрезок $[-1, 1]$, поэтому достаточно доказать утверждение теоремы для отрезка $[-1, 1]$. Имеем $P_0(x) = 1$ и в (9.4)

$$a_{10} = \frac{\int_{-1}^1 p(x)x dx}{2} = 0,$$

поэтому $P_1(x) = x$ — нечетная функция. Предположим теперь, что $P_{k-1}(x)$ — нечетная функция, а $P_k(x)$ — четная функция. Тогда в (9.6) $d_{k+1,k} = 0$, так как (9.7) в числителе стоит интеграл от нечетной функции на симметричном относительно нуля интервале. Поэтому

$$P_{k+1}(x) = xP_k(x) - d_{k+1,k-1}P_{k-1}(x)$$

$$\begin{aligned} P_{k+1}(-x) &= -xP_k(-x) - d_{k+1,k-1}P_{k-1}(-x) = \\ &= -xP_k(x) + d_{k+1,k-1}P_{k-1}(x) = -P_{k+1}(x). \end{aligned}$$

Другой случай рассматривается аналогично.

Ортогональные многочлены являются важным инструментом в теории интерполяции и смежных задачах. Изучая вопросы, связанные с оптимальным выбором узлов интерполяции, мы ввели в рассмотрение многочлены Чебышева. Нетрудно заметить, что многочлен $\bar{T}_m(x)$ является ортогональным многочленом на $[-1, 1]$ с весовой функцией

$$p(x) = \frac{1}{\sqrt{1-x^2}}.$$

При построении квадратурных формул мы убедимся, что наиболее точные формулы для приближенного вычисления интеграла получаются при использовании в качестве узлов квадратур нулей ортогональных многочленов. Приведем наиболее употребительные в вычислительной практике ортогональные многочлены.

Многочлены Якоби. Эти многочлены получаются при построении системы ортогональных многочленов на $[-1, 1]$ с весовой функцией

$$p(x) = (1-x)^\alpha(1+x)^\beta \quad \alpha, \beta > -1.$$

Многочлены Якоби имеют вид

$$P_n(x) = \frac{(-1)^n}{2^n n!} (1-x)^{-\alpha} (1+x)^{-\beta} \frac{d^n}{dx^n} [(1-x)^{-\alpha+n} (1+x)^{-\beta+n}].$$

Частным случаем при $p(x)=1$ являются многочлены Лежандра, с которыми мы встретимся при построении формул приближенного вычисления интегралов.

§ 10. Численное дифференцирование

Задача приближенного вычисления производных от функции, заданной таблично, обычно решается вычислением производных от интерполяционного многочлена. Если известны значения функции в точках x_1, x_2, \dots, x_n , то

$$f^{(k)}(x) \approx L_n^{(k)}(x). \quad (10.1)$$

Оценим погрешность формулы (10.1). Из (7.3) имеем равенство

$$f(x) - L_n(x) = f(x_1; x_2; \dots; x_n; x) \omega_n(x).$$

Отсюда

$$f^{(k)}(x) - L_n^{(k)}(x) = \sum_{i=0}^k C_k^i (f(x_1; \dots; x_n; x))^{(i)} \omega_n^{(k-i)}(x).$$

Выясним смысл выражения

$$(f(x_1; \dots; x_n; x))^{(i)}.$$

В § 8 получена формула

$$\underbrace{f(x_i; x_i; \dots; x_i)}_{m \text{ раз}} = \frac{f^{(m-1)}(x_i)}{(m-1)!},$$

из которой следует, что

$$(f(x_1; \dots; x_n; x))^{(i)} = i! f(x_1; \dots; x_n; \underbrace{x; \dots; x}_{i+1 \text{ раз}}).$$

Поэтому

$$f^{(k)}(x) - L_n^{(k)}(x) = \sum_{i=0}^k C_k^i f(x_1; \dots; x_n; \underbrace{x; \dots; x}_{i+1 \text{ раз}}) \omega_n^{(k-i)}(x).$$

В практических вычислениях обычно используются более простые соображения. Так, если в точках x_1 и x_2 заданы $f(x_1)$ и $f(x_2)$, то

$$L_2(x) = f(x_1) \frac{x - x_2}{x_1 - x_2} + f(x_2) \frac{x - x_1}{x_2 - x_1}$$

и

$$f'(x) \approx \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

При этом

$$|f'(x) - L'_2(x)| = \frac{f''(\xi)}{2} \omega'(x)|.$$

Если $h = |x_2 - x_1|$, то

$$\|f' - L'_2\| = f''(\xi) \|h\|.$$

Более подробно с разностными аналогами производных и соответствующими погрешностями мы познакомимся при изучении сеточных уравнений и разностных методов решения краевых задач.

§ 11. Численное интегрирование. Квадратурные формулы Ньютона — Котеса

В этом разделе мы получим формулы для приближенного вычисления интеграла

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

Функция $p(x)$ называется весовой, и предполагается, что $p(x) > 0$. Квадратурные формулы бывают простые и составные. Методика получения простых квадратурных формул состоит в том, что функция $f(x)$ на заданном отрезке под интегралом заменяется на ее интерполяционный многочлен, что позволяет получить достаточно простые формулы приближенного вычисления интеграла и оценить погрешность. Составные формулы получаются, когда весь отрезок интегрирования разбивается на отрезки длины h , и на каждом отрезке используются простые квадратурные формулы. Займемся выводом простых квадратурных формул. Все построения мы будем проводить на отрезке $[0, h]$. Поскольку в некоторых случаях мы будем исследовать дополнительные свойства квадратурных формул в случае симметрии функции $p(x)$ и симметричного расположения узлов, удобно ввести отображение отрезка $[-1, 1]$ на $[0, h]$.

$$x = \frac{h}{2}(t + 1). \quad (11.1)$$

Изложим общую схему построения этих формул. Рассматриваемые ниже формулы приближенного вычисления интеграла называются квадратурными формулами Ньютона — Котеса. Все квадратурные формулы имеют вид

$$\int_0^h f(x)dx \approx \sum_{i=1}^n C_i f(x_i).$$

Зададим некоторые узлы $x_i = \frac{h}{2}(t_i + 1)$ $i = 1, 2, \dots, n$ и построим интерполяционный многочлен

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

Будем считать, что

$$I(f) = \int_0^h f(x)dx \approx \int_0^h L_n(x)dx = S_n(f). \quad (11.2)$$

Здесь

$$S_n(f) = \sum_{i=1}^n C_i f(x_i),$$

а коэффициенты C_I вычисляются по формуле

$$C_i = \int_0^h p(x) \frac{\prod_{\substack{j=1 \\ j \neq i}}^n (x - x_j)}{\prod_{\substack{j=1 \\ j \neq i}}^n (x_i - x_j)} dx \quad (11.3)$$

или, сделав замену (11.1),

$$C_i = \frac{h}{2} \int_{-1}^1 p^0(t) \frac{\prod_{\substack{j=1 \\ j \neq i}}^n (t - t_j)}{\prod_{\substack{j=1 \\ j \neq i}}^n (t_i - t_j)} dt = \frac{h}{2} D_i,$$

где

$$D_i = \int_{-1}^1 p^0(t) \frac{\prod_{\substack{j=1 \\ j \neq i}}^n (t - t_j)}{\prod_{\substack{j=1 \\ j \neq i}}^n (t_i - t_j)} dt. \quad (11.3')$$

Оценим погрешность построенной квадратурной формулы. В силу (5.4)

$$|R_n(f)| \leq \frac{\max |f^{(n)}|(x)}{n!} \int_0^h |p(x)| |\omega_n(x)| dx. \quad (11.4)$$

После замены (11.1)

$$|R_n(f)| \leq \frac{h^{n+1} \max |f^{(n)}(x)|}{2^{n+1}n!} \int_{-1}^1 |p^0(t)| |\omega_n(t)| dt. \quad (11.4')$$

Следующая теорема иллюстрирует свойства квадратурных формул в случае симметричного расположения узлов.

Теорема 1. Пусть весовая функция $p(x)$ — четная относительно середины отрезка $[0, h]$, а узлы расположены симметрично относительно середины отрезка. Тогда коэффициенты C_i , соответствующие симметричным узлам, равны.

Доказательство достаточно проделать для отрезка $[-1, 1]$. Пусть узлы на $[-1, 1]$ расположены симметрично: $t_i = t_{n-i+1}$. Рассмотрим

$$D_i = \int_{-1}^1 p^0(t) \frac{\prod_{\substack{j=1 \\ j \neq i}}^n (t - t_j)}{\prod_{\substack{j=1 \\ j \neq i}}^n (t_i - t_j)} dt = \int_{-1}^1 p^0(t) \frac{\omega_n(t)}{(t - t_i) \omega'_n(t_i)} dt,$$

где $\omega_n(t) = (t - t_1)(t - t_2) \dots (t - t_n)$. В силу симметричного расположения узлов функция $\omega_n(t)$ есть четная функция при четном n и нечетная функция при нечетном n . Далее, производная от четной функции есть нечетная функция, а от нечетной — четная. Поэтому $\omega'_n(t_i) = -\omega'_n(t_{n-i+1})$ при четном n и $\omega'_n(t_i) = \omega'_n(t_{n-i+1})$ при нечетном n . Сделаем замену переменной интегрирования $t = -\tau$ и, заменив t_i на $-t_{n-i+1}$, получим

$$D_i = \int_{-1}^1 p^0(t) \frac{\omega_n(t)}{(t - t_i) \omega'_n(t_i)} dt = \int_{-1}^1 p^0(-\tau) \frac{\omega_n(-\tau)}{(-\tau + t_{n-i+1}) \omega'_n(-t_{n-i+1})} d\tau.$$

В обоих случаях получаем

$$D_i = \int_{-1}^1 p^0(\tau) \frac{\omega_n(\tau)}{(\tau - t_{n-i+1}) \omega'_n(t_{n-i+1})} d\tau = D_{n-i+1}.$$

Теорема доказана.

Будем говорить, что квадратурная формула точна для некоторого класса функций, если с ее помощью вычисляется точное значение интеграла.

Квадратурные формулы Ньютона — Котеса точны для многочленов степени $n - 1$. В самом деле, интерполяционный многочлен, построенный по n узлам для функции, которая сама есть многочлен степени $m < n$, совпадает с этим многочленом в силу теоремы единственности для интерполяционного многочлена.

Симметричное расположение узлов позволяет не только уменьшить работу по вычислению коэффициентов квадратурных формул, но в ряде случаев увеличить точность квадратурной формулы. В самом деле, пусть $p(x)$ — четная функция и число узлов n нечетно. Тогда такая квадратурная формула точна уже для многочленов степени n . Для доказательства этого утверждения достаточно заметить, что для элементарного одночлена x^n

$$\int_{-1}^1 p(x)x^n dx = 0$$

в силу нечетности подынтегральной функции. С другой стороны,

$$\sum_{i=1}^n C_i x_i^n = 0$$

в силу нечетности n и симметричного расположения узлов.

Займемся выводом наиболее употребительных формул при $p(x)=1$. Сразу заметим, что константа является четной функцией относительно любого интервала.

Формула прямоугольников. Различают формулу прямоугольников с левосторонним, правосторонним и центральным узлом. Выведем формулу прямоугольников с центральным узлом. В этом случае $t_1 = 0$, $x_1 = \frac{h}{2}$ и

$$C_1 = \frac{h}{2} \int_{-1}^1 1 dt = h$$

и получаем

$$\int_1^h f(x) \approx hf\left(\frac{h}{2}\right).$$

Оценка погрешности согласно (11.4)

$$R_1 = \max_{[0,h]} f'(x) \frac{h^2}{4}.$$

Интересно отметить, что здесь мы сталкиваемся с описанной выше ситуацией: число узлов нечетно и весовая функция четна. Это обстоятельство позволяет предположить, что на самом деле оценка погрешности может быть лучше. Действительно, построим квадратурную формулу прямоугольников с кратным центральным узлом. Интерполяционный многочлен в этом случае имеет вид

$$L_2(x) = f\left(\frac{h}{2}\right) + f'\left(\frac{h}{2}\right)\left(x - \frac{h}{2}\right).$$

Вычисляя коэффициенты C_1 и C_2 , замечаем, что $C_1 = h$, а $C_2 = 0$. Таким образом, получена та же квадратурная формула

$$\int_1^h f(x) \approx hf\left(\frac{h}{2}\right),$$

но с оценкой погрешности

$$R_2 = \frac{h^3 \max |f''(x)|}{16} \int_{-1}^1 |t^2| dt = \frac{h^3 \max |f''(x)|}{24}.$$

Составная формула прямоугольников получается, когда весь отрезок интегрирования $[a, b]$ разбивается точками x_i на отрезки длиной h , с последующим применением полученной формулы на каждом интервале. Составная формула прямоугольников имеет вид:

$$S_n(f) = h \sum_{i=0}^{n-1} f\left(x_i + \frac{h}{2}\right). \quad (11.5)$$

Здесь

$$h = \frac{b-a}{n}, \quad x_i = a + ih, \quad i = 0, 1, \dots, n-1.$$

Поскольку $nh = b - a$, погрешность составной формулы прямоугольников

$$R \leq \frac{h^2 \max |f''(x)|}{24}.$$

Формула трапеций. Аппроксимируя функцию $f(x)$ многочленом $L_2(x)$ с узлами $x_1 = 0$ и $x_2 = h$ ($t_1 = -1$ и $t_2 = 1$), получаем

$$C_1 = \frac{h}{2} \int_{-1}^1 \frac{1-t}{2} dt = \frac{h}{2}, \quad C_2 = C_1.$$

В результате приходим к формуле

$$\int_0^h f(x) dx \approx \frac{h}{2}(f(0) + f(h))$$

с оценкой погрешности

$$R_2 \leq \frac{h^3 \max |f''(x)|}{16} \int_{-1}^1 |t^2 - 1| dt = \frac{h^3 \max |f''(x)|}{12}.$$

Составная формула трапеций имеет вид

$$S_n(f) = h \left(\frac{f(a) + f(b)}{2} + \sum_{i=1}^{N-1} f(x_i) \right) \quad x_i = a + ih \quad h = \frac{b-a}{n} \quad (11.6)$$

с оценкой погрешности

$$R \leq \frac{h^2 \max |f''(x)|}{12}.$$

Формула Симпсона. Так же, как и в случае формулы прямоугольников с центральным узлом, можно заметить, что интерполяционный многочлен с тремя узлами $x_1 = 0$, $x_2 = \frac{h}{2}$, $x_3 = h$ ($t_1 = -1$, $t_2 = 0$, $t_3 = 1$) и интерполяционный многочлен с четырьмя узлами $x_1 = 0$, $x_2 = x_3 = \frac{h}{2}$, $x_4 = h$ ($t_1 = -1$, $t_2 = t_3 = 0$, $t_4 = 1$) дает одну и ту же квадратурную формулу. В самом деле,

$$L_4(x) = L_3(x) + f(0; \frac{h}{2}; \frac{h}{2}; h) x (x - \frac{h}{2}) (x - h),$$

где $L_3(x)$ — интерполяционный многочлен, построенный по трем точкам $x_1 = 0$, $x_2 = \frac{h}{2}$, $x_3 = h$. Но

$$\int_0^h x(x - \frac{h}{2}) (x - h) dx = 0,$$

поскольку под интегралом нечетная относительно середины отрезка функция. Поэтому квадратурная формула Симпсона с тремя узлами такая же, как и формула с четырьмя узлами, когда средний узел двукратный.

Вычислим коэффициенты квадратурной формулы Симпсона, соответствующие узлам $t_1 = -1$, $t_2 = 0$, $t_3 = 1$:

$$C_3 = C_1 = \frac{h}{2} \int_{-1}^1 \frac{t(t+1)}{2} dt = \frac{h}{6},$$

$$C_2 = \frac{h}{2} \int_{-1}^1 (1-t^2) dt = \frac{4h}{6}.$$

Получена квадратурная формула

$$\int_0^h f(x) dx \approx \frac{h}{6} [f(0) + 4f(\frac{h}{2}) + f(h)].$$

Оценим погрешность квадратурной формулы Симпсона. Согласно (11.4')

$$R_4 \leq \frac{h^5 \max |f^{(4)}(x)|}{2^5 5!} \int_{-1}^1 |t^2(t^2-1)| dt = \frac{h^5 \max |f^{(4)}(x)|}{2880}.$$

Составная формула Симпсона имеет вид

$$S_N(f) =$$

$$= \frac{h}{3} [f(a) + f(b) + 2f(x_1) + 4f(x_2) + 2f(x_3) + 4f(x_4) + \dots + 2f(x_{N-1})] \quad (11.7)$$

Здесь N чётно, $x_i = a + ih$, $h = \frac{b-a}{n}$. Погрешность составной формулы трапеций

$$R \leq \frac{h^4 \max |f^{(4)}(x)|}{2880}.$$

§ 12. Квадратурные формулы Гаусса

Эффект увеличения точности квадратурных формул при специальном выборе узлов позволяет предположить, что возможно построение квадратурных формул, точных для многочленов степени $2N-1$. Это предположение обосновывается следующими соображениями. В этом случае квадратурная формула должна давать точное значение при вычислении интегралов от элементарных одночленов $1, x, x^2, \dots, x^{2N-1}$. Следовательно,

$$\int_a^b p(x)dx = C_1 + C_2 + \dots + C_N$$

$$\int_a^b p(x) \, x dx = C_1 x_1 + C_2 x_2 + \dots + C_N x_N$$

$$\int_a^b p(x) x^2 dx = C_1 x_1^2 + C_2 x_2^2 + \dots + C_N x_N^2$$

• • • • •

$$\int_a^b p(x) x^{2N-1} dx = C_1 x_1^{2N-1} + C_2 x_2^{2N-1} + \dots + C_N x_N^{2N-1}$$

Получена нелинейная система из $2N$ уравнений относительно $2N$ неизвестных $C_1, C_2, \dots, C_N, x_1, x_2, \dots, x_N$. Если эта система имеет решение и узлы x_i принадлежат отрезку $[a, b]$, то такая квадратурная формула может быть построена.

Предположим, что такая квадратурная формула существует.

Теорема 1. Пусть квадратурная формула

$$S_N(f) = C_1 f(x_1) + C_2 f(x_2) + \dots + C_N f(x_N)$$

точна для многочленов до $N-1$ степени и пусть

$$\omega_N(x) = (x - x_1)(x - x_2) \dots (x - x_N)$$

— многочлен, корнями которого являются узлы этой квадратурной формулы. Тогда для любого многочлена $P_{N-1}(x)$ степени не выше $N-1$ имеет место

$$\int_a^b p(x)\omega_N(x)P_{N-1}(x)dx = 0.$$

Доказательство. Рассмотрим многочлен

$$Q_{2N-1}(x) = \omega_N(x)P_{N-1}(x),$$

где $P_{N-1}(x)$ — многочлен степени не выше $N-1$. По предположению теоремы квадратурная формула точна для многочленов до $2N-1$ степени. Следовательно,

$$\int_a^b p(x)\omega_N(x)P_{N-1}(x)dx = C_1\omega_N(x_1)P_{N-1}(x_1) + \dots + C_N\omega_N(x_N)P_{N-1}(x_N).$$

Но $\omega(x_j) = 0$ при $j = 1, 2, \dots, N$, что и доказывает теорему.

Эта теорема показывает, что узлы квадратурной формулы, точной для многочленов до $2N - 1$ степени, являются корнями ортогонального многочлена, которые рассмотрены в § 9.

Теорема 2. Пусть

$$S_N(f) = C_1f(x_1) + C_2f(x_2) + \dots + C_Nf(x_N)$$

— квадратурная формула Ньютона — Котеса, узлами которой являются корни ортогонального многочлена $P_N(x)$. Тогда эта квадратурная формула точна для многочленов степени $2N - 1$.

Доказательство. Квадратурная формула Ньютона — Котеса точна для многочленов до $N-1$ степени. Пусть x_i $i = 1, 2, \dots, N$ — корни ортогонального многочлена, которые являются узлами квадратурной формулы, и пусть

$$\omega_N = (x - x_1)(x - x_2) \dots (x - x_N).$$

Поскольку старший коэффициент у многочлена ω_N равен единице, то многочлен ω_N и есть ортогональный многочлен $P_N(x)$. Разделим произвольный многочлен $Q_{2N-1}(x)$ на $\omega_N(x)$

$$\frac{Q_{2N-1}(x)}{\omega_N(x)} = T_{N-1}(x) + \frac{R_{N-1}(x)}{\omega_N(x)}.$$

Здесь $T_{N-1}(x)$ и $R_{N-1}(x)$ — многочлены степени не выше $N-1$. Поэтому $Q_{2N-1}(x)$ можно представить в виде

$$Q_{2N-1}(x) = \omega_N(x)T_{N-1}(x) + R_{N-1}(x).$$

Покажем, что квадратурная формула точна для многочлена $Q_{2N-1}(x)$. В силу того, что $\omega_N(x) = P_N(x)$

$$\int_a^b p(x)\omega_N(x)T_{N-1}(x)dx = 0$$

и

$$\int_a^b p(x)Q_{2N-1}(x)dx = \int_a^b p(x)R_{N-1}(x)dx.$$

Но квадратурная формула $S_N(f)$ точна для многочленов степени $N-1$, и поэтому она точна для многочленов степени до $2N-1$. Теорема доказана.

Таким образом, квадратурные формулы Гаусса — это квадратуры Ньютона — Котеса, узлами которых являются корни ортогонального многочлена.

Из других свойств квадратур Гаусса отметим положительность коэффициентов C_i . В самом деле, рассмотрим многочлен

$$T_{2N-2}(x) = \prod_{\substack{i=1 \\ i \neq k}}^N (x - x_i)^2.$$

Этот многочлен степени $2N-2$ равен нулю в точках x_i , $i \neq k$. Квадратурная формула Гаусса точна для таких многочленов. Поэтому

$$\int_a^b p(x)T_{2N-2}(x)dx = C_k T_{N-2}(x_k)$$

и положительность C_k при $b > a$ очевидна.

Построим в качестве примера квадратурную формулу Гаусса с тремя узлами для вычисления интеграла

$$\int_{-1}^1 f(x)dx \approx c_1 f(x_1) + c_2 f(x_2) + c_3 f(x_3).$$

Эта квадратура должна вычислять точное значение интеграла в случае, если

$$f(x) = 1, \quad f(x) = x, \quad f(x) = x^2, \quad f(x) = x^3, \quad f(x) = x^4, \quad f(x) = x^5.$$

Подставляя x^m , $m = 0, 1, 2, 3, 4, 5$ в квадратурную формулу, получаем

$$c_1 + c_2 + c_3 = 2$$

$$c_1 x_1 + c_2 x_2 + c_3 x_3 = 0$$

$$c_1 x_1^2 + c_2 x_2^2 + c_3 x_3^2 = \frac{2}{3}$$

$$c_1 x_1^3 + c_2 x_2^3 + c_3 x_3^3 = 0$$

$$c_1 x_1^4 + c_2 x_2^4 + c_3 x_3^4 = \frac{2}{5}$$

$$c_1 x_1^5 + c_2 x_2^5 + c_3 x_3^5 = 0$$

Поскольку $p(x)=1$ — четная функция, узлы расположены симметрично (теорема 5 § 9), а коэффициенты, соответствующие симметричным узлам, равны (теорема 1 § 11). Второе, четвертое и шестое уравнения — следствие этого свойства. По этому же свойству

$x_2 = 0$ и $x_1 = -x_2$ и получаем

$$2c_1 + c_2 = 2, \quad 2c_1 x_1^2 = \frac{2}{3}, \quad 2c_1 x_1^4 = \frac{2}{5}.$$

Решив эту простую систему, получим

$$x_1 = -\sqrt{\frac{3}{5}} \quad x_2 = 0 \quad x_3 = \sqrt{\frac{3}{5}} \quad c_1 = \frac{5}{9} \quad c_2 = \frac{8}{9} \quad c_3 = \frac{5}{9}.$$

Так как отрезок $[-1,1]$ отображением

$$x = \frac{h}{2}t + \frac{h}{2}$$

переводится в отрезок $[0,h]$, то для интервала интегрирования $[0,h]$ в квадратурной формуле

$$\int_0^h f(x)dx \approx d_1 f(x_1) + d_2 f(x_2) + d_3 f(x_3)$$

для узлов и коэффициентов этой квадратуры получаем

$$x_1 = -\frac{h}{2} \quad x_2 = \frac{h}{2} \quad x_3 = \frac{h}{2} \quad d_1 = \frac{5}{9} \quad d_2 = \frac{8}{9} \quad d_3 = \frac{5}{9}$$

Составная формула Гаусса с тремя узлами, которая является аналогом формулы Симпсона, имеет вид

$$\int_a^b f(x)dx \approx \sum_{i=0}^{n-1} [d_1 f(y_i + x_1) + d_2 f(y_i + x_2) + d_3 f(y_i + x_3)]. \quad (11.8)$$

Здесь

$$y_i = a + ih, \quad h = \frac{b-a}{n}.$$

С оценками погрешности квадратурных формул Гаусса можно ознакомиться, например, в [2].

§ 13. Наилучшее приближение в гильбертовом пространстве

Рассмотрим следующую задачу. Пусть в гильбертовом пространстве H задана линейно независимая система g_1, g_2, \dots, g_n . Требуется заданный элемент f приблизить линейной комбинацией элементов g_1, g_2, \dots, g_n так, чтобы

$$\Delta = \left\| f - \sum_{i=1}^n c_i^0 g_i \right\| = \inf_{c_1, \dots, c_n} \left\| f - \sum_{i=1}^n c_i g_i \right\|.$$

Элемент

$$h = \sum_{i=1}^n c_i^0 g_i$$

называется элементом наилучшего приближения, а величина Δ называется величиной наилучшего приближения. Далее мы будем предполагать, что f не принадлежит линейной оболочке g_1, g_2, \dots, g_n , то есть $\Delta > 0$. Задача отыскания такого элемента решается достаточно просто, если в пространстве задано скалярное произведение.

Теорема 1. Вещественное гильбертово пространство строго нормировано.

Доказательство. Напомним, что пространство называется строго нормированным, если равенство

$$\|x + y\| = \|x\| + \|y\| \quad (13.1)$$

выполняется тогда и только тогда, когда $y = \lambda x$ $\lambda > 0$. В одну сторону, если $y = \lambda x$, $\lambda > 0$ неравенство треугольника очевидно превращается в равенство. Пусть теперь неравенство треугольника превратилось в равенство. Докажем, что существует такое $\lambda > 0$, что $y = \lambda x$. Заметим, что из $\|x + y\| = \|x\| + \|y\|$ следует, что

$$(x, y) = \|x\| \|y\|. \quad (13.2)$$

В самом деле, возводя (13.1) в квадрат, получаем

$$\|x + y\|^2 = (x + y, x + y) = \|x\|^2 + \|y\|^2 + 2(x, y) = \|x\|^2 + \|y\|^2 + 2\|x\| \|y\|.$$

В случае строгого неравенства треугольника

$$(x, y) < \|x\| \|y\|.$$

Рассмотрим $\|y - \lambda x\|$:

$$\|y - \lambda x\|^2 = (y - \lambda x, y - \lambda x) = \lambda^2 \|x\|^2 - 2\lambda(x, y) + \|y\|^2.$$

Последнее выражение есть многочлен второй степени относительно λ . Найдем его корни:

$$\lambda^2 \|x\|^2 - 2\lambda(x, y) + \|y\|^2 = 0.$$

В силу (13.2) дискриминант этого уравнения равен нулю и

$$\lambda = \frac{\|y\|}{\|x\|} > 0.$$

При выполнении строгого неравенства треугольника дискриминант отрицательный и уравнение вещественных корней не имеет. Таким образом показано, что при выполнении (13.1) и только в этом случае существует $\lambda > 0$, что $\|y - \lambda x\| = 0$. Теорема доказана. Эта теорема справедлива и для комплексного гильбертова пространства. Доказательство проводится по той же схеме, что предлагается проделать самостоятельно.

Теорема 2. В гильбертовом пространстве элемент наилучшего приближения единственен.

Доказательство. Предположим противное: существуют два элемента наилучшего приближения h_1 и h_2 , такие, что $\|f - h_1\| = \Delta$ и $\|f - h_2\| = \Delta$. Тогда

$$\left\| f - \frac{h_1 + h_2}{2} \right\| = \left\| \frac{f - h_1}{2} + \frac{f - h_2}{2} \right\| \leq \frac{\|f - h_1\|}{2} + \frac{\|f - h_2\|}{2} = \Delta.$$

Но

$$\left\| f - \frac{h_1 + h_2}{2} \right\| \geq \Delta$$

по определению величины наилучшего приближения. Следовательно,

$$\left\| \frac{f - h_1}{2} + \frac{f - h_2}{2} \right\| = \frac{\|f - h_1\|}{2} + \frac{\|f - h_2\|}{2}.$$

Поскольку гильбертово пространство строго нормированно,

$$\frac{f - h_1}{2} = \lambda \frac{f - h_2}{2}.$$

Если $\lambda = 1$, то $h_1 = h_2$. Если же $\lambda \neq 1$, то

$$f = \frac{h_1 - \lambda h_2}{1 - \lambda}.$$

Но h_1 и h_2 есть линейная комбинация элементов g_1, g_2, \dots, g_n , и тогда и f есть линейная комбинация этих элементов, т. е. принадлежит их линейной оболочке. Получено противоречие. Теорема доказана.

Следующие две теоремы дают практический аппарат построения наилучшего приближения в гильбертовом пространстве.

Теорема 3. Пусть в гильбертовом пространстве H существует элемент h_0 , реализующий наилучшее приближение к элементу f . Тогда разность $f - h_0$ ортогональна всем элементам гильбертова пространства H .

Доказательство. Предположим противное, что в H есть элемент x , такой, что

$$(f - h_0, x) = a \neq 0.$$

Без ограничения общности можно считать, что $\|x\| = 1$. Пусть $y = h_0 - ax$. Тогда

$$\begin{aligned} \|f - y\|^2 &= \\ (f - h_0 - ax, f - h_0 - ax) &= \|f - h_0\|^2 + a\bar{a}\|x\|^2 - a(x, f - h_0) - \bar{a}(f - h_0, x). \end{aligned}$$

Но $(f - h_0, x) = a$ и $\|x\| = 1$. Поэтому

$$\|f - y\|^2 = \|f - h_0\|^2 + a\bar{a} - a\bar{a} - \bar{a}a.$$

В итоге получаем

$$\|f - y\|^2 < \|f - h_0\|^2.$$

Последнее неравенство означает, что h_0 не есть элемент наилучшего приближения. Это противоречие и доказывает теорему.

Теорема 4. Пусть $(f - h_0, x) = 0$ для любого $x \in H$. Тогда x есть элемент наилучшего приближения.

Доказательство. Для произвольного $x \in H$ имеем

$$\begin{aligned} \|f - x\|^2 &= (f - h_0 + h_0 - x, f - h_0 + h_0 - x) = (f - h_0, f - h_0) + (f - h_0, h_0 - x) + \\ &+ (h_0 - x, f - h_0) + (h_0 - x, h_0 - x). \end{aligned}$$

Но $f - h_0$ ортогонально всем элементам из H , в том числе и $h_0 - x$. Поэтому второе и третье слагаемые обращаются в нуль. В итоге имеем

$$\|f - x\|^2 = \|f - h_0\|^2 + \|h_0 - x\|^2 > \|f - h_0\|^2.$$

Это неравенство и доказывает теорему.

Эти две теоремы дают практический алгоритм нахождения наилучшего приближения для заданного элемента f . В самом деле, пусть

$$h_0 = \sum_{i=1}^n c_i g_i.$$

Умножая скалярно выражение

$$f = \sum_{i=1}^n c_i g_i$$

на g_i $i = 1, 2, \dots, n$, получаем систему для определения коэффициентов c_i .

$$c_1(q_1, q_1) + c_2(q_2, q_1) + \dots + c_n(q_n, q_1) = (f, q_1)$$

$$c_1(q_1, q_2) + c_2(q_2, q_2) + \dots + c_n(q_n, q_2) = (f, q_2)$$

• • • • •

$$c_1(q_1, q_n) + c_2(q_2, q_n) + \dots + c_n(q_n, q_n) = (f, q_n).$$

Матрица этой системы есть матрица Грамма. Если система g_1, g_2, \dots, g_n линейно независима, то ее определитель отличен от нуля. Поэтому полученная система разрешима единственным образом.

Рассмотрим пример. Пусть в пространстве $L_{2,(0,4)}$ со скалярным произведением

$$(f, g) = \int_0^4 f(x)g(x)dx$$

для функции $f(x) = |x - 1|$ нужно найти наилучшее приближение многочленом первой степени. Для элемента наилучшего приближения $h_0(x) = a_0 + a_1x$ получаем систему для определения коэффициентов a_0, a_1 .

$$a_0(g_1, g_1) + a_1(g_2, g_1) = (|x - 1|, g_1)$$

$$a_0(g_1, g_2) + a_1(g_2, g_2) = (|x - 1|, g_2).$$

Здесь $g_k = x^k$, $k = 0, 1$, а

$$(g_i, g_j) = \int_0^4 x^i x^j dx \quad (|x-1, g_i) = \int_0^4 |x-1| x^i dx.$$

Вычисляя интегралы, получаем систему для определения коэффициентов многочлена наилучшего приближения.

$$4a_0 + 8a_1 = 5,$$

$$8a_0 + \frac{64}{3}a_1 = 13\frac{2}{3}.$$

Отсюда

$$a_0 = -\frac{1}{8}, \quad a_1 = \frac{11}{16},$$

и многочлен имеет вид

$$h(x) = \frac{11}{16}x - \frac{1}{8}.$$

§ 14. Наилучшее равномерное приближение

Рассмотрим пространство C вещественных непрерывных функций, заданных на отрезке $[a, b]$, с нормой

$$\|f\| = \max_{[a,b]} |f(x)|$$

Будем искать наилучшее приближение заданной функции $f(x)$ многочленами степени n . Далее будем предполагать, что $f(x)$ не есть многочлен степени n .

В пространстве C норма элемента определяется не через скалярное произведение, что значительно осложняет задачу построения многочлена наилучшего приближения.

Итак, для заданной функции $f(x)$ будем искать многочлен наилучшего приближения

$$Q_n(x) = \sum_{k=0}^n q_k x^k,$$

такой, что для любых многочленов $P_n(x)$ степени n имеет место

$$E_n(f) = \|f - Q_n\| \leq \|f - P_n\|.$$

Многочлен $Q_n(x)$ будем называть многочленом наилучшего равномерного приближения, а величину $E_n(f)$ — величиной наилучшего приближения.

Теорема 1 (Валле - Пуссен). Пусть существуют $n+2$ точки $a \leq x_0 < x_1 < \dots < x_{n+1} \leq b$, такие, что величина $f(x) - P_n(x)$ последовательно меняет знак при переходе от точки x_i к x_{i+1} , где $P_n(x)$ — некоторый многочлен степени n , и пусть

$$\mu = \min_{0 \leq i \leq n+1} |f(x_i) - P_n(x_i)|.$$

Тогда $E_n(f) \geq \mu$.

Доказательство. Пусть $\mu > 0$. (Случай $\mu = 0 \rightarrow E_n \geq 0$ тривиален.) Предположим противное, т. е. $E_n(f) = \|f - Q_n\| < \mu$, и рассмотрим многочлен

$$T_n(x) = P_n(x) - Q_n(x)$$

в точках $x_i \quad i = 0, 1, \dots, n + 1$. Здесь $Q_n(x)$ — многочлен наилучшего равномерного приближения. Имеем

$$P_n(x) - Q_n(x) = (P_n(x) - f(x)) - (Q_n(x) - f(x)).$$

Поскольку $Q_n(x)$ — многочлен наилучшего равномерного приближения,

$$|f(x) - Q_n(x)| \leq |f(x) - P_n(x)|$$

во всех точках отрезка $[a, b]$. Поэтому знак $P_n(x) - Q_n(x)$ в точках $x_i \quad i = 0, 1, \dots, n + 1$ определяется знаком выражения $(P_n(x) - f(x))$. Но в этих точках последнее выражение $n + 2$ раза меняет знак. Мы получили, что многочлен $T_n(x)$ степени n меняет знак $n + 2$ раза и, следовательно, имеет $n + 1$ корень, что невозможно. Теорема доказана.

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

$$f(x_k) - Q_n(x_k) = \alpha(-1)^k \|f - Q_n\| \quad k = 0, 1, \dots, n + 1$$

$\alpha = 1$ или $\alpha = -1$ одновременно для всех k .

Точки, удовлетворяющие условиям теоремы, называются точками чебышевского альтернанса.

Иными словами, теорема утверждает, что для многочлена наилучшего равномерного приближения в точках чебышевского альтернанса разность $f(x) - Q_n(x)$ должна по абсолютной величине достигать своего максимального значения, равного $\|f - Q_n\|$, и менять знак при последовательном переходе от точки x_k к точке x_{k+1} .

Доказательство. Мы докажем только достаточность. С доказательством необходимости можно ознакомиться в [2]. Пусть $Q_n(x)$ — некоторый многочлен степени n , для которого в точках чебышевского альтернанса выполнено условие теоремы, и пусть $L = \|f - Q_n\|$. Согласно теореме Валле — Пуссена $L = \mu \leq E_n(f)$. Но, по определению, $E_n(f) \leq L$. Следовательно, $E_n(f) = L$ и многочлен $Q_n(x)$ есть многочлен наилучшего равномерного приближения.

В случае гильбертова пространства была доказана теорема единственности. Докажем эту теорему для пространства C .

Теорема. Многочлен наилучшего приближения непрерывной функции един-

наилучшего приближения

$$Q_n^1(x) \neq Q_n^2(x) \quad \text{такие, что} \quad \|f - Q_n^1\| = \|f - Q_n^2\| = E_n(f).$$

Тогда

$$\left\| f - \frac{Q_n^1 + Q_n^2}{2} \right\| \leq \left\| \frac{f - Q_n^1}{2} \right\| + \left\| \frac{f - Q_n^2}{2} \right\| = E_n(f),$$

следовательно, многочлен

$$Q_n^0 = \frac{Q_n^1 + Q_n^2}{2}$$

также является многочленом наилучшего равномерного приближения. Пусть x_0, x_1, \dots, x_{n+1} — точки чебышевского альтернанса многочлена $Q_n^0(x)$. В этих точках

$$\left| f(x_k) - \frac{Q_n^1(x_k) + Q_n^2(x_k)}{2} \right| = E_n(f) \quad k = 0, 1, \dots, n+1$$

или

$$|(f(x_k) - Q_n^1(x_k)) + f(x_k) - Q_n^2(x_k)| = 2E_n(f) \quad k = 0, 1, \dots, n+1.$$

Но в этих точках

$$|f(x_k) - Q_n^i(x_k)| \leq E_n(f) \quad i = 1, 2.$$

Поэтому последнее равенство возможно, лишь если

$$f(x_k) - Q_n^1(x_k) = f(x_k) - Q_n^2(x_k) \quad k = 0, 1, \dots, n+1.$$

Но это означает, что два многочлена совпадают в $n+2$ точках, что возможно лишь в случае, если они равны. Теорема доказана.

Рассмотрим примеры наилучшего равномерного приближения.

1. Непрерывная функция $f(x)$ приближается многочленом нулевой степени.

Пусть $M = \max_{[a,b]} f(x)$ и $m = \min_{[a,b]} f(x)$. Тогда таким многочленом будет

$$y = \frac{M - m}{2},$$

поскольку для него есть две точки чебышевского альтернанса x_1 и x_2 , такие, что $f(x_1) = m$ и $f(x_2) = M$.

2. $f(x) = |x - 1|$ на отрезке $[0, 4]$ приближается многочленом наилучшего равномерного приближения первой степени $y = ax + b$.

Пусть $f(0) - b = d$, $f(1) - (a + b) = -d$, $f(4) - (4a + b) = d$. Очевидно, что точки $x_1 = 0$, $x_2 = 1$ и $x_3 = 4$ являются точками чебышевского альтернанса. Решая систему

$$1 - b = d, \quad -(a + b) = -d, \quad 3 - (4a + b) = d,$$

находим $a = 1/2$, $b = 3/4$ и многочлен первой степени наилучшего равномерного приближения имеет вид

$$y = \frac{x}{2} + \frac{1}{4}.$$

Сравните с многочленом, полученным в § 12 для этой же функции в гильбертовом пространстве L^2 на том же отрезке.

Эта же методика построения многочлена наилучшего равномерного приближения 1-й степени может быть применена в случае, если функция $f(x)$ есть многочлен 2-й степени.

3. Выпуклая (вогнутая) дифференцируемая функция $f(x)$ на отрезке $[a, b]$ приближается многочленом 1-й степени.

Построим прямую, проходящую через точки $(a, f(a))$ и $(b, f(b))$. Уравнение этой прямой

$$y = \frac{f(b) - f(a)}{b - a}(x - a) + f(a).$$

Найдем точку x_1 , касательная в которой к графику функции $f(x)$ параллельна к этой прямой, т. е. решим уравнение

$$f'(x_1) = \frac{f(b) - f(a)}{b - a}.$$

Уравнение касательной в этой точке

$$y = \frac{f(b) - f(a)}{b - a}(x - x_1) + f(x_1).$$

Проведем теперь третью прямую, параллельную касательной на равном расстоянии от первой прямой и касательной. Уравнение этой прямой и будет искомым многочленом первой степени наилучшего равномерного приближения. В самом деле, для этого многочлена есть три точки чебышевского альтернанса a , x_1 , b , где разность $f(x) - P_1(x)$ максимальна и меняет знак. Найдем уравнение этой прямой. Для этого достаточно найти точки пересечения

первой прямой и касательной с осью OX и провести прямую, проходящую через середину полученного отрезка с заданным угловым коэффициентом

$$k = \frac{f(b) - f(a)}{b - a}.$$

Первая прямая пересекается с осью OX в точке

$$x_2 = a - f(a) \frac{b - a}{f(b) - f(a)},$$

а касательная — в точке

$$x_3 = x_1 - f(x_1) \frac{b - a}{f(b) - f(a)}.$$

Точка x_4 пересечения искомой прямой с осью Ox есть полусумма точек x_2 и x_3

$$x_4 = \frac{1}{2} \left[a + x_1 - \frac{b - a}{f(b) - f(a)} (f(a) - f(x_1)) \right].$$

Наконец, уравнение прямой, пересекающей ось OX в точке x_4 с угловым коэффициентом k , имеет вид

$$y = \frac{f(b) - f(a)}{b - a} (x - x_4).$$

Изложенная методика позволяет построить многочлен 1-й степени наилучшего равномерного приближения для дифференцируемой функции $f(x)$ и достигающей максимума (минимума) во внутренней точке отрезка $[a, b]$. Построение этого алгоритма предлагается проделать самостоятельно.

§ 15. Численные методы решения задачи Коши
для обыкновенных дифференциальных уравнений и систем

Пусть на отрезке $[t_0, T]$ требуется найти решение дифференциального уравнения

$$\dot{x} = f(t, x), \quad (15.1)$$

удовлетворяющего начальному условию

$$x(t_0) = x_0. \quad (15.2)$$

Разобьем отрезок $[t_0, T]$ на отрезки $[t_j, t_{j+1}]$, $j = 0, 1, \dots, n-1$, $t_{j+1} - t_j = h$ и рассмотрим алгоритмы приближенного нахождения $x_j = x(t_j)$ — значения решения задачи (15.1)— (15.2). Простейшим методом нахождения приближенного решения задачи Коши является метод Эйлера, согласно которому значение решения в точке t_{j+1} находится по формуле

$$x_{j+1} = x_j + hf(t_j, x_j). \quad (15.3)$$

Эту формулу можно получить из следующих соображений. Разложим $x_{j+1} = x(t_j + h)$ в ряд по степеням h , ограничиваясь лишь линейными членами разложения

$$x(t_j + h) = x(t_j) + h\dot{x}(t_j) + O(h^2).$$

Так как $\dot{x}(t_j) = f(t_j, x_j)$, получаем формулу (15.3). При этом замечаем, что погрешность на шаге имеет величину порядка h^2 . Для практических расчетов эта формула применяется достаточно редко из-за высокой погрешности на шаге. Если, например, требуется найти решение на отрезке длины 100 с шагом 0.01, то при расчете по методу Эйлера значений в конце интервала итоговая погрешность будет порядка единицы, что обычно неприемлемо. Более того, если искомое решение неустойчиво по Ляпунову, то расчет решений с помощью такого грубого метода просто теряет смысл. Эти соображения привели к построению более точных методов численного интегрирования дифференциальных уравнений.

Определение. Метод интегрирования дифференциальных уравнений будем называть методом k -го порядка, если погрешность на шаге порядка $O(h^{k+1})$.

Рассматриваемое ниже семейство методов носит название методов Рунге — Кутты.

Рассмотрим равенство

$$x(t+h) = x(t) + \int_0^h \dot{x}(t+\tau) d\tau. \quad (15.4)$$

Воспользуемся формулами приближенного вычисления интеграла, например формулой трапеций

$$x(t+h) = x(t) + \frac{h}{2}[\dot{x}(t) + \dot{x}(t+h)] + O(h^3)$$

или

$$x(t+h) = x(t) + \frac{h}{2}[f(t, x(t)) + f(t+h, x(t+h))] + O(h^3),$$

что приводит к формуле

$$x_{j+1} = x_j + \frac{h}{2}[f(t_j, x_j) + f(t_{j+1}, x_{j+1})]$$

с погрешностью на шаге порядка $O(h^3)$. Однако полученная расчетная формула неприменима, поскольку правая часть неявно зависит от x_{j+1} . Заменяем в правой части x_{j+1} на некоторую другую известную величину, такую, чтобы итоговая погрешность на шаге по-прежнему была порядка $O(h^3)$. Если в качестве такой величины взять x^* , вычисленную по методу Эйлера

$$x^* = x_j + hf(t_j, x_j),$$

то итоговая погрешность по-прежнему останется порядка $O(h^3)$. В самом деле, $x(t+h) - x^* = O(h^2)$ и по теореме Лагранжа о конечных приращениях

$$f(t+h, x(t+h)) - f(t+h, x^*) = \frac{\partial f(t, x)}{\partial x}(x(t+h) - x^*).$$

В результате этих рассуждений приходим к расчетным формулам

$$Q_1 = hf(t_k, x_k), \quad Q_2 = hf(t_k + h, x_k + Q_1),$$

$$x_{k+1} = x_k + \frac{Q_1 + Q_2}{2}. \quad (15.5)$$

Вычисляя интеграл (15.4) по формуле прямоугольников

$$x_{j+1} = x_j + hf\left(t_j + \frac{h}{2}, x\left(t_j + \frac{h}{2}\right)\right)$$

и заменяя $x(t_j + \frac{h}{2})$ на

$$x^* = x_j + \frac{h}{2}f(t_j, x_j),$$

вычисленную по формуле Эйлера, получаем другие расчетные формулы

$$\begin{aligned} Q_1 &= hf(t_k, x_k), & Q_2 &= hf(t_k + \frac{h}{2}, x_k + \frac{Q_1}{2}), \\ x_{k+1} &= x_k + Q_2. \end{aligned} \quad (15.6)$$

В случае системы дифференциальных уравнений, например

$$\begin{aligned} \dot{x}(t) &= f_1(t, x, y) & x(t_0) &= x_0 \\ \dot{y}(t) &= f_2(t, x, y) & y(t_0) &= y_0, \end{aligned}$$

те же соображения приводят к формулам

$$\begin{aligned} Q_1^x &= hf_1(t_k, x_k, y_k), & Q_1^y &= hf_2(t_k, x_k, y_k), \\ Q_2^x &= hf_1(t_k + h, x_k + Q_1^x, y_k + Q_1^y), \\ Q_2^y &= hf_2(t_k + h, x_k + Q_1^x, y_k + Q_1^y), \\ x_{k+1} &= x_k + \frac{Q_1^x + Q_2^x}{2}, \\ y_{k+1} &= y_k + \frac{Q_1^y + Q_2^y}{2}. \end{aligned} \quad (15.5').$$

Формула (15.6) в случае системы двух уравнений имеет вид

$$\begin{aligned} Q_1^x &= hf_1(t_k, x_k, y_k), & Q_1^y &= hf_2(t_k, x_k, y_k), \\ Q_2^x &= hf_1(t_k + \frac{h}{2}, x_k + \frac{Q_1^x}{2}, y_k + \frac{Q_1^y}{2}), \\ Q_2^y &= hf_2(t_k + \frac{h}{2}, x_k + \frac{Q_1^x}{2}, y_k + \frac{Q_1^y}{2}), \\ x_{k+1} &= x_k + Q_2^x, \\ y_{k+1} &= y_k + Q_2^y. \end{aligned} \quad (15.6')$$

Полученные формулы реализуют методы Рунге — Кутта 2-го порядка.

Изложим схему получения вычислительных формул для методов более высокого порядка. Будем последовательно вычислять

$$Q_1(h) = hf(t, x),$$

$$Q_2(h) = hf(t + \alpha_2 h, x + \beta_{21} Q_1(h)),$$

.

$$Q_p(h) = hf(t + \alpha_p h, x + \beta_{p1} Q_1(h) + \dots + \beta_{p,p-1} Q_{p-1}(h))$$

и положим

$$x(t+h) \approx z(h) = x(t) + \sum_{i=1}^p d_i Q_i(h).$$

Параметры α_i , β_{ij} , d_i выберем так, чтобы величина $\varphi(h) = y(t+h) - z(h)$ была порядка $O(h^{s+1})$. Величина s называется порядком метода. Условие, наложенное на $\varphi(h)$, означает, что $\varphi(0) = \varphi'(0) = \dots = \varphi^{(s)}(0) = 0$, а $\varphi^{(s+1)}(0) \neq 0$.

Пусть $p=1$. В этом случае $z(h) = x(t) + d_1 h f(t, x)$ и

$$\varphi(h) = x(t+h) - x(t) - d_1 h f(t, x).$$

Тогда

$$\varphi(0) = 0, \quad \varphi'(0) = x'(t) - d_1 f(t, x) = f(t, x)(1 - d_1).$$

Для того чтобы равенство $\varphi'(0) = 0$ выполнялось при любой функции $f(t, x)$, надо, чтобы $d_1 = 1$, и мы получаем формулу Эйлера

$$x(t+h) = x(t) + hf(t, x).$$

Построим формулы 2-го порядка. Пусть $p=2$. В этом случае

$$z(h) = x(t) + d_1 Q_1(h) + d_2 Q_2(h) = x(t) + d_1 h f(t, x) + d_2 h f(t + \alpha_2 h, x + \beta_{21} h f(t, x))$$

и

$$\varphi(h) = x(t+h) - x(t) - d_1 h f(t, x) - d_2 h f(t + \alpha_2 h, x + \beta_{21} h f(t, x)).$$

Введем обозначения

$$\bar{t} = t + \alpha_2 h, \quad \bar{x} = x + \beta_{21} h f(t, x).$$

Вычислим $\varphi'(h)$ и $\varphi''(0)$:

$$\varphi'(h) = x'(t+h) - d_1 f(t, x) - d_2 f(\bar{t}, \bar{x}) - d_2 h [\alpha_2 f'_t(\bar{t}, \bar{x}) + \beta_{21} f'_x(\bar{t}, \bar{x}) f(t, x)],$$

$$\varphi''(h) = x''(t+h) - 2d_2 [\alpha_2 f'_t(\bar{t}, \bar{x}) + \beta_{21} f'_x(\bar{t}, \bar{x}) f(t, x)] - d_2 h [\dots].$$

Поскольку мы будем вычислять $\varphi'(0)$ и $\varphi''(0)$, выражение в последних квадратных скобках для нас интереса не представляет.

При $h = 0$ получаем

$$\varphi(0) = x(t) - x(t) = 0,$$

$$\varphi'(0) = x'(t) - d_1 f(t, x) - d_2 f(t, x) = f(t, x)(1 - d_1 - d_2).$$

Далее, поскольку $x''(t) = f'_t(t, x) + f'_x(t, x)x'(t) = f'_t(t, x) + f'_x(t, x)f(t, x)$

$$\varphi''(h) = f'_t(t, x) + f'_x(t, x)f(t, x) - 2d_2\alpha_2 f'_t(t, x) - 2d_2\beta_{21} f'_x(t, x)f(t, x).$$

Вычислим эти выражения при $h = 0$.

$$\varphi(0) = x(t) - x(t) = 0,$$

$$\varphi'(0) = (1 - d_1 - d_2)f(t, x),$$

$$\varphi''(0) = (1 - 2d_2\alpha_2)f'_t(t, x) + (1 - 2d_2\beta_{21})f'_x(t, x)f(t, x).$$

Равенство нулю $\varphi'(0)$ и $\varphi''(0)$ при произвольной функции $f(t, x)$ возможно, если

$$1 - d_1 - d_2 = 0,$$

$$1 - 2d_2\alpha_2 = 0,$$

$$1 - 2d_2\beta_{21} = 0.$$

Мы получили три уравнения относительно четырех неизвестных. Задавая один из параметров, получаем семейство методов Рунге — Кутты 2-го порядка. В частности, при

$$d_1 = d_2 = \frac{1}{2}, \quad \alpha_2 = \beta_{21} = 1$$

получаем формулу (15.5)

$$Q_1 = hf(t_k, x_k), \quad Q_2 = hf(t_k + h, x_k + Q_1), \quad x_{k+1} = x_k + \frac{Q_1 + Q_2}{2}.$$

Формула (15.6) получается при

$$d_1 = 0, \quad d_2 = 1, \quad \alpha_1 = \beta_{21} = \frac{1}{2}.$$

$$Q_1 = hf(t_k, x_k), \quad Q_2 = hf(t_k + \frac{h}{2}, x_k + \frac{Q_1}{2}), \quad x_{k+1} = x_k + Q_2.$$

Мы построили методы второго порядка. Методы более высокого порядка получаются при $p > 2$. Построение расчетных формул для методов связано с

громоздкими вычислениями. Приведем наиболее употребительные формулы для методов 3-го и 4-го порядка.

Методы 3-го порядка.

1.

$$Q_1 = hf(t_k, x_k), \quad Q_2 = hf\left(t_k + \frac{h}{2}, x_k + \frac{Q_1}{2}\right), \quad Q_3 = hf(t_k + h, x_k - Q_1 + 2Q_2),$$

$$x_{k+1} = x_k + \frac{Q_1 + 4Q_2 + Q_3}{6}. \quad (15.7)$$

2.

$$Q_1 = hf(t_k, x_k), \quad Q_2 = hf\left(t_k + \frac{h}{3}, x_k + \frac{Q_1}{3}\right), \quad Q_3 = hf\left(t_k + \frac{2h}{3}, x_k - Q_1 + \frac{2Q_2}{3}\right),$$

$$x_{k+1} = x_k + \frac{Q_1 + 3Q_3}{4}. \quad (15.8)$$

Методы 4-го порядка.

1.

$$Q_1 = hf(t_k, x_k), \quad Q_2 = hf\left(t_k + \frac{h}{2}, x_k + \frac{Q_1}{2}\right),$$

$$Q_3 = hf\left(t_k + \frac{h}{2}, x_k + \frac{Q_2}{2}\right), \quad Q_4 = hf(t_k + h, x_k + Q_3),$$

$$x_{k+1} = x_k + \frac{Q_1 + 2Q_2 + 2Q_3 + Q_4}{6}. \quad (15.9)$$

2.

$$Q_1 = hf(t_k, x_k), \quad Q_2 = hf\left(t_k + \frac{h}{3}, x_k + \frac{Q_1}{3}\right),$$

$$Q_3 = hf\left(t_k + \frac{2h}{3}, x_k - \frac{Q_1}{2} + Q_2\right), \quad Q_4 = hf(t_k + h, x_k + Q_1 - Q_2 + Q_3),$$

$$x_{k+1} = x_k + \frac{Q_1 + 3Q_2 + 3Q_3 + Q_4}{8}. \quad (15.10)$$

Напомним, что в случае системы

$$\dot{x}(t) = f_1(t, x, y)$$

$$\dot{y}(t) = f_2(t, x, y)$$

величины Q_i становятся двумерными векторами $Q_i = (Q_i^x, Q_i^y)^T$, причем при вычислении Q_i^x используется функция $f_1(t, x, y)$, а при вычислении Q_i^y — функция $f_2(t, x, y)$ так же, как и при получении формул (15.5') и (15.6') для систем из формул (15.5) и (15.6) для скалярного уравнения.

В заключение отметим, что методы Рунге — Кутта относятся к так называемым одношаговым методам. В одношаговых методах решение в точке $t_1 = t_0 + h$ строится с помощью информации о решении, полученной на предыдущем шаге. В многошаговых методах при вычислении x_k используется информация о решении, полученная на нескольких предыдущих шагах.

Опишем коротко одну из процедур получения двушагового метода. Пусть для простоты $t_{i+1} = t_i + h$. Первый шаг сделаем по одному из методов Рунге — Кутта, вычислив x_1 в точке $t_1 = t_0 + h$. Построим интерполяционный многочлен в форме Ньютона по двукратной точке t_0 и t_1 :

$$L_3(t) = x_0 + \dot{x}_0(t - t_0) + \frac{x_1 - x_0 - h\dot{x}_0}{h^2}(t - t_0)^2.$$

Положим $t_2 = t_1 + h$ и будем считать, что $x_2 = x(t_1 + h) = L_3(t_1 + h)$. Получим

$$x_2 = -3x_0 + 4x_1 - 2h\dot{x}_0$$

и мы приходим к формуле

$$x_{k+1} = -3x_{k-1} + 4x_k - 2h\dot{x}_{k-1}.$$

Из оценки интерполяционного многочлена следует, что построенный метод есть метод 2-го порядка.

Для увеличения точности нужно строить интерполяционный многочлен более высокого порядка, сделав предварительно несколько шагов по методу Рунге — Кутта. Естественно, что при увеличении числа точек для построения интерполяционного многочлена формула становится более громоздкой. Кроме того, какой смысл переходить к новым расчетным формулам, а не продолжать расчеты по методу Рунге — Кутта? Все эти причины и приводят к тому, что многошаговые методы используются в вычислительной практике значительно реже, чем методы Рунге — Кутта.

§ 16. Разностные методы решения краевых задач
для обыкновенных дифференциальных уравнений 2-го порядка
Понятие аппроксимации, устойчивости и сходимости разностных схем

Рассмотрим следующую задачу. Требуется найти решение дифференциального уравнения

$$y'' + p(x)y' + q(x)y = f(x) \quad 0 \leq x \leq 1, \quad (16.1)$$

удовлетворяющее на концах отрезка краевым условиям вида

$$\begin{aligned} \alpha_1 y(0) + \beta_1 y'(0) &= \gamma_1, \\ \alpha_2 y(1) + \beta_2 y'(1) &= \gamma_2. \end{aligned} \quad (16.2)$$

Будем говорить, что уравнение (16.1) и краевые условия (16.2) задают в пространстве дважды непрерывно дифференцируемых функций дифференциальный оператор

$$Ly = f.$$

Разобьем отрезок $[0,1]$ точками x_i

$$x_i = a + ih, \quad i = 1, 2, \dots, n, \quad h = \frac{1}{n}$$

и аппроксимируем в точках x_i $i = 0, 2, \dots, n$ производную $y'(x_i)$ выражением

$$y'(x_i) \approx \frac{y_{i+1} - y_i}{h}, \quad \text{где } y_i = y(x_i), \quad (16.3)$$

или выражением

$$y'(x_i) \approx \frac{y_i - y_{i-1}}{h}. \quad (16.4)$$

Введем обозначения

$$\Delta y_i = y_{i+1} - y_i, \quad \nabla y_i = y_i - y_{i-1},$$

$$y_{x,i} = \frac{\Delta y_i}{h} = \frac{y_{i+1} - y_i}{h},$$

$$y_{\bar{x},i} = \frac{\nabla y_i}{h} = \frac{y_i - y_{i-1}}{h}.$$

Второй производной $y''(x_i)$ поставим в соответствие выражение

$$y_{x\bar{x},i} = y_{\bar{x}x,i} = \frac{\Delta \nabla y_i}{h^2} = \frac{\nabla \Delta y_i}{h^2} = \frac{y_{i-1} - 2y_i + y_{i+1}}{h^2}$$

и дифференциальному уравнению (16.1) в точках x_i можно поставить в соответствие следующее уравнение, которое будем называть разностным

$$\frac{y_{i-1} - 2y_i + y_{i+1}}{h^2} + p_i \frac{y_{i+1} - y_i}{h} + q_i y_i = f_i \quad i = 1, 2, \dots, n-1 \quad (16.5)$$

После приведения подобных членов уравнение (16.5) запишем в виде

$$b_i y_{i+1} - c_i y_i + d_i y_{i-1} = -\varphi_i \quad i = 1, 2, \dots, n-1, \quad (16.6)$$

а краевые условия в виде

$$y_0 = \chi_1 y_1 + \mu_1 \quad y_n = \chi_2 y_{n-1} + \mu_2. \quad (16.7)$$

Здесь

$$b_i = 1 + p_i h, \quad c_i = 2 + p_i h - q_i h^2, \quad d_i = 1, \quad \varphi_i = f_i h^2, \\ \chi_1 = \frac{\beta_1}{\beta_1 - \alpha_1 h}, \quad \mu_1 = \frac{\gamma_1 h}{\alpha_1 h - \beta_1}, \quad \chi_2 = \frac{\beta_2}{\alpha_2 h + \beta_2}, \quad \mu_2 = \frac{\gamma_2 h}{\alpha_2 h + \beta_2}.$$

При аппроксимации $y'(x_n)$ в правом граничном условии использовались формулы (16.4).

Задачу (16.6)–(16.7) будем в дальнейшем называть разностной или сеточной краевой задачей, а функцию целочисленного аргумента y_i — сеточной функцией. Эта задача представляет собой с учетом граничных условий (16.7) систему $n+1$ уравнений относительно $n+1$ неизвестного y_i $i = 0, 1, \dots, n$. Ниже будут рассмотрены методы решения таких задач, однако, прежде чем приступить к их рассмотрению, следует выяснить, насколько решение непрерывной задачи (16.1) – (16.2) отличается от решения сеточной задачи (16.6) – (16.7).

Пусть $u_i = u(x_i)$ — значение решения задачи (16.1) – (16.2) в узле x_i , а y_i — значение решения сеточной задачи в том же узле.

Определение 1. Будем говорить, что решение задачи (16.6) – (16.7) сходится (сходится со скоростью h^k), если $z_i = y_i - u_i \rightarrow 0$ ($z_i = O(h^k)$) при $h \rightarrow 0$.

Обозначим задачу (16.1) – (16.2) оператором

$$Lu = f,$$

а задачу (16.6) - (16.7) — оператором

$$ly_i = \varphi_i.$$

Пусть по-прежнему $z_i = y_i - u_i$, т. е. $y_i = z_i + u_i$. Тогда для z_i получаем задачу

$$lz_i = \phi_i,$$

где $\phi_i = \varphi_i - lu_i$.

Определение 2. Будем говорить, что разностная задача $ly_i = \varphi_i$ аппроксимирует задачу $Lu = f$ (аппроксимирует с порядком h^k), если для $z_i = y_i - u_i$ в задаче $lz_i = \phi_i$ для правой части имеет место оценка $\phi_i = O(h^k)$.

Определение 3. Будем говорить, что разностная схема

$$ly_i = \varphi_i$$

устойчива, если имеет место оценка

$$\|y_i\| \leq A\|\varphi_i\|,$$

причем постоянная A не зависит от h . Если A зависит от h , схема называется условно устойчивой.

Из аппроксимации разностной схемы и ее устойчивости очевидно следует ее сходимость.

Выясним порядок аппроксимации разностной схемой (16.6) - (16.7) исходной задачи (16.1) - (16.2). Пусть $u(x_i)$ — решение задачи (16.1) - (16.2) в узле x_i . Тогда

$$u(x_i \pm h) = u(x_i) \pm hu'(x_i) + \frac{h^2}{2!}u''(x_i) \pm \frac{h^3}{3!}u'''(x_i) + \frac{h^4}{4!}u^{(4)}(x_i) + O(h^5).$$

Поэтому

$$\begin{aligned} \frac{u_{i+1} - u_i}{h} &= u'(x_i) + O(h), \\ \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} &= u''(x_i) + O(h^2) \end{aligned}$$

и левая часть в (16.5) есть

$$u''(x_i) + p_i u'(x_i) + q_i u(x_i) + O(h).$$

Таким образом, задача (16.6) - (16.7) аппроксимирует задачу (16.1) - (16.2) с порядком $O(h)$, причем если $p(x) = 0$ и $\beta_1 = \beta_2 = 0$, то порядок аппроксимации $O(h^2)$.

§ 17. Методы решения краевых задач 2-го порядка.

Методы прогонки, метод стрельбы

Как отмечалось, краевая задача

$$b_i y_{i+1} - c_i y_i + d_i y_{i-1} = -\varphi_i \quad i = 1, 2, \dots, n-1 \quad (17.1)$$

$$y_0 = \chi_1 y_1 + \mu_1 \quad y_n = \chi_2 y_{n-1} + \mu_2 \quad (17.2)$$

представляет собой линейную систему $n+1$ уравнения относительно неизвестных y_0, y_1, \dots, y_n . Особенностью матрицы этой системы является то, что эта матрица является трехдиагональной, т. е. у этой матрицы отличны от нуля только элементы на главной и двух соседних диагоналях. Ниже будут предложены простые методы решения таких систем.

Метод прямой (правой) прогонки.

Теорема 1. Решение задачи (17.1) - (17.2) может быть представлено в виде

$$y_i = \alpha_{i+1} y_{i+1} + \beta_{i+1} \quad i = n-1, n-2, \dots, 1, 0, \quad (17.3)$$

где

$$y_n = \frac{\mu_2 + \chi_2 \beta_n}{1 - \alpha_n \chi_2}, \quad (17.4)$$

а α_i и β_i определяются из рекуррентных соотношений

$$\alpha_{i+1} = \frac{b_i}{c_i - \alpha_i d_i} \quad i = 1, 2, \dots, n-1, \quad \alpha_1 = \chi_1, \quad (17.5)$$

$$\beta_{i+1} = \frac{d_i \beta_i + \varphi_i}{c_i - \alpha_i d_i} \quad i = 1, 2, \dots, n-1, \quad \beta_1 = \mu_1. \quad (17.6)$$

Доказательство. Левое граничное условие $y_0 = \chi_1 y_1 + \mu_1$ имеет вид (17.3), где $\alpha_1 = \chi_1$, а $\beta_1 = \mu_1$. Уравнение (17.1) при $i=1$

$$b_1 y_2 - c_1 y_1 + d_1 y_0 = -\varphi_1$$

после подстановки y_0 приобретает вид

$$b_1 y_2 - c_1 y_1 + d_1 (\chi_1 y_1 + \mu_1) = -\varphi_1.$$

Или

$$y_1 = \frac{b_1}{c_1 - d_1 \chi_1} y_2 + \frac{d_1 + \varphi_1 \mu_1}{c_i - \alpha_i d_i},$$

т. е. при $i=1$ уравнение (17.1) представимо в виде (17.3), где α_1 и β_1 имеют вид (17.5), (17.6). Предположим теперь, что $i-1$ -е уравнение (17.1) представимо в виде (17.3). Подставим y_{i-1} в i -е уравнение (17.1), получим

$$b_i y_{i+1} - c_i y_i + d_i(\alpha_i y_i + \beta_i) = -\varphi_i$$

или

$$y_i = \frac{b_i}{c_i - \alpha_i d_i} y_{i+1} + \frac{d_i \beta_i + \varphi_i}{c_i - \alpha_i d_i}.$$

Следовательно, и i -е уравнение (17.1) представимо в виде (17.3) и коэффициенты α_i и β_i представимы в виде (17.5), (17.6). Последнее уравнение (17.3) и правое граничное условие представляют собой систему двух уравнений относительно y_{n-1} и y_n

$$y_n = \chi_2 y_{n-1} + \mu_2$$

$$y_{n-1} = \alpha_n y_n + \beta_n,$$

решая которую получаем (17.4).

Теорема доказана. Эта теорема дает удобный алгоритм нахождения решения задачи (17.1) - (17.2). В самом деле, левое граничное условие задает α_1 и β_1 . Далее, в цикле $i = 1, 2, \dots, n-1$ определяются по формулам (17.5), (17.6) прогоночные коэффициенты α_{i+1} и β_{i+1} . Затем по (17.4) определяется y_n и далее обратным ходом по формулам (17.3) определяются y_i $i = n-1, n-2, \dots, 0$.

Рассмотренный в теореме 1 метод, при котором определение y_i производится последовательно справа налево, называют методом прямой или правой прогонки. Аналогично доказывается

Теорема 2 (метод обратной или левой прогонки). Решение задачи (17.1) (17.2) представимо в виде

$$y_{i+1} = \xi_{i+1} y_i + \eta_{i+1} \quad i = 0, 1, \dots, n-1, \quad (17.7)$$

где

$$y_0 = \frac{\mu_1 + \chi_1 \eta_1}{1 - \xi_1 \chi_1}, \quad (17.8)$$

$$\xi_i = \frac{d_i}{c_i - b_i \xi_{i+1}} \quad i = n-1, n-2, \dots, 1, \quad \xi_n = \chi_2, \quad (17.9)$$

$$\eta_i = \frac{b_i \eta_{i+1} + \varphi_i}{c_i - b_i \xi_{i+1}} \quad i = n-1, n-2, \dots, 1, \quad \eta_n = \mu_2. \quad (17.10)$$

Иногда используется метод встречной прогонки, который состоит в следующем. Зафиксируем некоторое $i = l$ и, используя метод прямой прогонки, определим α_i и β_i $i = 1, 2, \dots, l + 1$. Затем, используя метод обратной прогонки, определим ξ_i и η_i $i = n - 1, n - 2, \dots, l + 1$. Далее, из системы

$$y_l = \alpha_{l+1}y_{l+1} + \beta_{l+1}$$

$$y_{l+1} = \xi_{l+1}y_l + \eta_{l+1}$$

определим y_l и y_{l+1} и обратным ходом определим остальные y_i .

Очевидно, что методы прогонки применимы, если в процессе вычислений знаменатели в формулах (17.4) — (17.6) и аналогичных им в методе обратной прогонки не обращаются в нуль. Достаточный признак применимости метода прямой прогонки формулируется в следующей теореме.

Теорема 3. При выполнении условий

$$|c_i| \geq |b_i| + |d_i| \quad i = 1, 2, \dots, n - 1 \quad (17.11)$$

$$|\chi_1| \leq 1, \quad |\chi_2| \leq 1,$$

причем одно из этих неравенств строгое, знаменатели в формулах (17.4) - (17.6) не обращаются в нуль и $|\alpha_i| \leq 1$, $i = 1, \dots, n$.

Доказательство. Заметим, что $\alpha_0 = \chi_1$ и по условию $|\alpha_0| \leq 1$. Предположим, что $|\alpha_i| \leq 1$, и покажем, что и $|\alpha_{i+1}| \leq 1$. Рассмотрим формулу

$$\alpha_{i+1} = \frac{b_i}{c_i - \alpha_i d_i}$$

и покажем, что $|c_i - d_i \alpha_i| \geq |b_i|$. В самом деле, используя (17.11) и индуктивное предположение, получаем

$$|c_i - d_i \alpha_i| - |b_i| \geq |c_i| - |d_i| |\alpha_i| - |b_i| \geq |d_i| (1 - |\alpha_i|) \geq 0,$$

что и доказывает, что $|c_i - d_i \alpha_i| \geq |b_i|$ и $|\alpha_{i+1}| \leq 1$.

Обратимся к формуле (17.4). Заметим, что если $|\alpha_k| < 1$, то $|\alpha_i| < 1$, при $i > k$. Поэтому, если $|\chi_1| < 1$, то и $|\alpha_n| < 1$ и знаменатель в (17.4) не обращается в нуль. Если же $|\alpha_n| \leq 1$, то условие $|\chi_2| < 1$ гарантирует отличие от нуля знаменателя в (17.4).

Теорема доказана.

Метод стрельбы. Рассмотрим систему уравнений (17.1)

$$b_i y_{i+1} - c_i y_i + d_i y_{i-1} = -\varphi_i \quad i = 1, 2, \dots, n - 1.$$

Последовательность этих уравнений можно рассматривать как систему $n-1$ уравнений относительно $n+1$ неизвестного y_0, y_1, \dots, y_{n+1} . Рассмотрим сначала однородную систему

$$b_i y_{i+1} - c_i y_i + d_i y_{i-1} = 0 \quad i = 1, 2, \dots, n-1. \quad (17.12)$$

Очевидно, что разность двух решений неоднородной системы есть решение однородной системы. Поэтому, для того чтобы описать все решения неоднородной системы, надо к формуле общего решения однородной системы прибавить любое частное решение неоднородной системы (17.1). Мы будем предполагать, что коэффициенты системы (17.12) b_i, c_i, d_i таковы, что ранг этой системы в точности равен $n-1$. Тогда эта система имеет два линейно независимых векторных решения — вектора $U = (u_0, u_1, \dots, u_{n+1})^T$ и $V = (v_0, v_1, \dots, v_{n+1})^T$. Если U и V есть решения (17.12) и если положить $u_0 = 1$, а $u_1 = 0$, то последовательность уравнений (17.12) становится рекуррентной для определения u_2, u_3, \dots, u_n . Аналогично, если положить $v_0 = 0$, $v_1 = 1$, то также из (17.12) можно рекуррентно определить v_2, v_3, \dots, v_n . Очевидно, что два этих вектора линейно независимы и их линейная комбинация

$$C_1 U + C_2 V$$

есть общее решение однородной системы (17.12). Частное решение неоднородной системы (17.1) — вектор $W = (w_0, w_1, \dots, w_n)^T$ также рекуррентно можно определить из (17.1), положив $w_0 = 0$, $w_1 = 0$. Таким образом, общее решение (17.1) можно представить в виде

$$y_i = c_1 u_i + c_2 v_i + w_i. \quad (17.13)$$

Определим теперь c_1 и c_2 так, чтобы решение удовлетворяло краевым условиям (17.2)

$$y_0 = \chi_1 y_1 + \mu_1, \quad y_n = \chi_2 y_{n-1} + \mu_2.$$

Из левого краевого условия получаем

$$c_1 = \chi_1 c_2 + \mu_1,$$

из правого краевого условия —

$$c_1 u_n + c_2 v_n + w_n = \chi_2 (c_1 u_{n-1} + c_2 v_{n-1} + w_{n-1}) + \mu_2.$$

Таким образом, получена система

$$c_1 = \chi_1 c_2 + \mu_1$$

$$c_1(u_1 - \chi_2 u_{n-1}) + c_2(v_n - \chi_2 v_{n-1}) = \mu_2 + \chi_2 w_{n-1} - w_n$$

Найдем из этой системы c_1 и c_2 и подставим в (17.13). Получим решение уравнения (17.1), удовлетворяющее краевым условиям (17.2).

Мы изучили методы решения краевых задач для разностных уравнений. Но для того, чтобы решение разностной схемы сходилось к решению исходной задачи, необходимо, чтобы эта разностная схема была устойчивой. Для исследования на устойчивость нужно получить оценку решения разностной задачи через правую часть. Для этого исследуем некоторые общие свойства разностных уравнений.

§ 18. Принцип максимума для разностных уравнений 2-го порядка.
Оценка решений первой краевой задачи

Рассмотрим следующую краевую задачу для разностного уравнения 2-го порядка:

$$ly_i \equiv b_i y_{i+1} - c_i y_i + d_i y_{i-1} = -\varphi_i \quad i = 1, 2, \dots, n-1, \quad (18.1)$$

$$y_0 = \mu_1, \quad y_n = \mu_2, \quad (18.2)$$

причем

$$b_i > 0, \quad d_i > 0, \quad c_i \geq b_i + d_i. \quad (18.3)$$

Теорема 1 (Принцип максимума). Пусть выполнены условия (18.3) и $\varphi_i \leq 0$ ($\varphi_i \geq 0$) $i = 1, 2, \dots, n-1$. Тогда решение задачи (18.1) — (18.2), отличное от постоянной, не может принимать наибольшее положительное (наименьшее отрицательное) значение во внутренних точках.

Доказательство. Пусть $\varphi_i \leq 0$. $i = 1, 2, \dots, n-1$. Предположим противное, что в некоторой внутренней точке $i = p$ выполнено $y_p = \max_{0 \leq i \leq n} y_i = M > 0$. Так как $y_i \neq \text{const}$, найдется внутренний узел k , в котором достигает своего положительного максимума $y_k = M > 0$. Тогда найдется соседний узел, например $k-1$, в котором $y_{k-1} < M$. Может случиться, что в нескольких узлах, соседних с узлом k , $y_i = M$. Тогда в качестве узла k возьмем меньший по номеру и запишем (18.1) в узле k :

$$b_k y_{k+1} - c_k y_k + d_k y_{k-1} = b_k (y_{k+1} - y_k) - d_k (y_k - y_{k-1}) - (c_k - b_k - d_k) y_k < 0.$$

Но $\varphi_i \leq 0$, $i = 1, 2, \dots, n-1$, т. е. это выражение должно быть неотрицательно. Получено противоречие. Второе утверждение доказывается аналогично. Теорема доказана.

Следствием доказанной теоремы являются следующие утверждения.

Теорема 2. Пусть выполнены условия (18.3) и $\varphi_i \geq 0$. $i = 1, 2, \dots, n-1$ и $\mu_1 \geq 0$, $\mu_2 \geq 0$. Тогда $y_i \geq 0$, $i = 1, 2, \dots, n-1$. Если же $\varphi_i \leq 0$, $i = 1, 2, \dots, n-1$ и $\mu_1 \leq 0$, $\mu_2 \leq 0$. Тогда $y_i \leq 0$, $i = 1, 2, \dots, n-1$.

Доказательство. Пусть $\varphi_i \geq 0$, $i = 1, 2, \dots, n-1$ и $y_i < 0$ хотя бы в одном внутреннем узле. Тогда y_i достигает наименьшего отрицательного значения во внутреннем узле, что невозможно.

Теорема единственности. Если выполнены условия (18.3), то краевая задача

$$b_i y_{i+1} - c_i y_i + d_i y_{i-1} = 0 \quad i = 1, 2, \dots, n-1,$$

$$y_0 = 0, \quad y = 0,$$

имеет только тривиальное решение $y_i = 0 \quad i = 0, 1, \dots, n$.

Доказательство. По теореме 2 из $\varphi_i \leq 0, \mu_1 \leq 0, \mu_2 \leq 0$ следует $y_i \leq 0$. Аналогично, из $\varphi_i \geq 0, \mu_1 \geq 0, \mu_2 \geq 0$ следует $y_i \geq 0$. Следовательно, $y_i = 0$. Что касается задачи (18.1) - (18.2), то решение этой задачи единственно, поскольку разность двух решений удовлетворяет однородному уравнению, которое имеет только тривиальное решение.

Теорема сравнения. Пусть y_i — решение задачи (18.1) - (18.2), а \bar{y}_i — решение задачи

$$l\bar{y}_i = -\bar{\varphi}_i \quad i = 1, 2, \dots, n-1, \quad \bar{y}_0 = \bar{\mu}_1, \quad \bar{y}_n = \bar{\mu}_2,$$

коэффициенты которой удовлетворяют условиям (18.3) и выполнены условия

$$|\varphi_i| \leq \bar{\varphi}_i, \quad i = 1, 2, \dots, n-1, \quad |\mu_1| \leq \bar{\mu}_1, \quad |\mu_2| \leq \bar{\mu}_2.$$

Тогда $|y_i| \leq |\bar{y}_i| \quad i = 0, 1, 2, \dots, n$.

Доказательство. По теореме 2 $|\bar{y}_i| \geq 0$. Заметим далее, что для

$$u = \bar{y}_i - y_i \quad \text{и} \quad v = \bar{y}_i + y_i$$

получаем задачу (18.1) - (18.2) с правыми частями соответственно $\bar{\varphi}_i - \varphi_i \geq 0$ и $\bar{\varphi}_i + \varphi_i \geq 0$ и условиями на границе соответственно $\bar{\mu}_1 - \mu_1 \geq 0$ и $\bar{\mu}_2 + \mu_2 \geq 0$. По теореме 2 $u_i \geq 0$ и $v_i \geq 0$, т. е.

$$\bar{y}_i - y_i \geq 0, \quad \bar{y}_i + y_i \geq 0$$

и

$$-\bar{y}_i \leq y_i \leq \bar{y}_i.$$

Теорема доказана.

Оценка решения первой краевой задачи

Ниже в качестве нормы сеточной функции будет использована первая векторная норма

$$\|y_i\|_c = \max_{0 \leq i \leq n} |y_i|.$$

Теорема. Пусть выполнены условия (18.3), причем

$$g_i = |c_i| - |b_i| - |d_i| > 0 \quad i = 1, 2, \dots, n-1.$$

Тогда справедлива оценка

$$\|y\|_c \leq \left\| \frac{\varphi}{g} \right\| + \max(|\mu_1|, |\mu_2|).$$

Доказательство. Решение задачи (18.1) - (18.2) представим в виде $y^{(1)} + y^{(2)}$, где $y_i^{(1)}$ — решение однородного уравнения с заданными граничными условиями,

$$b_i y_{i+1}^{(1)} - c_i y_i^{(1)} + d_i y_{i-1}^{(1)} = 0 \quad i = 1, 2, \dots, n-1,$$

$$y_0^{(1)} = \mu_1, \quad y_n^{(1)} = \mu_2,$$

а $y_i^{(2)}$ — решение неоднородного уравнения с нулевыми граничными условиями

$$b_i y_{i+1}^{(2)} - c_i y_i^{(2)} + d_i y_{i-1}^{(2)} = -\varphi_i \quad i = 1, 2, \dots, n-1, \quad (18.4)$$

$$y_0^{(2)} = 0, \quad y_n^{(2)} = 0.$$

Покажем, что для $y_i^{(1)}$ справедлива оценка

$$\|y^{(1)}\|_c \leq \max(|\mu_1|, |\mu_2|).$$

В самом деле, для решения задачи

$$b_i \bar{y}_{i+1} - c_i \bar{y}_i + d_i \bar{y}_{i-1} = 0 \quad i = 1, 2, \dots, n-1,$$

$$\bar{y}_0 = \bar{\mu}, \quad \bar{y}_n = \bar{\mu}, \quad \bar{\mu} = \max(|\mu_1|, |\mu_2|)$$

замечаем, что $\bar{y}_i \geq 0$ и может достигать максимума только на границе, т. е. при $i = 0$ или $i = n$. Но по теореме сравнения $|y_i^{(1)}| \leq |\bar{y}_i|$.

Займемся оценкой $y_i^{(2)}$. Пусть максимум модуля y_i достигается в узле $i = p$:

$|y_i| \leq |y_p|, \quad i = 0, 1, \dots, n$. Перепишем (18.4) в виде

$$c_i y_i = b_i y_{i+1} + d_i y_{i-1} + \varphi_i.$$

В этом узле имеем

$$|c_p| |y_p| = |b_p y_{p+1} + d_p y_{p-1} + \varphi_p| \leq (|b_p| + |d_p|) |y_p| + |\varphi_p|.$$

Отсюда

$$|y_p| \leq \frac{|\varphi_p|}{(|c_p| - |b_p| - |d_p|)} \leq \left\| \frac{\varphi}{g} \right\|.$$

Теорема доказана.

Обсудим теперь вопросы аппроксимации, устойчивости и сходимости при решении сеточным методом первой краевой задачи для дифференциального уравнения

$$\begin{aligned} \ddot{u} + p(t)\dot{u} + q(t)u &= f(t) \\ u(0) &= \mu_1, \quad u(1) = \mu_2. \end{aligned}$$

Как показано выше, этой непрерывной задаче можно поставить в соответствие сеточную задачу

$$\begin{aligned} b_i y_{i+1} - c_i y_i + d_i y_{i-1} &= -\varphi_i \quad i = 1, 2, \dots, n-1 \\ y_0 &= \mu_1 \quad y_n = \mu_2, \end{aligned}$$

где

$$b_i = 1 + p_i h, \quad c_i = 2 + p_i h - q_i h^2, \quad d_i = 1 \quad \varphi_i = f_i h^2.$$

При $p(t) \geq c > 0$ условия (18.3) могут быть выполнены при любых ограниченных $q(t)$ за счет уменьшения h . При $p(t) = 0$ и $q(t) > 0$ условия (18.3) не выполнены. Как известно, в этом случае и непрерывная задача также может не иметь решения. Случай $p(t) < 0$ требует дополнительного исследования, однако в приложениях он обычно не реализуется, поскольку физический смысл этого коэффициента — трение, которое в приложениях неотрицательно.

Подведем итог. Мы показали, что задача (16.4)- (16.5) аппроксимирует задачу (16.1)- (16.2) с порядком $O(h)$. При выполнении условий (18.3) первая краевая задача устойчива и, следовательно, решение первой краевой задачи для сеточного уравнения (16.4) сходится к решению первой краевой задачи уравнения (16.1).

§ 19. Разностные уравнения

Определение. Линейное уравнение относительно функции дискретного аргумента (сеточной функции) $y_i = y(i)$ вида

$$a_{0i}y_i + a_{1i}y_{i+1} + \dots + a_{mi}y_{i+m} = f_i \quad i = 0, \pm 1, \pm 2, \dots \quad (19.1)$$

называется разностным уравнением m -го порядка. Если коэффициенты a_0, a_1, \dots, a_m не зависят от i , то уравнение называется линейным разностным уравнением с постоянными коэффициентами. Если все $f_i = 0$, то разностное уравнение называется однородным. Рассмотрим разностные уравнения 1-го и 2-го порядка.

Разностные уравнения и неравенства 1-го порядка.

Рассмотрим разностное уравнение 1-го порядка

$$a_i y_i + b_i y_{i+1} = f_i \quad a_i \neq 0, \quad b_i \neq 0$$

или

$$y_{i+1} = q_i y_i + \varphi_i. \quad (19.2)$$

Если задано y_0 , то это уравнение становится рекуррентным для определения y_i . По индукции нетрудно получить формулу

$$y_{i+1} = y_0 \prod_{k=0}^i q_k + \sum_{k=0}^i \left(\prod_{s=k+1}^i q_s \right) \varphi_k + \varphi_i,$$

доказательство которой предлагается проделать самостоятельно. В случае уравнения с постоянными коэффициентами $q_i = q$ эта формула имеет вид

$$y_{i+1} = q^{i+1} y_0 + \sum_{k=0}^i q^{i-k} \varphi_k. \quad (19.3)$$

Неравенства 1-го порядка.

Теорема. Пусть дано разностное уравнение 1-го порядка

$$u_{i+1} = q y_i + f_i \quad q \geq 0 \quad (19.4)$$

и неравенство первого порядка

$$v_{i+1} \leq q v_i + f_i \quad (19.5)$$

и пусть $u_0 = v_0$. Тогда имеет место оценка

$$v_i \leq u_i \quad (19.6)$$

и

$$v_i \leq q^i u_0 + \sum_{k=0}^{i-1} q^{i-k-1} f_k. \quad (19.7)$$

Доказательство. Вычтем из неравенства (19.5) равенство (19.4). Получаем рекуррентность

$$v_{i+1} - u_{i+1} \leq q(v_i - u_i) \leq q^2(v_{i-1} - u_{i-1}) \leq \dots \leq q^{i+1}(u_0 - v_0) = 0,$$

и неравенство (19.6) доказано. Неравенство (19.7) следует из (19.3).

Разностные уравнения 2-го порядка с постоянными коэффициентами

Рассмотрим разностное уравнение

$$by_{i+1} - cy_i + dy_{i-1} = f_i \quad i = 1, 2, \dots, n, \quad (19.8)$$

коэффициенты которого b, c, d не зависят от i и $b \neq 0$, $d \neq 0$. В § 17 при изучении метода стрельбы мы выяснили, что для нахождения общего решения исходного уравнения необходимо найти общее решение однородного разностного уравнения

$$by_{i+1} - cy_i + dy_{i-1} = 0 \quad i = 1, 2, \dots, n, \quad (19.9)$$

к которому следует прибавить любое частное решение исходного уравнения. Займемся однородным уравнением (19.9). Последовательность уравнений (19.9) можно рассматривать как систему n уравнений относительно $n + 2$ неизвестных y_0, y_1, \dots, y_{n+1} . Нетрудно заметить, что ранг этой системы равен n , поскольку в матрице

$$A = \begin{pmatrix} b & -c & d & \dots & 0 & 0 & 0 \\ 0 & b & -c & d & \dots & 0 & 0 \\ 0 & 0 & b & -c & d & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & b & -c & d \end{pmatrix}$$

есть отличный от нуля минор размерности n . Следовательно, система имеет 2 линейно независимых решения — вектор $Y^{(1)} = (y_0^{(1)}, y_1^{(1)}, \dots, y_n^{(1)})^T$ и вектор $Y^{(2)} = (y_0^{(2)}, y_1^{(2)}, \dots, y_n^{(2)})^T$ и общее решение однородного разностного уравнения имеет следующий вид

$$y_i = c_1 y_i^{(1)} + c_2 y_i^{(2)}. \quad (19.10)$$

Эти линейно независимые решения можно отыскивать так, как изложено в методе стрельбы, но мы сейчас рассмотрим другой прием. Будем искать ненулевые решения (19.9) в виде

$$y_i = q^i, \quad q \neq 0. \quad (19.11)$$

Подставляя (19.10) в (19.9), получаем уравнение для определения q , которое будем называть характеристическим

$$bq^2 - cq + d = 0. \quad (19.12)$$

Как обычно, рассмотрим три случая:

1. Корни характеристического многочлена вещественны и различны: $q_1 \neq q_2$. В этом случае имеем два решения $y_i^{(1)} = q_1^i$ и $y_i^{(2)} = q_2^i$, причем легко видеть, что эти вектора линейно независимы, поскольку есть минор 2-го порядка, например

$$\begin{pmatrix} 1 & 1 \\ q_1 & q_2 \\ \cdot & \cdot \\ \cdot & \cdot \end{pmatrix},$$

отличный от нуля.

2. Корни характеристического многочлена кратные: $q_1 = q_2 = q$. В этом случае второе линейно независимое решение имеет вид $y_i^{(2)} = iq^i$. В самом деле, подставляя $y_i^{(2)}$ в (19.9) и группируя, получаем

$$b(i+1)q^{i+1} - cid^i + d(i-1)q^{i-1} = q^{i-1}[(bq^2 - cq + d)i + (bq^2 - d)].$$

Выражение в первой скобке равно нулю, поскольку q есть корень характеристического многочлена. Выражение во второй скобке также равно нулю, поскольку в случае кратного корня $q = \frac{c}{2b}$, а поскольку корень кратный, дискриминант $D = c^2 - 4bd$ равен нулю и

$$bq^2 - d = \frac{c^2}{4b} - d = \frac{D}{4b} = 0.$$

Линейная независимость двух этих решений $y_i^{(1)} = q^i$ и $y_i^{(2)} = iq^i$ также легко проверяется.

3. Корни характеристического многочлена комплексные:

$$q_{1,2} = \frac{c \pm i\sqrt{|D|}}{2b}, \quad D = c^2 - 4bd.$$

Используя представление комплексного числа в виде

$$q_{12} = \rho e^{\pm i\varphi} = \rho(\cos \varphi \pm i \sin \varphi), \quad \rho = \sqrt{\frac{d}{b}}, \quad \varphi = \operatorname{arctg} \frac{\sqrt{|D|}}{c},$$

замечаем, что

$$q_{12}^k = \rho^k (\cos k\varphi \pm i \sin k\varphi).$$

Нетрудно видеть, что если комплекснозначная сеточная функция является решением линейного однородного разностного уравнения с вещественными коэффициентами, то тогда вещественная и мнимая части этой сеточной функции также являются решением этого уравнения. Поэтому решением являются функции

$$y_k^{(1)} = \rho^k \cos k\varphi \quad \text{и} \quad y_k^{(2)} = \rho^k \sin k\varphi.$$

Для того чтобы убедиться, что эти решения линейно независимы, достаточно в прямоугольной матрице

$$\begin{pmatrix} 1 & 0 \\ \rho \cos \varphi & \rho \sin \varphi \\ \rho^k \cos k\varphi & \rho^k \sin k\varphi \end{pmatrix}$$

найти минор 2-го порядка, отличный от нуля. Это предлагается сделать самостоятельно в качестве легкого упражнения.

Пример 1. Найти общее решение уравнения

$$y_{k+1} - y_k + 2y_{k-1} = a^k.$$

Найдем общее решение однородного уравнения

$$y_{k+1} - y_k + 2y_{k-1} = 0.$$

Корни характеристического уравнения

$$q^2 - q - 2 = 0$$

равны $q_1 = -1$, $q_2 = 2$ и общее решение однородного уравнения

$$\bar{y}_k = c_1(-1)^k + c_2 2^k.$$

Для отыскания частного решения неоднородного уравнения предположим, что $a \neq -1$ и $a \neq 2$. В этом случае решение можно найти в виде

$y_k^{(0)} = Aa^k$. Подставляя предполагаемое частное решение в исходное уравнение и сокращая на a^{k-1} , получаем

$$A = \frac{a}{a^2 - a - 2}.$$

Если же $a = 2$, решение будем искать в виде $y_k^{(0)} = Ak2^k$. Подставляя в исходное уравнение, получаем

$$A[(k+1)4 - k2 - 2(k-1)] = 2 \quad A = 1/3.$$

Случай $a = -1$ рассматривается аналогично. Таким образом, общее решение исходного уравнения имеет вид

$$y_k = c_1(-1)^k + c_22^k + y_k^{(0)}.$$

В этом примере явно просматривается аналогия между методами отыскания общего решения для линейных дифференциальных и линейных разностных уравнений. Эта аналогия не случайна, поскольку разностные уравнения есть дискретный аналог дифференциальных уравнений. В случае линейного дифференциального уравнения с постоянными коэффициентами решение однородного уравнения ищется в виде $e^{\lambda t}$. Если же время дискретно, $t = kh$, то $q = e^{\lambda h}$.

Пример 2. Пусть x — произвольное вещественное число, такое, что $|x| \leq 1$. Рассмотрим уравнение

$$y_{n+1} - 2xy_n + y_{n-1} = 0. \quad (19.13)$$

Будем искать решение этого уравнения, удовлетворяющее условиям $y_0 = 1$, $y_1 = x$. Поскольку $|x| \leq 1$, для дальнейшего удобно обозначить $\cos \delta = |x|$, $\delta = \arccos x$. Характеристическое уравнение имеет вид

$$q^2 - 2 \cos \delta q + 1 = 0.$$

Корни этого уравнения

$$q_{12} = \cos \delta \pm i \sin \delta = e^{\pm i\delta}$$

и общее решение

$$y_n = c_1 \cos(n \arccos x) + c_2 \sin(n \arccos x).$$

Из условий $y_0 = 1$, $y_1 = x$ получаем $c_1 = 1$, $c_2 = 0$, и решение имеет вид

$$y_n = \cos(n \arccos x).$$

В § 6 было доказано, что многочлены Чебышева $T_n(x) = \cos(n \arccos x)$ удовлетворяют рекуррентному соотношению (6.2). Здесь мы заметили, что решением разностного уравнения (19.13), которое и есть рекуррентное соотношение (6.2), при условиях $y_0 = 1$, $y_1 = x$ является многочлен Чебышева. Тем самым доказана эквивалентность двух определений многочлена Чебышева, о чем говорилось в § 6.

§ 20. Собственные значения простейшего разностного оператора

Напомним, что собственными значениями линейного оператора A называются те значения λ , при которых существуют ненулевые решения уравнения

$$Ax = \lambda x,$$

а сами эти ненулевые решения называются собственными функциями, соответствующими собственному значению λ .

В этом параграфе мы рассмотрим задачу на собственные значения для разностного оператора

$$\frac{u_{k+1} - 2u_k + u_{k-1}}{h^2} = \lambda u_k \quad h = \frac{l}{n} \quad k = 1, 2, \dots, n-1 \quad (20.1)$$

$$u_0 = 0 \quad u_n = 0. \quad (20.2)$$

Эта задача является разностным аналогом задачи на собственные значения для дифференциального оператора

$$y''(x) = \lambda y(x) \quad 0 \leq x \leq l$$

$$y(0) = 0 \quad y(l) = 0.$$

Из курса дифференциальных уравнений известно, что собственные значения этого оператора

$$\lambda_m = -\left(\frac{m\pi}{l}\right)^2,$$

а система собственных функций $e_m(x) = \sqrt{\frac{2}{l}} \sin \frac{m\pi x}{l}$ образует полную ортонормированную систему в пространстве $L^2_{(0,l)}$, что позволяет применять метод разделения переменных при решении задач математической физики.

Займемся собственными значениями разностного оператора. Перепишем уравнение (20.1) в виде

$$u_{k+1} - 2cu_k + u_{k-1} = 0, \quad c(\lambda) = 1 + \frac{\lambda h^2}{2}. \quad (20.3)$$

При $|c(\lambda)| > 1$ корни характеристического уравнения

$$q^2 - 2cq + 1 = 0 \quad (20.4)$$

вещественны и различны и общее решение имеет вид

$$u_k = c_1 q_1^k + c_2 q_2^k.$$

Из левого граничного условия $u_0 = 0$ получаем

$$c_1 + c_2 = 0,$$

а из правого граничного условия $u_n = 0$ следует

$$c_1 q_1^n + c_2 q_2^n = 0.$$

Отсюда следует, что $c_1 = c_2 = 0$, т. е. ненулевые решения невозможны при $|c(\lambda)| > 1$. При $|c(\lambda)| = 1$ корни кратные и общее решение имеет вид

$$u_k = c_1 q^k + c_2 k q^k.$$

С учетом граничных условий получаем $c_1 = 0$, $c_2 = 0$, и в этом случае также возможно только тривиальное решение.

Пусть $|c(\lambda)| < 1$. В этом случае корни характеристического многочлена комплексные. Как и раньше, обозначим

$$c(\lambda) = \cos \alpha \quad \cos \alpha = 1 + \frac{\lambda h^2}{2}. \quad (20.5)$$

Тогда корни уравнения (20.4) равны

$$q_{12} = \cos \alpha \pm i \sin \alpha$$

и общее решение имеет вид

$$u_k = c_1 \cos k\alpha + c_2 \sin k\alpha.$$

Из левого граничного условия $u_0 = 0$ следует $c_1 = 0$, т. е.

$$u_k = c_2 \sin k\alpha.$$

Из второго граничного условия следует, что

$$c_2 \sin n\alpha = 0$$

и, следовательно, ненулевые решения возможны в случае, когда

$$\alpha_m = \frac{m\pi}{n}.$$

Из (20.5)

$$-\lambda h^2 = 2(1 - \cos \alpha) = 4 \sin^2 \frac{\alpha}{2}.$$

Следовательно,

$$\lambda_m = -\frac{4}{h^2} \sin^2 \frac{\alpha_m}{2}.$$

Заметим далее, что $|c(\lambda)| = 1$ при $m=0$ и при $m=n$ и, следовательно, $m = 1, 2, \dots, n-1$. Далее значения λ повторяются. Таким образом, доказана следующая теорема.

Теорема 1. Разностный оператор (20.1) - (20.2) имеет $n-1$ собственное значение

$$\lambda_m = -\frac{4^2}{h} \sin^2 \frac{m\pi}{2n} \quad m = 1, 2, \dots, n-1 \quad (20.6)$$

или

$$\lambda_m = -\frac{4^2}{h} \sin^2 \frac{m\pi h}{2l} \quad m = 1, 2, \dots, n-1, \quad (20.6')$$

которым соответствуют собственные функции

$$e_k^m = c \sin k \frac{m\pi}{n} \quad m = 1, 2, \dots, n-1 \quad (20.7)$$

или, поскольку $\frac{1}{n} = \frac{h}{l}$ и $kh = x_k$,

$$e_k^m = c \sin \frac{m\pi x_k}{l} \quad m = 1, 2, \dots, n-1. \quad (20.7')$$

Заметим далее, что $|\lambda_m| < |\lambda_{m+1}|$ в силу того, что функция

$$\frac{4}{h^2} \sin^2 \frac{\pi m h}{2l}$$

монотонно возрастает при $1 \leq m \leq n-1$. Наименьшее по модулю собственное значение равно

$$\lambda_1 = -\frac{4}{h^2} \sin^2 \frac{\pi h}{2l}.$$

Обозначив

$$\xi = \frac{\pi h}{2l},$$

модуль λ_1 можно записать с виде

$$|\lambda_1| = \frac{\pi^2 \sin^2 \xi}{l^2 \xi^2}.$$

Элементарно показывается, что функция

$$\frac{\sin \xi}{\xi}$$

и имеем минимум в окрестности $\xi = \frac{\pi}{4}$, из чего следует, что $\lambda_1 \geq \frac{8}{l^2}$. Относительно λ_{n-1} имеем

$$|\lambda_{n-1}| = \sin^2 \frac{\pi(n-1)}{2n} = \frac{4}{h^2} \left(1 - \cos \frac{\pi h}{2l}\right)^2.$$

Без доказательства отметим, что если в пространстве сеточных функций ввести скалярное произведение по формуле

$$(u_i, v_i) = \sum_{i=1}^{n-1} h u_i v_i, \quad (20.8)$$

то в пространстве с этим скалярным произведением собственные функции

$$e_k^m = c \sin k \frac{m\pi}{n}$$

ортогональны. Для того чтобы норма собственных функций в этом пространстве была равна единице, константу c полагают равной $\sqrt{\frac{2}{l}}$.

§ 21. Разностные методы решения кривых задач для уравнения теплопроводности

Пусть задан стержень длины l и пусть $u(x, t)$ — температура в точке x ($0 \leq x \leq l$) стержня в момент времени t . Процесс распространения тепла в одномерном стержне описывается следующим уравнением:

$$c\rho \frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(k \frac{\partial u}{\partial x} \right) + f(x, t). \quad (21.1)$$

Функция $f(x, t)$ связана с плотностью внешних источников тепла. Величины k — коэффициент теплопроводности, c — теплоемкость единицы длины, ρ — плотность в общем случае могут зависеть от x и t . В простейшем случае, когда они константы, уравнение теплопроводности имеет вид

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x, t). \quad (21.2)$$

После замены $x_1 = \frac{x}{l}$, $t_1 = \frac{a^2}{l^2}t$, получим уравнение (индексы у t и x мы опускаем)

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(x, t). \quad (21.3)$$

Задача Коши для уравнения теплопроводности ставится следующим образом. Требуется найти решение уравнения (21.3), удовлетворяющее начальным

$$u(x, 0) = u_0(x) \quad (21.4)$$

и граничным условиям

$$\begin{aligned} \alpha_1 u(0, t) + \beta_1 u'_x(0, t) &= \gamma_1(t), \\ \alpha_2 u(1, t) + \beta_2 u'_x(1, t) &= \gamma_2(t). \end{aligned} \quad (21.5)$$

Говорят, что задана первая краевая задача, если $\beta_i = 0$ $i = 1, 2$. Если же $\alpha_i = 0$ $i = 1, 2$, то краевую задачу называют второй краевой задачей. Если все эти коэффициенты отличны от нуля, то говорят, что задана третья краевая задача.

Напомним, что в курсе математической физики решение уравнения (21.2) для первой или второй краевой задачи ищется в виде ряда по собственным функциям однородной краевой задачи, т. е. решение, вообще говоря, может быть найдено лишь приближенно. Кроме того, для ряда краевых условий

задача нахождения собственных значений приводит к необходимости нахождения корней некоторого нелинейного уравнения. Таким образом, задача нахождения решения для уравнения теплопроводности методом разделения переменных в ряде случаев оказывается нереализуемой.

Ниже будут описаны сеточные (разностные) методы решения краевых задач для уравнений математической физики, которые мы подробно изучим для уравнения теплопроводности. Пусть требуется найти значение решения краевой задачи (21.3), (21.4), (21.5) при $t = T$, т. е. $u(x_i, T)$. Разобьем отрезок $[0, 1]$ точками x_i , $i = 0, 1, \dots, n$ с шагом $h = \frac{1}{n}$, а отрезок $[0, T]$ точками t_j , $j = 0, 1, \dots, m$ с шагом $\tau = \frac{1}{m}$.

Совокупность точек (x_i, t_j) будем называть сеткой, а множество точек при $t = t_k$, т. е. точки $(t_k, x_0), (t_k, x_1) \dots (t_k, x_n)$, будем называть слоем.

Разностная схема с весами для уравнения теплопроводности

При фиксированном t заменим производную по x в точке $x = x_i$ разностным выражением

$$\frac{\partial u}{\partial x} \approx \frac{u_{i+1} - u_i}{h},$$

а при $x = x_n$ — выражением

$$\frac{\partial u}{\partial x} \approx \frac{u_n - u_{n-1}}{h}.$$

Вторую производную по x при $x = x_i$ $i = 1, 2, \dots, n-1$ заменим выражением

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2}.$$

В результате вместо задачи (21.3), (21.4), (21.5) получаем систему дифференциально-разностных уравнений

$$\frac{dy_i(t)}{dt} = \frac{y_{i+1}(t) - 2y_i(t) + y_{i-1}(t)}{h^2} + f_i(t) \quad i = 1, 2, \dots, n-1 \quad (21.6)$$

с начальными

$$y_i(0) = u_0(x_i) \quad (21.7)$$

и граничными условиями

$$\alpha_1 y_0(t) + \beta_1 \frac{y_1(t) - y_0(t)}{h} = \gamma_1(t), \quad (21.8)$$

$$\alpha_2 y_n(t) + \beta_2 \frac{y_n(t) - y_{n-1}(t)}{h} = \gamma_2(t). \quad (21.9)$$

Введем обозначения

$$\Lambda y_{it} = \frac{y_{i+1}(t) - 2y_i(t) + y_{i-1}(t)}{h^2}.$$

В этих обозначениях (21.6) имеет вид

$$\frac{dy_i(t)}{dt} = \Lambda y_{it} + f_i(t).$$

Заменим производную по t разностным выражением

$$\frac{dy_i(t_j)}{dt} \approx \frac{y_i(t_{j+1}) - y_i(t_j)}{\tau} = \frac{y_i^{j+1} - y_i^j}{\tau}. \quad (21.10)$$

Здесь и ниже $y_i^j = y(x_i, t_j)$. Правую часть в (21.6) заменим на линейную комбинацию значений на j -м и $j+1$ -м слое, т. е. на значения правой части при $t = t_j$ и $t = t_{j+1}$

$$\frac{y_i^{j+1} - y_i^j}{\tau} = \sigma \Lambda y_i^{j+1} + (1 - \sigma) y_i^j + \varphi_i^j. \quad (21.11)$$

Здесь σ — параметр, в зависимости от значений которого получается конкретная разностная схема, а φ — некоторая правая часть, например f_i^j . Ниже мы увидим, что в зависимости от σ правая часть φ может принимать и иные значения.

Вместе с начальными

$$y_i^0 = u_0(x_i) \quad (21.12)$$

и граничными условиями

$$\alpha_1 y_0^j + \beta_1 \frac{y_1^j - y_0^j}{h} = \gamma_1^j, \quad (21.13)$$

$$\alpha_2 y_n^j + \beta_2 \frac{y_n^j - y_{n-1}^j}{h} = \gamma_2^j \quad (21.14)$$

получаем задачу, которую называют разностной схемой с весами.

Методы решения разностной схемы с весами

Наиболее часто применяются разностные схемы при $\sigma = 0$, $\sigma = \frac{1}{2}$ и $\sigma = 1$. При $\sigma = 0$ схема называется явной, при $\sigma = \frac{1}{2}$ — симметричной, или схемой Кранка — Никольсона, при $\sigma = 1$ — неявной схемой. Любая разностная схема с весами позволяет реализовать алгоритм определения решения на неизвестном слое с помощью решения, полученного на предыдущем слое.

Рассмотрим этот алгоритм для явной схемы. Она имеет вид

$$\frac{y_i^{j+1} - y_i^j}{\tau} = \frac{y_{i+1}^j - 2y_i^j + y_{i-1}^j}{h^2} + f_i^j \quad i = 1, 2, \dots, n-1.$$

Пусть $j = 0$. Решение y_i^0 на нулевом слое задано начальными условиями (21.12.). Но для любой точки ($i = 1, 2, \dots, n-1$) первого слоя имеем

$$y_i^1 = y_i^0 + \frac{\tau}{h^2}(y_{i+1}^0 - 2y_i^0 + y_{i-1}^0) + \tau f_i^0 \quad i = 1, 2, \dots, n-1. \quad (21.15)$$

Далее, зная y_1^1 , с помощью левого граничного условия (21.13) при $j = 1$ определим y_0^1 и аналогично определим y_n^1 с помощью правого граничного условия (21.14). Таким образом, найдены решения во всех точках первого слоя. Повторяя этот алгоритм, определим решения на втором слое и т. д.

Неявная схема. При $\sigma = 1$ разностное уравнение (21.11) имеет вид

$$\frac{y_i^{j+1} - y_i^j}{\tau} = \frac{y_{i+1}^{j+1} - 2y_i^{j+1} + y_{i-1}^{j+1}}{h^2} + f_i^{j+\frac{1}{2}} \quad i = 1, 2, \dots, n-1.$$

Ниже будет показано, что более точная аппроксимация получается при $\varphi_i = f_i^{j+\frac{1}{2}}$. Обозначим через $z_i = y_i^{j+1}$. В этих обозначениях после приведения подобных членов получаем

$$z_{i+1} - (2 + \frac{h^2}{\tau})z_i + z_{i-1} = -\frac{h^2}{\tau}y_i^j - f_i^{j+\frac{1}{2}}h^2 \quad i = 1, 2, \dots, n-1. \quad (21.16)$$

Правая часть в этом уравнении определена, поскольку y_i^j — значения на предыдущем слое. Вместе с граничными условиями

$$\alpha_1 z_0 + \beta_1 \frac{z_1 - z_0}{h} = \gamma_1^{j+1}, \quad (21.17)$$

$$\alpha_n z_0 + \beta_2 \frac{z_n - z_{n-1}}{h} = \gamma_2^{j+1} \quad (21.18)$$

приходим к разностной задаче, которая решается методами прогонки или стрельбы, которые изложены в §17.

Симметричная схема (схема Кранка — Никольсона). При $\sigma = \frac{1}{2}$ в (21.11) получаем

$$\frac{y_i^{j+1} - y_i^j}{\tau} = \frac{y_{i+1}^{j+1} - 2y_i^{j+1} + y_{i-1}^{j+1}}{2h^2} + \frac{y_{i+1}^j - 2y_i^j + y_{i-1}^j}{2h^2} + f_i^{j+\frac{1}{2}} \quad i = 1, \dots, n-1.$$

Пусть, как и в неявной схеме, $z_i = y_i^{j+1}$. Приводя подобные члены, приходим к уравнению

$$z_{i+1} - \left(2 + \frac{2h^2}{\tau}\right)z_i + z_{i-1} = \varphi_i \quad i = 1, 2, \dots, n-1, \quad (21.19)$$

где

$$\varphi_i = -\frac{2h^2}{\tau}y_i^j - 2f_i^{j+\frac{1}{2}}h^2 - (y_{i+1}^j - 2y_i^j + y_{i-1}^j). \quad (21.20)$$

Вспоминая о краевых условиях (21.17), (21.18), замечаем, что и эта разностная задача решается методами стрельбы или прогонки.

§22. Аппроксимация, устойчивость и сходимость разностной схемы с весами для уравнения теплопроводности

Напомним, что для того, чтобы решение разностной задачи сходилось к решению задачи (21.3) — (21.5), надо, чтобы разностная задача аппроксимировала исходную задачу с порядком $O(h^k, \tau^p)$, где h и τ — шаги сетки по x и t соответственно, и была устойчива по правой части и начальным условиям.

Порядок аппроксимации

Ниже для упрощения изложения мы будем рассматривать первую краевую задачу, т. е. будем считать, что краевые условия в (21.5) имеют вид:

$$u(0, t) = \bar{\gamma}_1(t),$$

$$u(1, t) = \bar{\gamma}_2(t).$$

Введем обозначения в безындексной форме

$$y = y_i^j, \quad \widehat{y} = y_i^{j+1}, \quad \Lambda y = \frac{y_{i+1} - 2y_i + y_{i-1}}{h^2},$$

$$y_t = \frac{y_i^{j+1} - y_i^j}{\tau}, \quad y^{(\sigma)} = \sigma y_i^{j+1} + (1 - \sigma) y_i^j.$$

В этих обозначениях разностное уравнение (21.11) имеет вид

$$y_t = \Lambda y^{(\sigma)} + \varphi. \quad (22.1)$$

Начальные и граничные условия аппроксимируются точно

$$y_i^0 = u_0(x_i), \quad y_0^j = \gamma_1(t_j), \quad y_n^j = \gamma_2(t_j). \quad (22.2)$$

Пусть $u = u_i^j$ — точное решение исходной задачи в узлах сетки и пусть $z = y - u$. Тогда, подставляя $y = z + u$ в (22.1), получаем разностную задачу для погрешности z

$$z_t = \Lambda z^{(\sigma)} + \psi, \quad z_i^0 = 0, \quad z_o^k = z_n^k = 0, \quad \text{для всех } k, \quad (22.3)$$

где

$$\psi = \Lambda u^{(\sigma)} + \varphi - u_t. \quad (22.4)$$

Заметим, что имеет место равенство

$$u^{(\sigma)} = \sigma \widehat{u} + (1 - \sigma) u = \frac{u + \widehat{u}}{2} + \left(\sigma - \frac{1}{2}\right) \tau u_t.$$

В самом деле,

$$\frac{u + \hat{u}}{2} + (\sigma - \frac{1}{2})\tau u_t = \frac{u + \hat{u}}{2} + (\sigma - \frac{1}{2})(\hat{u} - u) = \sigma \hat{u} + (1 - \sigma)u.$$

Наша цель — выяснить порядок аппроксимации исходной задачи разностной схемой с весами, а для этого необходимо получить оценку функции ψ через шаги сетки h и τ . С этой целью разложим ψ в ряд по степеням h и τ в окрестности точки $(x_i, \bar{t} = t_j + \frac{1}{2}\tau)$.

Пусть $\bar{v} = v(t_j + \frac{\tau}{2})$. Тогда

$$\hat{v} = v(t_j + \tau) = \bar{v} + \frac{1}{2}\tau \frac{\partial \bar{v}}{\partial t} + \frac{\tau^2}{8} \frac{\partial^2 \bar{v}}{\partial t^2} + \frac{\tau^3}{48} \frac{\partial^3 \bar{v}}{\partial t^3} + O(\tau^4),$$

$$v = v(t_j) = \bar{v} - \frac{1}{2}\tau \frac{\partial \bar{v}}{\partial t} + \frac{\tau^2}{8} \frac{\partial^2 \bar{v}}{\partial t^2} - \frac{\tau^3}{48} \frac{\partial^3 \bar{v}}{\partial t^3} + O(\tau^4)$$

и

$$\Lambda u = \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} = \frac{\partial^2 u}{\partial x^2} + O(h^4).$$

Следовательно,

$$\frac{u + \hat{u}}{2} = \bar{u} + O(\tau^2),$$

$$u_t = \frac{\hat{u} - u}{\tau} = \frac{\partial \bar{u}}{\partial t} + O(\tau^2).$$

$$\tau u_t = \hat{u} - u = \tau \frac{\partial \bar{u}}{\partial t} + O(\tau^3).$$

Поэтому

$$\psi = \Lambda u^{(\sigma)} + \varphi - u_t = \Lambda \left(\frac{u + \hat{u}}{2} + (\sigma - \frac{1}{2})\tau u_t \right) + \varphi - u_t =$$

$$= \left(\frac{\partial^2 \bar{u}}{\partial x^2} - \frac{\partial \bar{u}}{\partial t} \right) + \varphi + (\sigma - \frac{1}{2})\tau \frac{\partial^2}{\partial x^2} \left(\frac{\partial \bar{u}}{\partial t} \right) + O(h^2 + \tau^2).$$

Поскольку \bar{u} есть решение исходной задачи в точке $(x_i, t_j + 1/2)$, то

$$\left(\frac{\partial^2 \bar{u}}{\partial x^2} - \frac{\partial \bar{u}}{\partial t} \right) = -\bar{f}.$$

Отсюда видно, что если в качестве φ взять \bar{f} , то ϕ будет иметь следующий порядок:

$$\phi = O(\tau + h^2) \quad \text{при } \sigma = 0,$$

$$\begin{aligned}\phi &= O(\tau + h^2) \quad \text{при } \sigma = 1, \\ \phi &= O(\tau^2 + h^2) \quad \text{при } \sigma = 1/2.\end{aligned}$$

В заключение заметим, что в случае краевых условий 2-го типа порядок аппроксимации по h снижается на единицу, поскольку в граничной точке приходится аппроксимировать 1-ю производную по x конечно-разностным соотношением с погрешностью порядка h .

Устойчивость разностной схемы с весами

Мы будем исследовать на устойчивость первую краевую задачу с нулевыми граничными условиями для уравнения

$$\frac{y_i^{j+1} - y_i^j}{\tau} = \sigma \Lambda y_i^{j+1} + (1 - \sigma) \Lambda y_i^j + \varphi_i^j \quad i = 1, 2, \dots, n-1 \quad (22.5)$$

$$y_0^j = 0 \quad y_n^j = 0$$

и заданными начальными условиями

$$y_i^0 = u_0(x_i). \quad (22.6)$$

Решение этой задачи можно представить в виде

$$y_i = v_i + w_i,$$

где v_i — решение однородного уравнения с исходными краевыми условиями

$$\frac{v_i^{j+1} - v_i^j}{\tau} = \sigma \Lambda v_i^{j+1} + (1 - \sigma) \Lambda v_i^j \quad i = 1, 2, \dots, n-1 \quad (22.7)$$

$$v_0^j = 0 \quad v_n^j = 0$$

и заданными начальными условиями

$$v_i^0 = u_0(x_i), \quad (22.8)$$

а w_i — решение неоднородного уравнения с исходными граничными

$$\frac{w_i^{j+1} - w_i^j}{\tau} = \sigma \Lambda w_i^{j+1} + (1 - \sigma) \Lambda w_i^j + \varphi_i \quad i = 1, 2, \dots, n-1 \quad (22.9)$$

$$w_0^j = 0 \quad w_n^j = 0$$

и нулевыми начальными условиями

$$w_i^0 = 0. \quad (22.10)$$

Будем говорить, что разностная схема с весами устойчива, если для y_i имеет место оценка

$$\|y_i\| \leq M_1 \|u_0\| + M_2 \|\varphi_i\|.$$

Очевидно, что для получения этой оценки необходимо получить оценки для v_i и w_i :

$$\|v_i\| \leq M_1 \|u_0(x_i)\| \quad \text{и} \quad \|w_i\| \leq M_2 \|\varphi_i\|.$$

Если имеет место оценка для v_i , то говорят об устойчивости по начальным данным, а если имеет место оценка для w_i , то говорят, что разностная схема устойчива по правой части. Если M_1 и M_2 не зависят от шагов сетки h и τ , то разностную схему называют абсолютно устойчивой. Если же эти оценки выполняются при определенном соотношении между h и τ , то говорят, что разностная схема условно устойчива.

Для исследования на устойчивость по начальным данным разложим $v_i^{(j)}$ — значение сеточной функции на j -м слое по собственным функциям сеточного оператора

$$\Lambda v_i \equiv \frac{v_{i+1} - 2v_i + v_{i-1}}{h^2} \quad v_0 = v_n = 0.$$

В §20 было показано, что собственными значениями этого разностного оператора являются

$$\lambda_m = -\frac{4}{h^2} \sin^2 \frac{\pi m h}{2} \quad m = 1, 2, \dots, n-1,$$

которым соответствуют собственные функции

$$e_i^m = \sqrt{2} \sin i \frac{m\pi}{n}.$$

Эти собственные функции ортогональны в сеточном скалярном произведении

$$(u, v) = \sum_{i=1}^{n-1} u_i v_i h$$

и нормированы на единицу в пространстве с нормой, порожденной этим скалярным произведением. Пусть

$$v_i^{(j)} = \sum_{k=1}^{n-1} b_k^{(j)} e_i^k. \quad (22.11)$$

Коэффициенты этого разложения, естественно, зависят от слоя j , что и отражает верхний индекс. Разложение (22.11) подставим в уравнение (22.7). Приравнявая коэффициенты при одинаковых e_i^k , получаем

$$\frac{b_k^{(j+1)} - b_k^{(j)}}{\tau} = \sigma \lambda_k b_k^{(j+1)} + (1 - \sigma) \lambda_k b_k^{(j)}.$$

Или

$$b_k^{j+1} = q_k b_k^j,$$

где

$$q_k = \frac{1 + (1 - \sigma) \lambda_k \tau}{1 - \sigma \tau \lambda_k}.$$

Таким образом,

$$b_k^{(j+1)} = q_k^{j+1} b_k^{(0)}$$

и

$$v_i^{(j)} = \sum_{k=1}^{n-1} q_k^j b_k^{(0)} e_i^k.$$

Но

$$\varphi_i = v_i^{(0)} = \sum_{k=1}^{n-1} b_k^{(0)} e_i^k.$$

В силу равенства Парсеваля

$$\|v_i^j\|^2 = \sum_{k=1}^{n-1} (q_k^j b_k^{(0)})^2.$$

Пусть $q = \max_k |q_k|$. Тогда

$$\|v_i^j\|^2 \leq q^{2j} \sum_{k=1}^{n-1} (b_k^{(0)})^2 = q^{2j} \|\varphi_i\|^2.$$

Отсюда ясно, что для того, чтобы получить ограниченную оценку $v_i^{(j)}$ для любого слоя j , необходимо, чтобы выполнялось неравенство $|q_k| \leq 1$. Выясним, при каких σ выполняется условие $|q| \leq 1$ или $-1 \leq q_k \leq 1$. Неравенство $q_k \leq 1$ означает, что

$$\frac{1 + (1 - \sigma) \lambda_k \tau}{1 - \sigma \tau \lambda_k} \leq 1$$

или

$$\frac{\lambda_k \tau}{1 - \sigma \tau \lambda_k} \leq 0.$$

Напомним, что $\tau > 0$, так как это шаг сетки, а $\lambda_k < 0$, что следует из результатов §20. Поэтому условие $q_k \leq 1$ означает, что

$$1 - \sigma \tau \lambda_k > 0,$$

т. е.

$$\sigma > \frac{1}{\tau \lambda_k}.$$

Таким образом, неравенство $q_k \leq 1$ выполняется при любом неотрицательном σ . Рассмотрим неравенство $q_k \geq -1$ или

$$q + 1 = \frac{2 - (2\sigma - 1)\tau \lambda_k}{1 - \sigma \tau \lambda_k} \geq 0.$$

При $\sigma \geq 0$ знаменатель дроби положителен и для выполнения неравенства необходимо, чтобы

$$2 - (2\sigma - 1)\tau \lambda_k \geq 0$$

или

$$\sigma \geq \frac{1}{2} + \frac{1}{\tau \lambda_k}.$$

Поскольку

$$\lambda_k = -\frac{4}{h^2} \sin^2 \frac{k\pi h}{2},$$

то

$$\sigma \geq \frac{1}{2} - \frac{h^2}{4\tau \sin^2 \frac{\pi k}{2}} \geq \frac{1}{2} - \frac{h^2}{4\tau}$$

или

$$\sigma \geq \frac{1}{2} - \frac{h^2}{4\tau} = \sigma_0. \quad (22.12)$$

Полученное соотношение позволяет сделать вывод об устойчивости по начальным данным рассмотренных выше разностных схем.

1. Явная схема $\sigma = 0$. Неравенство (22.12) выполняется, если шаги сетки удовлетворяют соотношению

$$\frac{\tau}{h^2} \leq 0.5. \quad (22.13)$$

Таким образом, явная схема условно устойчива.

Ее решение сходится к решению исходной задачи лишь при выполнении условия (22.13).

2. Схема Кранка — Никольсона $\sigma = \frac{1}{2}$. Поскольку h и τ положительны, соотношение (22.12) выполняется всегда. Схема Кранка — Никольсона безусловно устойчива, и ее решение сходится к решению исходной задачи со скоростью $O(\tau^2 + h^2)$.

3. Неявная схема $\sigma = 1$ безусловно устойчива. Решение неявной схемы сходится к решению сходной задачи со скоростью $O(\tau + h^2)$.

Исследование на устойчивость по правой части проводится по той же методике. Можно показать, что условие (22.12) является достаточным и для устойчивости по правой части. Более подробно эта тема изложена в [11].

Пусть $u(x, t)$ — положение точки x однородной струны в момент времени t . Тогда при малых отклонениях от положения равновесия положение этой точки x описывается уравнением

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + f(x, t) \quad 0 < x < l, \quad t > 0,$$

или после соответствующих замен

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + f(x, t) \quad 0 < x < 1, \quad 0 < t < T. \quad (23.1)$$

При этом задана начальная форма струны и начальное распределение скоростей

$$u(x, 0) = v(x), \quad \frac{\partial u(x, 0)}{\partial t} = w(x). \quad (23.2)$$

Граничные условия имеют вид

$$u(0, t) = \mu_1(t), \quad u(1, t) = \mu_2(t). \quad (23.3)$$

Введем в области $\bar{D} = (0 \leq x \leq 1, 0 \leq t \leq T)$ сетку с узлами (x_i, t_j) $x = ih, \quad h = \frac{1}{n}, \quad t_j = j\tau, \quad \tau = \frac{1}{m}$ и рассмотрим задачу нахождения решения для разностной схемы с весами

$$\frac{y_i^{j+1} - 2y_i^j + y_i^{j-1}}{\tau^2} = \Lambda(\sigma y_i^{j+1} + (1 - 2\sigma)y_i^j + \sigma y_i^{j-1}) + f_i^j, \quad (23.4)$$

причем

$$y_0^j = \mu_1(t_j), \quad y_n^j = \mu_2(t_j), \quad y_i^0 = v(x_i), \quad \frac{y_i^1 - y_i^0}{\tau} = \tilde{w}_i. \quad (23.5)$$

Напомним, что

$$\Lambda z_i = \frac{z_{i+1} - 2z_i + z_{i-1}}{h^2}.$$

Краевые и первое начальное условие аппроксимируются точно, а функцию \tilde{w} выберем так, чтобы погрешность аппроксимации производной по t разностным выражением была минимальной. Поскольку

$$u(x_i, \tau) = u(x_i, 0) + \tau \dot{u}(x_i, 0) + 0.5\tau^2 \ddot{u}(x_i, 0) + O(\tau^3)$$

из уравнения (23.1)

$$\ddot{u}(x, 0) = u''(x, 0) + f(x, 0),$$

поэтому

$$u(x_i, \tau) = u(x_i, 0) + \tau \dot{u}(x_i, 0) + 0.5\tau^2(u''(x_i, 0) + f(x_i, 0)) + O(\tau^3)$$

и

$$\frac{u(x_i, \tau) - u(x_i, 0)}{\tau} = w(x_i) + 0.5\tau^2(v''(x_i) + f(x_i, 0)) + O(\tau^2).$$

Следовательно, если

$$\tilde{w}_i = w(x_i) + 0.5\tau^2(v''(x_i) + f(x_i, 0)),$$

то погрешность аппроксимации второго начального условия разностным выражением будет порядка $O(h^2)$.

Таким образом, задаче (23.1) — (23.3) поставлена в соответствие разностная задача (23.4) — (23.5), которая решается, как и разностная задача для уравнения теплопроводности, методами прогонки или стрельбы. Применяя те же методы, что и в предыдущем параграфе, можно показать, что разностная задача (23.4) — (23.5) аппроксимирует исходную задачу (23.1) — (23.3) с порядком $O(\tau^2 + h^2)$.

Исследование на устойчивость разностной схемы (23.4) — (23.5) проводится так же, как и исследование на устойчивость разностной схемы для уравнения теплопроводности. Исследование на устойчивость по начальным данным приводит к следующим результатам.

При $\sigma = 0$ (явная схема) для устойчивости по начальным данным необходимо, чтобы шаги сетки удовлетворяли соотношению

$$\frac{\tau^2}{h^2} \leq \frac{1}{1 + \varepsilon}, \quad \varepsilon > 0,$$

т. е. явная схема условно устойчива.

При

$$\sigma \geq \frac{1 + \varepsilon}{4} - \frac{h^2}{4\tau^2} \quad (23.6)$$

разностная схема безусловно устойчива.

Условие (23.6) позволяет для 1-й краевой задачи доказать устойчивость по правой части при $\sigma \geq 0$.

§ 24. Разностные схемы для уравнения Пуассона

Рассмотрим задачу Дирихле уравнения Пуассона

$$\Delta u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = -f(x, y). \quad (24.1)$$

Требуется найти решение этого уравнения внутри прямоугольника

$$\bar{G} = G \cup \Gamma = \{0 \leq x \leq l_1, \quad 0 \leq y \leq l_2\},$$

принимającego на границе Γ заданные значения:

$$u|_{\Gamma} = \mu(s). \quad (24.2)$$

Для численного решения задачи (24.1) — (24.2) разобьем отрезки $[0, l_1]$ и $[0, l_2]$ на n_1 и n_2 частей соответственно с шагами $h_1 = 1/n_1$ и $h_2 = 1/n_2$, определив тем самым сетку ϖ с границей γ . Граничными узлами будем считать узлы $(0, y_j)$, (l_1, y_j) , $(x_i, 0)$, (x_i, l_2) , $i = 0, 1, \dots, n_1$, $j = 0, 1, \dots, n_2$, т. е. узлы на сторонах прямоугольника. Обозначим через v_{ij} сеточную функцию, заданную на сетке ϖ . Аппроксимируем вторые производные выражением

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h_1^2} + O(h_1^2),$$

$$\frac{\partial^2 u}{\partial y^2} \approx \frac{u_{i,j+1} - 2u_{ij} + u_{i,j-1}}{h_2^2} + O(h_2^2).$$

Пользуясь этими выражениями, заменим исходное уравнение (24.1) разностным уравнением

$$\frac{v_{i+1,j} - 2v_{ij} + v_{i-1,j}}{h_1^2} + \frac{v_{i,j+1} - 2v_{ij} + v_{i,j-1}}{h_2^2} = -f_{ij}. \quad (24.3)$$

С учетом граничных условий

$$v_{i0} = \mu(x_i, 0), \quad v_{in_2} = \mu(x_i, l_2), \quad v_{0j} = \mu(0, y_j), \quad v_{n_1j} = \mu(l_1, y_j) \quad (24.4)$$

получаем разностную задачу (24.3) — (24.4). Мы уже оценивали аппроксимацию 2-й производной разностным выражением. Поскольку граничные условия аппроксимируются точно, легко можно получить, что при $f = \varphi$ разностная задача (24.3) — (24.4) аппроксимирует исходную задачу (24.1) — (24.2) с порядком $O(h_1^2 + h_2^2)$.

Если сетка квадратная, $h_1 = h_2$, то уравнение (24.3) можно переписать в виде

$$v_{ij} = \frac{v_{i-1,j} + v_{i+1,j} + v_{i,j-1} + v_{i,j+1} + h^2 f_{ij}}{4}.$$

Отсюда видно, что в случае однородного уравнения значение в каждой точке есть среднее арифметическое от значений в соседних точках.

Среди методов решения задачи (24.3) – (24.4) разделяют прямые методы и итерационные методы. Прямые методы применяются, когда легко можно найти собственные значения и собственные функции некоторой разностной задачи, что возможно, если исходное уравнение есть уравнение с постоянными коэффициентами и область G есть прямоугольник. Если это не так, то применяют итерационные методы, поскольку задача (24.3) – (24.4) есть линейная система $(n_1 - 1)(n_2 - 1)$ уравнений относительно $(n_1 - 1)(n_2 - 1)$ неизвестных v_{ij} $i = 1, 2, \dots, n_1 - 1$, $j = 1, 2, \dots, n_2 - 1$.

Прямые методы. Метод разделения переменных

Пусть w_{ij} — сеточная функция, заданная на сетке ϖ и равная нулю на границе сетки, а вне границы удовлетворяющая разностному уравнению

$$-\left(\frac{w_{i+1,j} - 2w_{ij} + w_{i-1,j}}{h_1^2} + \frac{w_{i,j+1} - 2w_{ij} + w_{i,j-1}}{h_2^2}\right) = \varphi_{ij}, \quad (24.5)$$

где

$$\varphi = f + \frac{\mu(l_1, y)}{h_1^2} \quad \text{при} \quad x = l_1 - h_1, \quad 0 < y < l_2$$

$$\varphi = f + \frac{\mu(0, y)}{h_1^2} \quad \text{при} \quad x = h_1, \quad 0 < y < l_2$$

$$\varphi = f + \frac{\mu(x, l_2)}{h_2^2} \quad \text{при} \quad y = l_2 - h_2, \quad 0 < x < l_1$$

$$\varphi = f + \frac{\mu(x, 0)}{h_2^2} \quad \text{при} \quad y = h_2, \quad 0 < x < l_1$$

и

$$\varphi = f$$

во всех остальных точках сетки. Пусть

$$\Lambda_x w_{ij} = \frac{w_{i+1,j} - 2w_{ij} + w_{i-1,j}}{h_1^2},$$

$$\Lambda_y w_{ij} = \frac{w_{i,j+1} - 2w_{ij} + w_{i,j-1}}{h_2^2}.$$

В этих обозначениях уравнение (24.5) можно переписать в виде

$$-\Lambda_x w_{ij} - \Lambda_y w_{ij} = \varphi_{ij}, \quad (24.6)$$

$$w|_\gamma = 0. \quad (24.7)$$

Решение задачи (24.5) — (24.6) при фиксированном i будем искать в виде разложения решения по собственным функциям оператора

$$\Lambda_y e_j = \lambda e_j \quad e_0 = e_n = 0.$$

В §20 для собственных значений и собственных функций этого оператора получены формулы

$$\lambda_k = -\frac{4}{h_2^2} \sin^2 \frac{k\pi h_2}{2l_2}, \quad (24.8)$$

$$e_j^{(k)} = \sqrt{\frac{2}{l_2}} \sin \frac{k\pi y_j}{l_2} \quad k = 1, 2, \dots, n_2 - 1. \quad (24.9)$$

Пусть

$$w_{ij} = \sum_{k=1}^{n_2-1} \alpha_i^{(k)} e_j^{(k)}. \quad (24.10)$$

Разложим функцию φ_{ij} при фиксированном i по собственным функциям оператора Λ_y :

$$\varphi_{ij} = \sum_{k=1}^{n_2-1} \beta_i^k e_j^{(k)}. \quad (24.11)$$

Как следует из рассмотрений §20, собственные функции $e_j^{(k)}$ ортогональны в пространстве со скалярным произведением

$$(v_i, w_i) = \sum_{k=1}^{n_2-1} h_2 v_i^{(m)} w_i^{(n)}$$

и нормированы на единицу в пространстве с нормой, порожденной этим скалярным произведением. Это позволяет найти коэффициенты β_i^k .

$$\beta_i^k = \sum_{j=1}^{n_2-1} h_2 \varphi_{ij} e_j^{(k)} = \sqrt{\frac{2}{l_2}} \sum_{j=1}^{n_2-1} h_2 \sin \frac{k\pi j h_2}{l_2} \varphi_{ij}. \quad (24.12)$$

Подставим эти разложения в (24.6)

$$-\sum_{k_2=1}^{n_2-1} e_j^{(k)} \Lambda_x \alpha_i^{(k)} - \sum_{k_2=1}^{n_2-1} \lambda_k e_j^{(k)} \alpha_i^{(k)} = \sum_{k_2=1}^{n_2-1} \beta_i^{k_2} e_j^{(k_2)}.$$

Подставим эти разложения в (24.6)

$$-\sum_{k_2=1}^{n_2-1} e_j^{(k)} \Lambda_x \alpha_i^{(k)} - \sum_{k_2=1}^{n_2-1} \lambda_k e_j^{(k)} \alpha_i^{(k)} = \sum_{k_2=1}^{n_2-1} \beta_i^k e_j^{(k)}.$$

Или

$$\sum_{k_2=1}^{n_2-1} e_j^{(k)} \left(\frac{\alpha_{i+1}^{(k)} - 2\alpha_i^{(k)} + \alpha_{i-1}^{(k)}}{h_1^2} + \lambda_k \alpha_i^{(k)} + \beta_i^k \right) = 0.$$

Следовательно,

$$\frac{\alpha_{i+1}^{(k)} - 2\alpha_i^{(k)} + \alpha_{i-1}^{(k)}}{h_1^2} + \lambda_k \alpha_i^{(k)} + \beta_i^k = 0. \quad (24.13)$$

При этом

$$\alpha_0^{(k)} = \alpha_{n_1}^{(k)} = 0.$$

Если ввести обозначение $z_i = \alpha_i^{(k)}$, то для z_i получаем краевую задачу для разностного уравнения 2-го порядка

$$z_{i+1} - (2 - \lambda_k h_1^2) z_i + z_{i-1} + h_1^2 \beta_i^k = 0 \quad i = 1, 2, \dots, n_1 - 1 \quad (24.14)$$

$$z_0 = z_{n_1} = 0, \quad (24.15)$$

которая решается методами прогонки или стрельбы, изложенными в § 17.

В заключение заметим, что в случае, если область не прямоугольная или уравнение с переменными коэффициентами, то изложенный метод обычно неприменим. В этом случае для решения сеточной задачи применяются итерационные методы решения линейных систем, которые изложены в § 3. Сходимость итерационных процессов обусловлена тем, что сеточный оператор для уравнения Пуассона является симметричным и положительно определенным в сеточном скалярном произведении. Более подробно эти вопросы изложены в [9].

Лабораторная работа № 1

Эта работа посвящена отысканию максимального по модулю собственного значения матрицы A . Для нахождения максимального по модулю собственного значения необходимо задать любой ненулевой вектор $X^{(0)}$ и запустить итерационный процесс (2.1). Далее, в случае сходимости последовательностей (2.4') или (2.5') или (2.7) — (2.8) вычислить максимальное по модулю собственное значение и соответствующий ему собственный вектор, нормированный на единицу в 1-й векторной норме.

Задание 1

$$\begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & -1 \\ 1 & -2 & 2 \end{pmatrix}$$

Задание 6

$$\begin{pmatrix} -3 & 4 & -2 \\ 1 & 0 & 1 \\ 6 & -6 & 5 \end{pmatrix}$$

Задание 2

$$\begin{pmatrix} 2.1 & 1 & 1.1 \\ 1 & 2.6 & 1.1 \\ 1.1 & 1.1 & 3.1 \end{pmatrix}$$

Задание 7

$$\begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \\ -1 & 2 & 3 \end{pmatrix}$$

Задание 3

$$\begin{pmatrix} 1 & -2 & -1 \\ -1 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$$

Задание 8

$$\begin{pmatrix} -3 & 2 & 2 \\ -3 & -1 & 1 \\ -1 & 2 & 0 \end{pmatrix}$$

Задание 4

$$\begin{pmatrix} -3 & 4 & -2 \\ 1 & 0 & 1 \\ 6 & -6 & 5 \end{pmatrix}$$

Задание 9

$$\begin{pmatrix} 3 & -3 & 1 \\ 3 & -2 & 2 \\ -1 & 2 & 0 \end{pmatrix}$$

Задание 5

$$\begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}$$

Задание 10

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

Задание 11

$$\begin{pmatrix} 2 & 1 & 1 \\ -2 & 0 & -1 \\ 2 & 1 & 2 \end{pmatrix}$$

Задание 12

$$\begin{pmatrix} 2 & -1 & -1 \\ 1 & 0 & -1 \\ 3 & -1 & -2 \end{pmatrix}$$

Задание 13

$$\begin{pmatrix} 1 & -2 & 2 \\ 1 & 4 & -2 \\ 1 & 5 & - \end{pmatrix}$$

Задание 14

$$\begin{pmatrix} 1 & -2 & -1 \\ -1 & 1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$$

Задание 15

$$\begin{pmatrix} 2.7 & 1 & 1.7 \\ 1 & 3.2 & 1.7 \\ 1.7 & 1.7 & 3.7 \end{pmatrix}$$

Задание 16

$$\begin{pmatrix} 1.4 & 1.2 & -1.3 \\ 1.2 & 0.9 & 0.5 \\ -1.3 & 0.4 & 0.8 \end{pmatrix}$$

Задание 17

$$\begin{pmatrix} 2.9 & 1 & 1.8 \\ 1 & 3.3 & 1.8 \\ 1.8 & 1.8 & 3.8 \end{pmatrix}$$

Задание 18

$$\begin{pmatrix} 2.4 & 1.2 & -0.3 \\ 1.2 & 1.9 & 1.4 \\ -0.3 & 1.4 & 0.8 \end{pmatrix}$$

Задание 19

$$\begin{pmatrix} 1.6 & 1.2 & -1.1 \\ 1.2 & 1.1 & 0.6 \\ -1.1 & 0.6 & 0.8 \end{pmatrix}$$

Задание 20

$$\begin{pmatrix} 3.3 & 1 & 2.3 \\ 1 & 3.8 & 2.3 \\ 2.3 & 2.3 & 4.3 \end{pmatrix}$$

Задание 21

$$\begin{pmatrix} 2.9 & 1 & 1.9 \\ 1 & 3.4 & 1.9 \\ 1.9 & 1.9 & 3.9 \end{pmatrix}$$

Задание 22

$$\begin{pmatrix} 2.6 & 1.2 & -0.1 \\ 1.2 & 2.1 & 1.6 \\ -0.1 & 1. & 0.8 \end{pmatrix}$$

Задание 23

$$\begin{pmatrix} 1.8 & 1.2 & -0.9 \\ 1.2 & 1.3 & 0.8 \\ -0.9 & 0.8 & 0.8 \end{pmatrix}$$

Задание 24

$$\begin{pmatrix} 3.4 & 1 & 2.4 \\ 1 & 3.9 & 2.4 \\ 2.4 & 2.4 & 4.9 \end{pmatrix}$$

Лабораторная работа № 2.

Лабораторная работа посвящена итерационным методам решения линейных систем. Для нахождения решения студенту предлагается в зависимости от варианта реализовать итерационный процесс по методу простой итерации, либо по методу Зейделя, либо по методу скорейшего спуска. В качестве начального приближения к решению можно взять любой вектор $X^{(0)}$. Итерационный процесс прекращается при выполнении условия

$$\frac{\|X^{(k+1)} - X^{(k)}\|_1}{\|X^{(k)}\|_1} \leq 0.01.$$

На печать выводится результат — вектор $X^{(k+1)}$ и вектор невязки

$$R^{(k+1)} = AX^{(k+1)} - b.$$

Кроме того, при сдаче работы необходимо владеть материалом, изложенным в §3.

Метод простой итерации

Задание 1.1

$$\begin{aligned} 3.6x_1 + x_2 + 2.4x_3 &= 0.26 \\ x_1 + 3.9x_2 + 2.4x_3 &= 0.31 \\ 2.4x_1 + 2.4x_2 + 4.9x_3 &= 0.76 \end{aligned}$$

Задание 1.3

$$\begin{aligned} 3.5x_1 + x_2 + 2.1x_3 &= 0.56 \\ x_1 + 4x_2 + 2.5x_3 &= 0.61 \\ 2.1x_1 + 2.5x_2 + 4.7x_3 &= 0.96 \end{aligned}$$

Задание 1.2

$$\begin{aligned} 3.1x_1 + x_2 + 1.6x_3 &= 2.16 \\ x_1 + 3.1x_2 + 1.6x_3 &= 2.21 \\ 1.6x_1 + 1.6x_2 + 3.9x_3 &= 1.46 \end{aligned}$$

Задание 1.4

$$\begin{aligned} 2x_1 + 0.9x_2 + x_3 &= 2.2 \\ 0.9x_1 + 1.8x_2 + 0.3x_3 &= 2.31 \\ x_1 + 0.3x_2 + 1.9x_3 &= 2.06 \end{aligned}$$

Задание 1.5

$$1.3x_1 + 0.4x_2 + 0.5x_3 = 0.25$$

$$0.4x_1 + 1.3x_2 + 0.3x_3 = 0.51$$

$$0.5x_1 + 0.3x_2 + 1.3x_3 = 1.44$$

Задание 1.10

$$3.1x_1 + x_2 + 1.8x_3 = 3.2$$

$$x_1 + 3.3x_2 + 1.8x_3 = 2.1$$

$$1.8x_1 + 1.8x_2 + 3.9x_3 = 1.6$$

Задание 1.6

$$1.6x_1 + 0.7x_2 + 0.8x_3 = 1.2$$

$$0.7x_1 + 2.6x_2 + 0.3x_3 = 1.1$$

$$0.8x_1 + 0.3x_2 + 1.6x_3 = 2.3$$

Задание 1.11

$$3.6x_1 + x_2 + 2.3x_3 = 1.26$$

$$x_1 + 3.8x_2 + 2.3x_3 = 0.11$$

$$2.3x_1 + 2.3x_2 + 4.7x_3 = 1.76$$

Задание 1.7

$$2.6x_1 + x_2 + 1.2x_3 = 0.86$$

$$x_1 - 2.7x_2 + 1.2x_3 = 2.2$$

$$1.2x_1 + 1.2x_2 + 3.2x_3 = 1.44$$

Задание 1.12

$$3x_1 + x_2 + 1.9x_3 = 3.2$$

$$x_1 + 3.4x_2 + 1.9x_3 = 2.2$$

$$1.9x_1 + 1.9x_2 + 4x_3 = 0.6$$

Задание 1.8

$$3x_1 + x_2 + 1.5x_3 = 3.2$$

$$x_1 + 3.6x_2 - 1.5x_3 = 1$$

$$1.5x_1 - 1.5x_2 + 3.5x_3 = 2.6$$

Задание 1.13

$$5.6x_1 + x_2 + 2.3x_3 = 1.26$$

$$x_1 + 7.8x_2 + 2.3x_3 = 0.11$$

$$2.3x_1 + 2.3x_2 + 6.7x_3 = 2.76$$

Задание 1.9

$$1.4x_1 + 0.5x_2 + 0.6x_3 = 0.66$$

$$0.5x_1 + 1.4x_2 + 0.3x_3 = 0.37$$

$$0.6x_1 + 0.3x_2 + 1.4x_3 = -0.26$$

Задание 1.14

$$4x_1 + x_2 + 1.9x_3 = 3.2$$

$$x_1 + 4x_2 + 1.9x_3 = 5.2$$

$$1.9x_1 + 1.9x_2 + 6x_3 = 3.6$$

Задание 2.1

$$\begin{aligned}3.1x_1 + x_2 + 1.6x_3 &= 2.16 \\x_1 + 3.1x_2 + 1.6x_3 &= 2.21 \\1.6x_1 + 1.6x_2 + 3.9x_3 &= 1.46\end{aligned}$$

Задание 2.2

$$\begin{aligned}2x_1 + 0.9x_2 + x_3 &= 2.2 \\0.9x_1 + 1.8x_2 + 0.3x_3 &= 2.31 \\x_1 + 0.3x_2 + 1.9x_3 &= 2.06\end{aligned}$$

Задание 2.3

$$\begin{aligned}1.6x_1 + 0.7x_2 + 0.8x_3 &= 1.2 \\0.7x_1 + 2.6x_2 + 0.3x_3 &= 1.1 \\0.8x_1 + 0.3x_2 + 1.6x_3 &= 2.3\end{aligned}$$

Задание 2.4

$$\begin{aligned}3x_1 + x_2 + 1.5x_3 &= 3.2 \\x_1 + 3.6x_2 - 1.5x_3 &= 1 \\1.5x_1 - 1.5x_2 + 3.5x_3 &= 2.6\end{aligned}$$

Задание 2.5

$$\begin{aligned}3.1x_1 + x_2 + 1.8x_3 &= 3.2 \\x_1 + 3.3x_2 + 1.8x_3 &= 2.1 \\1.8x_1 + 1.8x_2 + 3.9x_3 &= 1.6\end{aligned}$$

Задание 2.6

$$\begin{aligned}3x_1 + x_2 + 1.9x_3 &= 3.2 \\x_1 + 3.4x_2 + 1.9x_3 &= 2.2 \\1.9x_1 + 1.9x_2 + 4x_3 &= 0.6\end{aligned}$$

Задание 2.7

$$\begin{aligned}4x_1 + x_2 + 1.9x_3 &= 3.2 \\x_1 + 4x_2 + 1.9x_3 &= 5.2 \\1.9x_1 + 1.9x_2 + 6x_3 &= 3.6\end{aligned}$$

Задание 2.8

$$\begin{aligned}3.5x_1 + x_2 + 2.4x_3 &= 1.2 \\x_1 + 4x_2 + 2.5x_3 &= 2.2 \\2.4x_1 + 2.5x_2 + 4.5x_3 &= 3.1\end{aligned}$$

Задание 2.9

$$\begin{aligned}3x_1 + x_2 + x_3 &= 0.2 \\x_1 + 3x_2 &= 2.2 \\x_1 + 3x_3 &= 3.6\end{aligned}$$

Задание 2.10

$$\begin{aligned}2x_1 + 0.9x_2 + x_3 &= -0.2 \\0.9x_1 + 2x_2 + 0.3x_3 &= 1.2 \\x_1 + 0.3x_2 + 2x_3 &= 2.6\end{aligned}$$

Задание 2.11

$$\begin{aligned}4.9x_1 + 0.7x_2 + x_3 &= -5.2 \\0.7x_1 + 5.1x_2 + x_3 &= 2.2 \\x_1 + x_2 + 6.1x_3 &= 3.6\end{aligned}$$

Задание 2.12

$$\begin{aligned}5.1x_1 + 2x_2 + x_3 &= 8.2 \\2x_1 + 6x_2 + 3x_3 &= 5.1 \\x_1 + 3x_2 + 5x_3 &= 3.6\end{aligned}$$

Метод скорейшего спуска

Задание 3.1

$$5x_1 + x_2 + 1.4x_3 = 1.2$$

$$x_1 + 4x_2 + 2.5x_3 = 2.2$$

$$1.4x_1 + 2.5x_2 + 4.5x_3 = 3.1$$

Задание 3.2

$$3x_1 + x_2 + x_3 = 0.2$$

$$x_1 + 3x_2 = 2.2$$

$$x_1 + 3x_3 = 3.6$$

Задание 3.3

$$4.9x_1 + 0.7x_2 + x_3 = -5.2$$

$$0.7x_1 + 5.1x_2 + x_3 = 2.2$$

$$x_1 + x_2 + 6.1x_3 = 3.6$$

Задание 3.4

$$4x_1 + x_2 + 1.9x_3 = 3.2$$

$$x_1 + 4x_2 + 1.9x_3 = 5.2$$

$$1.9x_1 + 1.9x_2 + 6x_3 = 3.6$$

Задание 3.5

$$2x_1 + 0.9x_2 + x_3 = -0.2$$

$$0.9x_1 + 2x_2 + 0.3x_3 = 1.2$$

$$x_1 + 0.3x_2 + 2x_3 = 2.6$$

Задание 3.6

$$5.1x_1 + 2x_2 + x_3 = 8.2$$

$$2x_1 + 6x_2 + 3x_3 = 5.1$$

$$x_1 + 3x_2 + 5x_3 = 3.6$$

Задание 3.7

$$x_1 + x_2 = 8.2$$

$$x_1 + 2x_2 + 2x_3 = 5.1$$

$$2x_2 + 5x_3 = 3.6$$

Задание 3.8

$$2x_1 + 2x_2 - 2x_3 = 10.2$$

$$2x_1 + 5x_2 - 4x_3 = -1.1$$

$$-2x_1 - 4x_2 + 5x_3 = 3$$

Задание 3.9

$$5x_1 - 2x_2 - 2x_3 = 1.2$$

$$-2x_1 + 6x_2 = 1.1$$

$$-2x_1 + 4x_3 = 0.9$$

Задание 3.10

$$6x_1 + 0.6x_2 + 0.7x_3 = 5.1$$

$$0.6x_1 + 3x_2 - 0.3x_3 = 5.1$$

$$0.7x_1 - 0.3x_2 + 4x_3 = 2.9$$

Лабораторная работа № 3

Лабораторная работа посвящена итерационным методам отыскания решения нелинейных уравнений или систем вида

$$f_1(x_1, x_2, \dots, x_n) = 0$$

.

$$f_n(x_1, x_2, \dots, x_n) = 0.$$

Или в векторной форме

$$F(x) = 0.$$

Для запуска итерационного процесса нулевое приближение к решению студент определяет самостоятельно, используя стандартные методы исследования функций, графические или иные соображения. Итерационный процесс прекращается при выполнении условия

$$\frac{\|X^{(k+1)} - X^{(k)}\|_1}{\|X^{(k)}\|_1} \leq 0.01.$$

На печать выводится результат — вектор $X^{(k+1)}$ и вектор невязки

$$R^{(k+1)} = F(X^{(k+1)}).$$

Кроме того, при сдаче работы необходимо владеть материалом, изложенным в §4.

Методы нахождения решения для одного уравнения

1. Используя метод секущих (§ 4, формула (4,4)), найти решение уравнения:

1.1 $x - \sin x = 0.25$

1.2 $\sqrt{x} - \cos 0.41x = 0$

1.3 $3x - \cos x - 1 = 0$

1.4 $x \ln x - 1.5 = 0$

1.5 $x + \ln x = 0.5$

1.6 $x^3 - \sin x = 0.25$

1.7 $\operatorname{tg} 0.58x + 0.1 = x^2$

1.8 $x^3 - 6x - 8 = 0$

1.9 $x^2 + 4 \sin x = 0$

1.10 $x^3 + 4x - 6 = 0$

1.11 $\operatorname{ctgx} = 0.5x$

1.12 $x^3 + 3x^2 + 12x + 3 = 0$

2. Используя метод Ньютона (§4, формула (4.5)), найти решение уравнения:

$$2.1. x^3 - 3x^2 + 9x - 8 = 0$$

$$2.2. \sqrt{x} - \cos(0.387x) = 0$$

$$2.3. 3x - \cos x - 1 = 0$$

$$2.4. x^2 - \cos^2 \pi x - 1 = 0$$

$$2.5. x^2 + 4 \sin x = 0$$

$$2.6. (x - 1) + 0.5e^x = 0$$

$$2.7. 5x^3 + 2x^2 - 15x - 6 = 0$$

$$2.8. x^4 - 4x^3 + 6x^2 - 4x + 1 = 0$$

$$2.9. tg(0.5x + 0.1) = x^2$$

$$2.10. 5x - e^x = 0$$

Итерационные методы для нелинейных систем

3. Используя метод Ньютона (§4, формула (4.8)), найти решение системы:

$$3.1. \begin{cases} \sin(x + 1) - y = 1 \\ 2x + \cos y = 2 \end{cases}$$

$$3.2. \begin{cases} \cos(x - 1) + y = 0.5 \\ x - \cos y = 3 \end{cases}$$

$$3.3. \begin{cases} \sin x + 2y = 2 \\ 2x + \cos(y - 1) + x = 0.7 \end{cases}$$

$$3.4. \begin{cases} \cos x + y = 1.5 \\ x - \cos y = 3 \end{cases}$$

$$3.5. \begin{cases} \sin(x + 0.5) - y = 1 \\ \cos y - 2 + x = 0 \end{cases}$$

$$3.6. \begin{cases} \cos(y - 1) + x = 0.8 \\ y - \cos x = 2 \end{cases}$$

$$3.7. \begin{cases} \sin(x + y) - 1.1x = 0.1 \\ x^2 + y^2 = 1 \end{cases}$$

$$3.8. \begin{cases} \frac{x^2}{9} + \frac{y^2}{16} = 1 \\ x - \cos y = 1 \end{cases}$$

$$3.9. \begin{cases} \sin(x + 1) - y = 1 \\ 2x + \cos y = 2 \end{cases}$$

$$3.10. \begin{cases} \cos(x + 1) + y = 1 \\ e^x - y = 0 \end{cases}$$

$$3.11. \begin{cases} \sin(x + 1) - y = 0 \\ x^2 + y^2 = 4 \end{cases}$$

$$3.12. \begin{cases} \cos(x - 1) + y = 0.5 \\ x - \cos y = 3 \end{cases}$$

4. Используя метод скорейшего спуска (на выбор: либо формулы (4.14) — (4.15), либо формулы (4.16) — (4.17) §4)) найти решение системы:

$$4.1. \quad \begin{aligned} y &= e^x - 1 \\ 4x^2 + y &= 4 \end{aligned}$$

$$4.2. \quad \begin{aligned} y &= \sin x + 1 \\ y &= x^2 - 1 \end{aligned}$$

$$4.3. \quad \begin{aligned} \sin(x+1) - y &= 1 \\ 2x + \cos y &= 2 \end{aligned} \quad 4.4. \quad \begin{aligned} \cos(x-1) + y &= 0.5 \\ x - \cos y &= 3 \end{aligned}$$

$$4.5. \quad \begin{aligned} \sin x - y &= 0 \\ x^2 + y^2 &= 4 \end{aligned} \quad 4.6. \quad \begin{aligned} x^2 + y^2 &= 16 \\ x - \cos y &= 3 \end{aligned}$$

$$4.7. \quad \begin{aligned} \frac{x^2}{16} + \frac{y^2}{25} &= 1 \\ (x-3)^2 + y^2 &= 4 \end{aligned} \quad 4.8. \quad \begin{aligned} y &= 1.5 + x^3 \\ x^2 + \frac{y^2}{9} &= 1 \end{aligned}$$

$$4.9. \quad \begin{aligned} \sin(x+1) - y &= 1 \\ x^2 + y^2 &= 9 \end{aligned} \quad 4.10. \quad \begin{aligned} \cos(x-1) + y &= 0.50 \\ x - \cos y &= 3 \end{aligned}$$

$$4.11. \quad \begin{aligned} y &= x^2 \\ x &= e^{-y} \end{aligned} \quad 4.12. \quad \begin{aligned} \frac{x^2}{4} + \frac{y^2}{9} &= 1 \\ (x-1)^2 + \frac{y^2}{4} &= 1 \end{aligned}$$

$$4.13. \quad \begin{aligned} x - y^2 &= 0 \\ y - \sin(x-1) &= 1 \end{aligned} \quad 4.14. \quad \begin{aligned} \frac{x^2}{4} + \frac{y^2}{9} &= 1 \\ x - 1 &= e^y \end{aligned}$$

Лабораторная работа № 4

В работе предлагается вычислить значения функции

$$F(t) = \int_a^b f(t, x) dx \quad c \leq t \leq d$$

в точках $t_j = c + j\tau$, $\tau = \frac{d-c}{m}$ двумя способами:

а) методом удвоения числа шагов для достижения заданной точности, используя стандартные квадратурные формулы;

б) используя составные квадратурные формулы Гаусса (11.8) с тремя узлами при фиксированном $n = 4$.

Пусть

$$g_j(x) = f(t_j, x).$$

Тогда задача сводится к вычислению в цикле по j интеграла от функции $g_j(x)$ по заданным квадратурным формулам.

Метод удвоения числа шагов состоит в следующем. Все квадратурные формулы имеют вид

$$\int_a^b g(x) dx \approx \sum_{i=1}^N c_i g(x_i) = S_N.$$

Разобьем отрезок интегрирования точками x_1, \dots, x_N и составим сумму S_N . Удвоим N и составим сумму S_{2N} . Если величина

$$|S_{2N} - S_N| < \varepsilon,$$

то считается, что заданная точность достигнута. В противном случае процесс продолжается.

Варианты заданий

Квадратурные формулы

Обозначим через $h = \frac{b-a}{N}$, $g_i = g(x_i)$, $g_{i+\frac{1}{2}} = g(x_i + \frac{h}{2})$, $g_0 = g(a)$, $g_N = g(b)$. Начальное разбиение N выбирается произвольно, но при использовании формул Симпсона начальное N четно.

А. Квадратурные формулы прямоугольников

$$S_N(g) = h \sum_{i=0}^{N-1} g_{i+\frac{1}{2}}.$$

Б. Квадратурные формулы трапеций

$$S_N(g) = h\left(\frac{g_0 + g_n}{2} + g_1 + g_2 + \dots + g_{N-1}\right).$$

В. Квадратурные формулы Симпсона (N - четно)

$$S_N(g) = \frac{h}{3}(g_0 + g_n + 2g_1 + 4g_2 + 2g_3 + 4g_4 + \dots + 2g_{N-1}).$$

Функция $f(x, t)$.

1. $f(x, t) = \sin\left(\frac{t}{1+x^2} + 0.001x\right)$
2. $f(x, t) = \cos\left(\frac{t}{1+x^2} + 0.001x\right)$
3. $f(x, t) = e^{-\left(\frac{t}{1+x^2} + 0.001x\right)}$
4. $f(x, t) = e^{\left(\frac{\sqrt{t}}{1+x^2}\right)}$
5. $f(x, t) = \ln\left(\frac{1+t}{1+x^2} + 0.001x\right)$
6. $f(x, t) = \sqrt{\frac{t}{6+x^2}} + x$
7. $f(x, t) = \frac{\sin(t+x^2+0.5)}{2+\cos(x^2+1)}$
8. $f(x, t) = \frac{\cos(t+x^2+0.5)}{2+\sin(x^2+1)}$
9. $f(x, t) = \frac{\cos(t+x)}{2x^2+\cos(x^2+1)}$
10. $f(x, t) = e^{\left(\frac{\sqrt{t+1}}{1+x^2}\right)} \sin 2x$

Значения a, b, c, d, m . Во всех вариантах $\varepsilon = 0.001$.

Номер	a	b	c	d	m
1	-1	0	0.5	1.5	20
2	0	2	0	1	20
3	0	1	2	3	25
4	1	2	1.5	2.5	25
5	-1	2	2	3	20
6	0.5	1.5	0	1.5	25
7	-2	2	2	3	25

Результаты следует оформить в виде таблицы с выводом значений функции $F(t_i)$, вычисленной методом удвоения числа шагов и значением N , при котором достигнута заданная точность, значениями, полученными с использованием квадратур Гаусса.

При сдаче работы необходимо знать оценки погрешности квадратурных формул и методы их получения, методы построения квадратурных формул Гаусса, свойства ортогональных многочленов.

Лабораторная работа № 5

В работе предлагается численно исследовать поведение решений системы дифференциальных уравнений в окрестности особых точек и сделать вывод о характере особой точки.

Пусть задана система

$$\begin{aligned}\frac{dx}{dt} &= P(x, y), \\ \frac{dy}{dt} &= Q(x, y).\end{aligned}$$

Напомним, что особой точкой (x_0, y_0) называется такая точка, в которой

$$P(x_0, y_0) = 0, \quad Q(x_0, y_0) = 0.$$

Необходимо найти все особые точки системы, численно исследовать поведение траекторий в окрестности особой точки и сделать вывод о том, является ли данная точка устойчивым положением равновесия и к какому типу (узел, седло, фокус, центр) относится эта особая точка. Результат численного исследования должен быть согласован с выводами, полученными аналитическим путем. Напомним, что аналитическое исследование проводится по следующей схеме.

1. Система линеаризуется в окрестности исследуемого положения равновесия.

Это означает, что в системе делается замена

$$u = x - x_0, \quad v = y - y_0$$

и в полученной системе отбрасываются все нелинейные члены.

2. В полученной системе вида

$$\begin{aligned}\frac{du}{dt} &= au + bv, \\ \frac{dv}{dt} &= cu + dv\end{aligned}$$

находятся собственные значения матрицы

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Лабораторная работа № 6

В лабораторной работе предлагается найти приближенное и точное решение дифференциального уравнения

$$y'' + py' + qy = f(x) \quad a \leq x \leq b,$$

удовлетворяющего граничным условиям:

$$\alpha_1 y(a) + \beta_1 y'(a) = \gamma_1,$$

$$\alpha_2 y(b) + \beta_2 y'(b) = \gamma_2,$$

на сетке $x_i = a + ih$, $i = 0, 1, \dots, n$, $h = \frac{(b-a)}{n}$, найдя точное решение в узлах сетки и построив и решив сеточную краевую задачу.

Построение точного решения задачи проводится элементарно, с использованием стандартных приемов нахождения решений линейных дифференциальных уравнений. Напомним, что общее решение дифференциального уравнения второго порядка имеет вид:

$$Y = C_1 y_1 + C_2 y_2 + y_0,$$

где y_1 и y_2 — два линейно независимых решения однородного уравнения, а y_0 — частное решение неоднородного уравнения. Два граничных условия, если определитель

$$\begin{vmatrix} \alpha_1 y_1(a) + \beta_1 y_1'(a), & \alpha_1 y_2(a) + \beta_1 y_2'(a) \\ \alpha_2 y_1(a) + \beta_2 y_1'(a), & \alpha_2 y_2(a) + \beta_2 y_2'(a) \end{vmatrix}$$

отличен от нуля, дают возможность определить константы C_1 и C_2 в формуле общего решения. Таким образом можно определить конкретный вид решения, и его следует вычислить в точках x_i .

Построение сеточной краевой задачи проводят по схеме, изложенной в §16. При сдаче работы необходимо вывести на экран точное и приближенное решение в точках x_i . Студент должен знать методы решения краевых задач для линейных разностных уравнений 2-го порядка.

Варианты заданий

I. Методы решения краевых задач

1. Метод прямой прогонки.
2. Метод обратной прогонки.
4. Метод встречной прогонки.
5. Метод стрельбы.

II. Уравнения и краевые условия

$$\begin{aligned} 1. \quad & y'' + 2y' + y = -2x \\ & y'(0) = 0 \quad y(1) + 2y'(1) = 0 \quad n = 20 \end{aligned}$$

$$\begin{aligned} 2. \quad & y'' + y = -2e^{-2x} \\ & y(1) + y'(1) = 0 \quad y'(2) = 1 \quad n = 20 \end{aligned}$$

$$\begin{aligned} 3. \quad & y'' + 4y' + 3y = \cos x \\ & y(1) = 1 \quad y'(2) = 1 \quad n = 20 \end{aligned}$$

$$\begin{aligned} 4. \quad & y'' + y = -2x \\ & y(0) + y'(0) = 1 \quad y'(\frac{\pi}{2}) = 0 \quad n = 20 \end{aligned}$$

$$\begin{aligned} 5. \quad & y'' + 4y' + 3y = -2x \\ & y'(1) = 0 \quad y'(2) - y(2) = 1 \quad n = 20 \end{aligned}$$

$$\begin{aligned} 6. \quad & y'' - 3y' + 2y = \sin(x) \\ & y(0) + 2y'(0) = -1 \quad y'(1) = 1 \quad n = 20 \end{aligned}$$

$$\begin{aligned} 7. \quad & y'' + y' - 2y = -2 \\ & y(-1) - y'(-1) = 0 \quad y(1) = 1 \quad n = 25 \end{aligned}$$

$$\begin{aligned} 8. \quad & y'' + 4y' + 5y = -2x \\ & y(0) = 1 \quad y'(1) = 1 \quad n = 25 \end{aligned}$$

$$9. y'' + 2y' - 3y = e^{4x}$$

$$y(0) = 0 \quad y'(1) = 1 \quad n = 20$$

$$10. y'' - y = -2e^{-2x}$$

$$y(0) = 0 \quad y(2) = 1 \quad n = 20$$

$$11. y'' - 5y' + 4y = -2e^{2x}$$

$$y'(0) = 0 \quad y(1) = 1 \quad n = 20$$

$$12. y'' - 3y' - 4y = -2e^{2x}$$

$$y'(0) = 0 \quad y'(1) = 1 \quad n = 20$$

$$13. y'' + 2y' - 3y = e^x$$

$$y(0) = 0 \quad y'(1) = 1 \quad n = 20$$

$$14. y'' - 4y' + 8y = -2$$

$$y'(0) = 1 \quad y(1) = 0 \quad n = 20$$

$$15. y'' - 9y = -2e^{3x}$$

$$y(0) = 0 \quad y(1) = 1 \quad n = 20$$

$$16. y'' + y = -2\sin(x)$$

$$y(0) = 0 \quad y'(\frac{\pi}{2}) = 1 \quad n = 25$$

$$17. y'' - 5y' = -2x^2$$

$$y'(0) = 1 \quad y(1) = 1 \quad n = 20$$

$$18. y'' - 5y' + 4y = 1$$

$$y'(1) = 0 \quad y'(2) = 0 \quad n = 20$$

Лабораторная работа № 7

Лабораторная работа посвящена изучению разностных методов решения краевых задач для уравнения теплопроводности. Пусть задано уравнение

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f(x, t) \quad 0 \leq x \leq 1$$

при начальных условиях

$$u(x, 0) = \phi(x)$$

и граничных условиях

$$\alpha_1 u(0, t) + \beta_1 \frac{\partial u(0, t)}{\partial x} = \gamma_1(t),$$

$$\alpha_2 u(1, t) + \beta_2 \frac{\partial u(1, t)}{\partial x} = \gamma_2(t).$$

Требуется:

– построить сеточную краевую задачу в соответствии с предложенной разностной схемой,

– вычислить при $t = t_0$ в точках x_i , $i = 0, 1, \dots, n$ приближенное решение, полученное решением построенной сеточной краевой задачи. Для сравнения результатов задается точное решение краевой задачи. Способы построения разностных схем и методы их решения изложены в §21.

При сдаче лабораторной работы студент должен быть знаком с основными понятиями теории разностных схем (аппроксимация, устойчивость, сходимость), уметь строить разностную схему для конкретной краевой задачи.

Варианты заданий

I. Типы разностных схем

1. Неявная схема.
2. Схема Кранка — Никольсона.

II. Методы решения сеточной задачи

1. Метод прямой прогонки.
2. Метод обратной прогонки.
3. Метод стрельбы.

Задание 8

$$\begin{aligned} \Delta u &= 2(y-2) & (u &= x^2(y-2)) \\ u|_{x=0} &= 0 \quad u|_{x=1} = y-2 & 0 \leq x \leq 1 \quad h_1 &= 0.1 \\ u|_{y=0} &= -2x^2 \quad u|_{y=2} = 0 & 0 \leq y \leq 2 \quad h_2 &= 0.1 \end{aligned}$$

Задание 9

$$\begin{aligned} \Delta u &= 12xy & (u &= x^3y + xy^3) \\ u|_{x=0} &= 0 \quad u|_{x=1} = y^3 + y & 0 \leq x \leq 1 \quad h_1 &= 0.1 \\ u|_{y=0} &= 0 \quad u|_{y=1} = x^3 + x & 0 \leq y \leq 1 \quad h_2 &= 0.1 \end{aligned}$$

Задание 10

$$\begin{aligned} \Delta u &= 2(x^2 + y^2 - 2x - y) & (u &= (x^2 - 2x)(y^2 - y)) \\ u|_{x=0} &= 0 \quad u|_{x=2} = 0 & 0 \leq x \leq 2 \quad h_1 &= 0.1 \\ u|_{y=0} &= 0 \quad u|_{y=1} = 0 & 0 \leq y \leq 1 \quad h_2 &= 0.1 \end{aligned}$$

Задание 11

$$\begin{aligned} \Delta u &= (2 - 6x)y & (u &= (x^2 - x^3)y) \\ u|_{x=0} &= 0 \quad u|_{x=1} = 0 & 0 \leq x \leq 1 \quad h_1 &= 0.1 \\ u|_{y=0} &= 0 \quad u|_{y=1} = x^2 - x^3 & 0 \leq y \leq 1 \quad h_2 &= 0.1 \end{aligned}$$

Задание 12

$$\begin{aligned} \Delta u &= (6y - 2)x & (u &= (y^3 - y^2)x) \\ u|_{x=0} &= 0 \quad u|_{x=1} = y^3 - y^2 & 0 \leq x \leq 1 \quad h_1 &= 0.1 \\ u|_{y=0} &= 0 \quad u|_{y=1} = 0 & 0 \leq y \leq 1 \quad h_2 &= 0.1 \end{aligned}$$

Задание 13

$$\begin{aligned} \Delta u &= 4 & (u &= (x^2 + y^2)) \\ u|_{x=0} &= y^2 \quad u|_{x=2} = 4 + y^2 & 0 \leq x \leq 2 \quad h_1 &= 0.1 \\ u|_{y=0} &= x^2 \quad u|_{y=1} = 1 + x^2 & 0 \leq y \leq 1 \quad h_2 &= 0.1 \end{aligned}$$

Литература

1. Бабенко, К.И. Основы численного анализа / К.И. Бабенко. — М.: Наука, 1986. — 774 с.
2. Бахвалов, Н.С. Численные методы. — 2-е изд. / Н.С.Бахвалов. — М.: Наука, 1975. — 632 с.
3. Березин, И.С. Методы вычислений. — Т. 1/ И.С. Березин, Н.П. Жидков. — М.: Наука, 1966. — 464 с.
4. Березин, И.С. Методы вычислений. — Т. 2/ И.С. Березин, Н.П. Жидков. — М.: Физматгиз, 1962. — 640 с.
5. Воеводин, В.В. Численные методы алгебры/ В.В. Воеводин. — М.: Наука, 1966. — 208 с.
6. Годунов, С.К. Разностные схемы. — 2-е изд. / С.К.Годунов, В.С. Рябенский. — М.: Наука, 1977. — 440 с.
7. Демидович, Б.П. Основы вычислительной математики. — 4-е изд./ Б.П. Демидович, И.А. Марон. — М.: Наука, 1970. — 664 с.
8. Крылов, В.И. Вычислительные методы высшей математики. — Т. 1/ В.И. Крылов, В.В. Бобкова, П.И. Монастырский. — Минск: Высшая школа, 1972. — 284 с.
9. Марчук, Г.И. Методы вычислительной математики. — 3-е изд./ Г.И. Марчук. — М.: Наука, 1989. — 608 с.
10. Самарский, А.А. Введение в численные методы. — 2-е изд./ А.А. Самарский. — М.: Наука, 1987. — 288 с.
11. Самарский, А.А. Теория разностных схем. — 2-е изд./ А.А. Самарский. — М.: Наука, 1983. — 600 с.
12. Тихомиров, В.В. Некоторые вопросы теории приближений/ В.В. Тихомиров. — М.: Изд-во МГУ, 1976. — 98 с.
13. Фадеев, Д.К. Вычислительные методы линейной алгебры. — 2-е изд./ Д.К. Фадеев, В.П. Фадеева. — М.: Физматгиз, 1963. — 734 с.

Учебное издание

Матвеев Владимир Николаевич

Методы вычислений

Учебное пособие

Редактор, корректор В.Н. Чулкова
Компьютерная верстка В.Н. Матвеева

Подписано в печать 27.04.2007 г. Формат 60x84/8.
Бумага тип. Усл. печ. л. 18,13. Уч.-изд. л. 9,0.
Тираж 100 экз. Заказ

Оригинал-макет подготовлен
в редакционно-издательском отделе ЯрГУ.

Ярославский государственный университет.
150000 Ярославль, ул. Советская, 14.

Отпечатано
ООО «Ремдер» ЛР ИД № 06151 от 26.10.2001.
г. Ярославль, пр. Октября, 94, оф. 37
тел. (4852) 73-35-03, 58-03-48, факс 58-03-49.