

Министерство образования Российской Федерации

**РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ
ГИДРОМЕТЕОРОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ**

**УЧЕБНО-МЕТОДИЧЕСКОЕ ПОСОБИЕ
ДЛЯ ВЫПОЛНЕНИЯ КОНТРОЛЬНОЙ РАБОТЫ**

по дисциплине

«МАТЕМАТИКА»

Раздел «Вычислительная математика»

Специальности: 012600 – метеорология
012700 – гидрология
012800 – океанология

САНКТ-ПЕТЕРБУРГ
2001

Одобрено Ученым советом РГГМУ

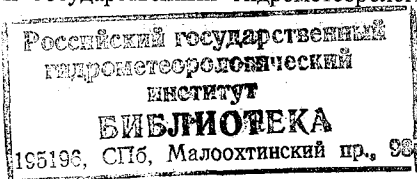
УДК 519.6

Учебно-методическое пособие для выполнения контрольной работы по дисциплине “Математика”. Раздел “Вычислительная математика” – СПб: изд. РГГМУ, 2000 – 56 с.

Приведены темы дисциплины “Математика” раздела “Вычислительная математика”, рекомендованная литература, основные теоретические сведения и примеры решения типичных задач.

Составитель: В. Н. Веретенников, канд. техн. наук, доц. РГГМУ

© Российский государственный гидрометеорологический университет (РГГМУ), 2000.



Уч. 1902

ПРЕДИСЛОВИЕ

“Вычислительная математика” как раздел дисциплины “Математика” является не только мощным средством решения прикладных гидрометеорологических задач, но и важнейшей компонентой интеллектуального развития. Именно в рамках математического образования студент получает навыки творческого подхода к решению интеллектуальных проблем, точному пониманию средств возможностей решения проблем, знакомится с современными информационными технологиями.

Внедрение компьютеров во все сферы человеческой деятельности требует от специалистов гидрометеорологического профиля овладения навыками использования вычислительной техники. Повышается уровень подготовки студентов, которые уже с первых курсов приобщаются к использованию ПК и простейших численных методов, не говоря уже о том, что при выполнении курсовых и дипломных проектов применение ПК становится нормой.

Основной дисциплиной, непосредственно связанной с вычислительной техникой, является вычислительная математика. Она изучает методы построения и исследования численных методов решения математических задач, которые моделируют различные гидрометеорологические процессы.

Главной задачей раздела “Вычислительная математика” дисциплины “Математика” является понимание основных понятий и идей численного анализа, особенностей и областей его применения.

Изучение вычислительной математики преследует следующие цели:

- усвоение и закрепление основных алгоритмов, понятий и определений вычислительной математики;
- практическое решение типичных задач вычислительной математики, требующих небольшого объема вычислений;

- решение достаточно сложных в вычислительном отношении задач, требующих для их численной реализации использование ПК.

В результате изучения дисциплины студент овладевает знаниями теории основных вычислительных алгоритмов, умением реально убедиться в действительных возможностях и свойствах на примере численного решения типичных модельных и прикладных задач.

В пособии излагаются основные теоретические сведения по численным методам решения задач алгебры, анализа, обыкновенных дифференциальных уравнений и уравнений математической физики. Изложение вычислительной математики не может не опираться на материал из математического анализа и линейной алгебры в объеме программы дисциплины "Математика". Впрочем, справочные сведения по линейной алгебре, в порядке напоминания, приводятся в теме 2.

Студент должен выполнить одну контрольную работу. Изучение дисциплины завершается зачетом.

ЛИТЕРАТУРА

1. Турчак А. И. Основы численных методов. – М.: Наука, 1987.–320 с.
2. Косарев В. И. 12 лекций по вычислительной математике. – М.: Изд-во МФТИ, 1995.–176 с.
3. Рябенкий В. С. Введение в вычислительную математику. – М.: Физматлит, 1994.–336 с.
4. Бахвалов Н. С., Жидков Н. П., Кобельков Г. М. Численные методы. – М.: Наука, 1987.–600 с.
5. Самарский А. А., Гулин А. В. Численные методы. – М.: Наука, 1989.–432 с.
6. Амосов А. А., Дубинский Ю. А., Копченова Н. В. Вычислительные методы для инженеров. – М.: Высшая школа, 1994.–544 с.
7. Азаров А. И. и др. Сборник задач по методам вычислений. – М.: Физматлит, 1994.–320 с.
8. Копченова Н. В., Марон И. А. Вычислительная математика в примерах и задачах. – М.: Наука, 1972.–369 с.
9. Веретенников В. Н. Программа дисциплины «МАТЕМАТИКА» для высших учебных заведений. – СПб.: изд. РГТМУ, 2000. – 28 с.

Тема 1. Элементы теории погрешностей

Литература: [1] гл. 1; [4] гл. 1; [5] ч. I, п. 2; [6] гл. 2; [7] гл. 1; [8] гл. 1.

Основные теоретические сведения

① V1. Под **погрешностью** понимается некоторая величина, характеризующая точность результата. Существует ^② три вида погрешностей:

- ✓ • **неустраняемая погрешность**, возникающая из-за неточности исходной информации, например, неточности измерений;
- ✓ • **погрешность метода**;
- ✓ • **погрешность вычислений**, возникающая из-за округлений.

2. Пусть a – точное (вообще говоря, неизвестное) значение некоторой величины, a^* – известное приближенное значение той же величины (**приближенное число**).

Ошибкой (или **погрешностью**) приближенного числа a^* называется разность $a - a^*$ между точным и приближенным значениями.

Количественной простейшей мерой ошибки является **абсолютная погрешность**

$$\Delta a^* = |a - a^*|. \quad (1.1)$$

Показателем качества вычисления служит **относительная погрешность**

$$\delta a^* = \frac{|a - a^*|}{|a|} = \frac{\Delta a^*}{|a|}, \quad (1.2)$$

Чем меньше относительная погрешность, тем выше точность вычислений. Так как значение a неизвестно, то непосредственное вычисление величин Δa^* и δa^* по формулам (1.1), (1.2) невозможно.

Более реальная и часто поддающаяся решению задача состоит в получении оценок погрешности вида

$$|a - a^*| \leq \bar{\Delta} a^*. \quad (1.3)$$

$$\frac{|a - a^*|}{|a|} \leq \bar{\delta} a^*, \quad (1.4)$$

где $\bar{\Delta}a^*$ и $\bar{\delta}a^*$ – известные величины, которые будем называть *верхними границами* (или просто *границами*) *абсолютной* и *относительной погрешностей*.

Абсолютные и относительные погрешности числа принято округлять только в большую сторону, так как при округлениях границы неопределенности числа, как правило, увеличиваются. По этой причине вычисления ведут с одним–двумя запасными знаками.

Значащими цифрами числа a^* называют все цифры в его записи, начиная с первой ненулевой слева. Значащая цифра называется *верной*, если абсолютная погрешность числа не превосходит $1/2$ единицы разряда, соответствующего этой цифре.

3. Запись приближенного числа. Вообще приближенное значение принято записывать так, чтобы все цифры в записи были верными. Такая запись приближенных чисел позволяет грубо судить об их погрешностях и потому не требует дополнительного выписывания погрешностей. В том случае, когда число содержит сомнительные цифры или имеет слишком много десятичных знаков, производят округление. Округление чисел, записанных в десятичной системе, производится по правилу первой отбрасываемой цифры:

- если первая из отбрасываемых цифр меньше 5, то оставляемые десятичные знаки сохраняются без изменения;
- если первая из отбрасываемых цифр больше 5, то последняя оставляемая цифра увеличивается на единицу;
- если первая из отбрасываемых цифр равна 5, а за ней идут нули, то последняя оставляемая цифра увеличивается на единицу;
- если первая из отбрасываемых цифр равна 5, и все значащие цифры, идущие за ней, – нули, то последняя оставляемая цифра увеличивается на единицу, если она нечетная, и остается без изменения, если – четная.

Это правило округления обеспечивает увеличение абсолютной погрешности не более чем на половину последнего сохраняемого разряда. При работе на ЭВМ используются как способ отбрасывания всех

разрядов, не помещающихся в ячейку, без изменения последней цифры, так и предшествующее отбрасыванию округление по правилу, аналогичному приведенному выше. Второй режим работы требует больших затрат машинного времени.

Пример 1. У чисел $a^* = \underline{1.275038}$, $a^* = 0.0204$ и $a^* = 0.0204000$ значащие цифры подчеркнуты. Первое число имеет 7, второе 3, а третье 6 значащих цифр.

Пример 2. Пусть $a^* = 12.396$ и известно, что $\Delta a^* = 0.03$. Сколько верных значащих цифр у числа a^* ?

В приближенном числе

$$a^* = \pm (a_1 \cdot 10^{k-1} + a_2 \cdot 10^{k-2} + \dots + a_m \cdot 10^{k-m} + a_{m+1} \cdot 10^{k-m-1} + \dots) \quad (1.5)$$
 цифра a_m называется верной (по данному выше определению), если

$$\Delta a^* \leq 0.5 \cdot 10^{k-m}.$$

Очевидно, если a_m — верная цифра, то и все предшествующие цифры числа верные. Пусть a_m — верная последняя цифра, тогда число верных знаков приближенного числа будет равно m .

Если неравенство $\Delta a^* < 0.5 \cdot 10^{k-m}$ не выполняется, то цифра a_m называется сомнительной.

В примере

$$a^* = 12.396 = 1 \cdot 10^1 + 2 \cdot 10^0 + 3 \cdot 10^{-1} + 9 \cdot 10^{-2} + 6 \cdot 10^{-3},$$

т. е. $k = 2$ (сравнивая с формулой (1.5)). Тогда имеем

$$\Delta a^* \leq 0.5 \cdot 10^{2-m}; \quad 0.03 \leq 0.5 \cdot 10^{2-m};$$

т. е. $m = 3$. Значит, у числа a^* верные знаки 1, 2, 3, а 9 и 6 сомнительные.

Пример 3. Пусть $a^* = 0.037862$ и $\Delta a^* = 0.007$.

Здесь $k = -1$, $0.007 \leq 0.5 \cdot 10^{-2-m}$, $m = -1$. Значит, у числа a^* все цифры сомнительные.

Тема 2. Численные методы решения систем линейных алгебраических уравнений

Литература: [1] гл. 4, п. 1–3; [2] л. 2, 3; [3] ч. 2, гл. 4, п. 2–4, гл. 5, п. 1; [4] гл. 6, п. 1, 3, 7, 11; [5] ч. 2, гл. 1, п. 1, 3, 6, гл. 2, п. 1–3; [6] гл. 5, п. 5.1–5.5, 5.9, гл. 6; [7] гл. 2, п. 1, 3; [8] гл. 3, п. 1, 2, 4, 9, 10.

Основные теоретические сведения

1. Решением системы линейных алгебраических уравнений (СЛАУ)

$$A \cdot X = B \quad (2.1)$$

является вектор $X = (x_1, x_2, \dots, x_n)^T$, который будем рассматривать как элемент векторного пространства R^n . Приближенное решение $X^* = (x_1^*, x_2^*, \dots, x_n^*)^T$ и погрешность $X - X^* = (x_1 - x_1^*, \dots, x_n - x_n^*)^T$ также являются элементами пространства R^n . Для того чтобы анализировать методы решения систем, необходимо уметь количественно оценивать «величины» векторов X^* и $X - X^*$, а также векторов B и $B - B^*$, где $B^* = (b_1^*, b_2^*, \dots, b_n^*)^T$ – вектор приближенно заданных правых частей. Удобной для этой цели количественной характеристикой является широко используемое понятие нормы вектора.

Нормой вектора X называется число, обозначаемое $\|X\|$ и удовлетворяющее условиям:

- $\|X\| \geq 0, \|X\| = 0 \Leftrightarrow X = O$,
- $\|\alpha X\| = |\alpha| \cdot \|X\|$, α – скаляр,
- $\|X + Y\| \leq \|X\| + \|Y\|$.

Примеры:

- $\|X\|_1 = \sum_{i=1}^n |x_i|$,
- $\|X\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}$ – евклидова норма,
- $\|X\|_\infty = \max_{1 \leq i \leq n} |x_i|$ – равномерная норма.

Пространство с введенной в нем нормой называют **нормированным**. Одновременно оно является **метрическим**, т. к. норма определяет **метрику** – расстояние между элементами пространства:

$$\rho(X, Y) = \|X - Y\|.$$

2. Абсолютная и относительная погрешность вектора. Пусть в пространстве n -мерных векторов R^n введена и фиксирована некоторая норма $\|X\|$ (например, одна из норм (2.2)). В этом случае в качестве меры степени близости векторов X и X^* естественно использовать величину $\|X - X^*\|$, являющуюся аналогом расстояния между точками X и X^* . Введем абсолютную и относительную погрешности вектора X^* с помощью формул

$$\|\Delta X^*\| = \|X - X^*\|, \quad \|\delta X^*\| = \|X - X^*\| / \|X\|. \quad (2.3)$$

Выбор той или иной конкретной нормы в практических задачах диктуется тем, какие требования предъявляются к точности решения. Выбор нормы $\|\Delta X\|_1$ фактически отвечает случаю, когда малой должна быть суммарная абсолютная ошибка в компонентах решения; выбор $\|\Delta X\|_2$ соответствует критерию малости среднеквадратичной ошибки, а принятие в качестве нормы $\|\Delta X\|_\infty$ означает, что малой должна быть максимальная из абсолютных ошибок в компонентах решения.

3. Нормой квадратной матрицы A называется число, обозначаемое $\|A\|$ и удовлетворяющее свойствам:

- $\|A\| \geq 0$, $\|A\| = 0 \Leftrightarrow A = O$,
- $\|\alpha A\| = |\alpha| \cdot \|A\|$, α – скаляр,
- $\|A + B\| \leq \|A\| + \|B\|$,
- $\|A \cdot B\| \leq \|A\| \cdot \|B\|$.

Норма матрицы $\|A\|$ **согласована с нормой вектора $\|X\|$** , если

$$\|A \cdot X\| \leq \|A\| \cdot \|X\|.$$

✓ Каждой из векторных норм $\|X\|$ соответствует своя согласованная норма матрицы $\|A\|$. Нормам $\|X\|_1, \|X\|_2, \|X\|_\infty$ подчинены нормы $\|A\|_1, \|A\|_2, \|A\|_\infty$, вычисляемые по формулам:

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|, \quad (2.4)$$

$$\|A\|_2 \leq \|A\|_E = \left(\sum_{i,j=1}^n |a_{ij}|^2 \right)^{1/2}, \quad (2.5)$$

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|. \quad (2.6)$$

Здесь $\|A\|_E$ – величина, называемая *евклидовой нормой матрицы* A .

4. Обусловленность СЛАУ. Чувствительность решения к возмущению (погрешностям) входных данных можно охарактеризовать с помощью так называемого числа $\mu(A)$ обусловленности.

✓ **Числом обусловленности матрицы** A называют число

$$\mu(A) = \|A\| \cdot \|A^{-1}\|. \quad (2.7)$$

Всегда

$$\mu(A) \geq 1.$$

Теорема 1. Пусть правая часть линейного уравнения (2.1) где A – невырожденная матрица, получила погрешность – приращение ΔB . Тогда решение X уравнения получит некоторое приращение ΔX , так что

$$A(X + \Delta X) = B + \Delta B. \quad (2.8)$$

Относительная погрешность $\|\Delta X\|/\|X\|$ решения удовлетворяет неравенству

$$\frac{\|\Delta X\|}{\|X\|} \leq \mu(A) \frac{\|\Delta B\|}{\|B\|}. \quad (2.9)$$

5. Метод исключения Гаусса. Опишем простой метод Гаусса (схема единственного деления). Для удобства изложения ограничимся рас-

смотрением системы четвертого порядка. Запишем рассматриваемую СЛАУ в виде

$$\sum_{j=1}^4 a_{ij} x_j + a_{i5} = 0, \quad 1 \leq i \leq 4, \quad (2.10)$$

где все члены перенесены в левую часть равенства. Из первого уравнения системы (2.10) выражаем x_1 через остальные неизвестные. Получим

$$x_1 = b_{12} x_2 + b_{13} x_3 + b_{14} x_4 + b_{15}, \quad (2.11)$$

где $b_{1j} = -a_{1j}/a_{11}$, $a_{11} \neq 0$, $2 \leq j \leq 5$. Исключим из оставшихся уравнений x_1 ; для этого достаточно подставить значение x_1 из (2.11) во 2, 3 и 4 уравнения системы (2.10). Тогда придем к СЛАУ из трех уравнений, не содержащих x_1 :

$$\sum_{j=2}^4 a_{ij}^{(1)} x_j + a_{i5}^{(1)} = 0. \quad (2.12)$$

Из способа получения этой системы видно, что

$$a_{ij}^{(1)} = a_{ij} + a_{i1} \cdot b_{1j}, \quad 2 \leq i \leq 4, \quad 2 \leq j \leq 5.$$

Систему (2.12) можно подвергнуть тому же преобразованию, что и первоначальную. Продолжая этот процесс дальше, приходим к уравнению $x_4 = b_{45}$. Итак, получаем четыре уравнения (вида (2.11)), которые объединяем в систему

$$\begin{cases} x_1 = b_{12} x_2 + b_{13} x_3 + b_{14} x_4 + b_{15}, \\ x_2 = b_{23} x_3 + b_{24} x_4 + b_{25}, \\ x_3 = b_{34} x_4 + b_{35}, \\ x_4 = b_{45}. \end{cases} \quad (2.13)$$

Из уравнений (2.13) последовательно находим все x_i , $1 \leq i \leq 4$. Указанный алгоритм может оказаться нереализуемым из-за деления на нуль или дать грубую ошибку в результате округлений.

Сформулируем достаточное условие, гарантирующее вычислительную устойчивость метода Гаусса. Матрица

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

называется *матрицей с диагональным преобладанием* (величина $\delta > 0$), если

$$|a_{ii}| \geq \sum |a_{ij}| + \delta, 1 \leq i \leq n. \quad (2.14)$$

Теорема 2. Пусть матрица A системы (2.1) является матрицей с диагональным преобладанием (величина $\delta > 0$). Тогда в алгоритме простого метода Гаусса не встречается деления на нуль.

6. Метод простых итераций (метод Якоби) состоит в том, что система уравнений (2.1) преобразуется к виду

$$\mathbf{X} = \mathbf{C} \cdot \mathbf{X} + \mathbf{Z} \quad (2.15)$$

и ее решение вычисляется как предел последовательности

$$\mathbf{X}^{(k+1)} = \mathbf{C} \cdot \mathbf{X}^{(k)} + \mathbf{Z}. \quad (2.16)$$

К виду (2.15) систему можно привести, например, выделив диагональные элементы:

$$x_i = \frac{1}{a_{ii}} \left(-a_{i,n+1} - \sum_{j \neq i}^n a_{ij} x_j \right), 1 \leq i \leq n,$$

или

$$x_i = z_i + \sum_{j \neq i}^n c_{ij} x_j. \quad (2.17)$$

Теорема 3. Для сходимости итераций (2.16) к решению системы (2.15) необходимо и достаточно, чтобы все собственные значения матрицы $\mathbf{C}$ по абсолютной величине были меньше единицы.

Теорема 4. Для сходимости итераций (2.16) к решению системы (2.15) достаточно, чтобы в какой-либо норме выполнялось условие

$$\|\mathbf{C}\| < 1. \quad (2.18)$$

Тогда независимо от выбора $\mathbf{X}^{(0)}$

$$\|\mathbf{X}^{(k)} - \mathbf{X}^*\| \leq \|C\|^k \cdot \|\mathbf{X}^{(0)} - \mathbf{X}^*\|, \quad (2.19)$$

где $\mathbf{X}^*$ — точное решение (2.15).

Очередное приближение в методе простых итераций вычисляется по формуле

$$x_i^{(k+1)} = \sum_{j=2}^n c_{ij} x_j^{(k)} + z_i, \quad 1 \leq i \leq n, \quad (2.20)$$

где $k = 0, 1, \dots$ — номер очередного приближения к решению.

В качестве критерия окончания итерационного процесса может быть использовано неравенство

$$\|\mathbf{X}^{(k)} - \mathbf{X}^{(k-1)}\| < \varepsilon_1, \quad (2.21)$$

где $\varepsilon_1 = \varepsilon(1 - \|C\|)/\|C\|$.

7. Метод Зейделя. Пусть система (2.1) приведена к виду (2.17). Метод Зейделя можно рассматривать как модификацию метода Якоби. Основная идея модификации состоит в том, что при вычислении очередного $(k+1)$ -го приближения к неизвестному x_i при $i > 1$ используют уже найденные $(k+1)$ -е приближения к неизвестным $x_1, \dots, x_{i-1}$, а не k приближения, как в методе Якоби.

На $(k+1)$ -й итерации компоненты приближения $x^{(k+1)}$ вычисляются по формулам

$$x_i^{(k+1)} = \sum_{j=1}^{i-1} c_{ij} x_j^{(k+1)} + \sum_{j=i}^n c_{ij} x_j^{(k)} + z_i, \quad 1 \leq i \leq n. \quad (2.22)$$

Вводя нижнюю и верхнюю треугольные матрицы

$$C_1 = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ b_{21} & 0 & 0 & \dots & 0 \\ b_{31} & b_{32} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & b_{n3} & \dots & 0 \end{pmatrix}, \quad C_2 = \begin{pmatrix} 0 & b_{12} & b_{13} & \dots & b_{1n} \\ 0 & 0 & b_{23} & \dots & b_{2n} \\ 0 & 0 & 0 & \dots & b_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

расчетные формулы метода примут компактный вид:

$$\mathbf{X}^{(k+1)} = C_1 \cdot \mathbf{X}^{(k+1)} + C_2 \cdot \mathbf{X}^{(k)} + \mathbf{Z}. \quad (2.23)$$

Теорема 5. (достаточное условие сходимости). Пусть $\|C\| < 1$, где $\|C\|$ — одна из норм $\|C\|_\infty, \|C\|_1$. Тогда при любом выборе начального приближения $\mathbf{X}^{(0)}$ метод Зейделя сходится со скоростью геометрической прогрессии, знаменатель которой $q \leq \|C\|$.

В качестве критерия окончания итерационного процесса может быть использовано неравенство

$$\|\mathbf{X}^{(k)} - \mathbf{X}^{(k-1)}\| < \varepsilon_2, \quad (2.24)$$

где $\varepsilon_2 = \varepsilon(1 - \|C\|)/\|C\|_2$.

Пример 1. Решить СЛАУ простым методом Гаусса, проведя все вычисления с шестью значащими цифрами.

$$\begin{cases} 87.526x_1 - 17.576x_2 + 18.000x_3 - 20.148 = 0, \\ -17.576x_1 + 18.000x_2 - 2.600x_3 + 28.980 = 0, \\ 18.000x_1 - 2.600x_2 + 6.000x_3 - 11.400 = 0. \end{cases}$$

Решение. При вычислениях могут произойти ошибки, поэтому необходим контроль вычислений, который основан на следующем соображении. Рассмотрим $\{y_j\}_{j=1}^3$, причем будем считать, что

$$y_j = x_j - 1 \quad (1 \leq j \leq 3).$$

Тогда

$$\begin{aligned} \sum_{j=1}^3 a_{ij} y_j &= \sum_{j=1}^3 a_{ij} (x_j - 1) = \sum_{j=1}^3 a_{ij} x_j - \sum_{j=1}^3 a_{ij} = -a_{i4} - \sum_{j=1}^3 a_{ij} = -\sum_{j=1}^4 a_{ij}, \\ \sum_{j=1}^3 a_{ij} y_j + a_{i5} &= 0, \quad a_{i5} = \sum_{j=1}^4 a_{ij}, \quad 1 \leq i \leq 3. \end{aligned}$$

Таким образом, имеем две системы

$$\sum_{j=1}^3 a_{ij} x_j + a_{i4} = 0 \text{ и } \sum_{j=1}^3 a_{ij} y_j + a_{i5} = 0 \quad (1 \leq i \leq 3),$$

у которых различны свободные члены. Поэтому для контроля в прямом ходе (приведение системы к треугольному виду (2.13)) необходимо дополнительно вести над элементами $\{a_{i5}\}_1^3$ такие же преобразования, как и над $\{a_{i4}\}_1^3$. Обратный ход (определение неизвестных из (2.13)) следует выполнить отдельно для вычисления $\{x_j\}_1^3$ и $\{y_j\}_1^3$. Если окажется, что $y_j = x_j - 1$ ($1 \leq j \leq 3$), то вычисления выполняются правильно. Кроме того, имеем возможность контролировать промежуточные вычисления, поскольку в процессе всех вычислений сумма элементов строки $\left(\sum_{j=1}^3 a_{ij}, 1 \leq i \leq 3 \right)$ должна быть всегда равна последнему контрольному элементу данной строки $(a_{i5}, 1 \leq i \leq 3)$.

Разберем схему записи вычислений при их практическом выполнении. Прямой ход в свою очередь делится на несколько этапов, которые условно обозначим через ПХ₁, ПХ₂, ПХ₃, ПХ₄; обратный ход обозначим через ОХ. Все результаты вычислений будем записывать в одну табл. 2.1, порядок заполнения которой следующий. В раздел ПХ₁ табл. 2.1 записываем расширенную матрицу системы (коэффициенты при неизвестных и свободные члены) и для каждой строки контрольные суммы $(K \Sigma_i)$, равные сумме элементов в данной строке. Затем к этим числам присоединяем строку элементов b_{ij} уравнения (2.11). Числа в этой строке получаются делением всех элементов первой строки на ее крайний левый элемент, взятый с противоположным знаком. Например, $b_{12} = -17.576 / (-87.526) = 0.200809$ (назовем эту строку «отмеченной»).

Элементы раздела ПХ₂ вычисляем по следующему правилу: *к каждому элементу матрицы раздела ПХ₁ (кроме первой и последней строки) прибавляется произведение крайнего левого элемента той же строки на крайний нижний элемент того же столбца.*

Элементы «отмеченной» строки раздела PX_2 получим путем деления всех элементов первой вновь полученной строки на ее крайний левый элемент, взятый с противоположным знаком. Этим завершаются вычисления раздела PX_2 . Аналогично выполняются вычисления раздела PX_3 . В разделе PX_3 находим неизвестное x_3 и на этом завершаем прямой ход вычислений.

Для вычисления элементов раздела ОХ, т. е. Для нахождения неизвестных, используем «отмеченные» строки, начиная с последней. На пересечении первой строки раздела ОХ со столбцом свободных членов записываем 1, а со столбцом контрольной суммы $(K\Sigma) - 0$.

Для вычисления неизвестных x_j можно воспользоваться следующим правилом: *каждое неизвестное x_j равно скалярному произведению уже вычисленной части строки неизвестных (первая строка ОХ) на соответствующую часть («отмеченной») строки схемы PX_j .*

Контроль осуществляется с помощью столбцов $K\Sigma$. Над столбцами $K\Sigma$ производятся те же действия, что и над остальными столбцами. В итоге сумма элементов каждой строки схемы (без столбца $K\Sigma$) должна быть равна элементу этой строки из столбца $K\Sigma$. Для контроля на обратном ходе на пересечении второй строки раздела ОХ со столбцом свободных членов записываем 0, а со столбцом контрольной суммы $(K\Sigma) - 1$. Для вычисления неизвестных y_j используем последнее правило, при этом, как уже отмечалось,

$$y_j = x_j - 1 \quad (1 \leq j \leq 3).$$

Пример 2. Используя метод простой итерации (Якоби), найти решение системы

$$\begin{cases} 4x_1 - x_2 + 2x_3 = 1.75, \\ x_1 - 5x_2 + 3x_3 = 2.5, \\ 2x_1 + x_2 - 8x_3 = -0.25, \end{cases} \quad (2.25)$$

с точностью $\varepsilon = 10^{-4}$ в норме $\|\cdot\|_\infty$.

Таблица 2.1

Этапы	Коэффициенты при			Свободные члены	К Σ
	x_1	x_2	x_3		
ПХ ₁	87.526000	-17.576000	18.000000	-20.148000	67.802000
	-17.576000	18.000000	-2.600000	28.980000	26.804000
	18.000000	-2.600000	6.000000	-11.400000	10.000000
	-1	0.200811	-0.205650	0.230190	-0.774650
ПХ ₂		14.470583	1.014560	24.934102	40.419245
		1.014560	2.298243	-7.256500	-3.943697
		-1	-0.070112	-1.723089	-2.793200
ПХ ₃			2.227110	-9.004677	-6.777567
			-1	4.043212	3.043212
ОХ	-1.004241	-2.006566	4.043212	1	0
	-2.004241	-3.006566	3.043212	0	1

Решение. Вычисляя коэффициенты по формулам

$$c_{ij} = -a_{ij}/a_{ii}, \quad z_i = b_i/a_{ii} \quad (1 \leq i, j \leq n, j \neq i), \quad (2.26)$$

приведем систему к виду, удобному для итераций. Для этого из первого уравнения системы (2.25) выразим неизвестное x_1 , из второго уравнения – неизвестное x_2 , из третьего – неизвестное x_3 .

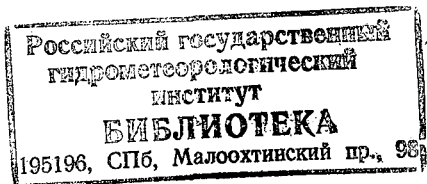
В результате получим систему

$$\begin{cases} x_1 = & 0.25x_2 - 0.5x_3 + 0.4375, \\ x_2 = 0.20x_1 + & 0.6x_3 - 0.5, \\ x_3 = 0.25x_1 + 0.125x_2 + & 0.03125. \end{cases} \quad (2.27)$$

В последнем уравнении коэффициенты даны с точностью до погрешности округления. Здесь

$$C = \begin{pmatrix} 0 & 0.25 & -0.5 \\ 0.2 & 0 & 0.6 \\ 0.25 & 0.125 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 0.4375 \\ -0.5 \\ 0.03125 \end{pmatrix}.$$

Достаточное условие сходимости метода простой итерации выполнено, так как $\|C\|_{\infty} = \max\{0.75, 0.8, 0.375\} = 0.8 < 1$.



Примем за начальное приближение к решению вектор

$$\mathbf{X}^{(0)} = (0.43750, -0.50000, 0.03125)^T$$

и будем вести итерации по формуле

$$x_i^{(k+1)} = z_i + \sum_{j=1}^3 c_{ij} x_j^{(k)}$$

до выполнения критерия окончания (2.21), где в данном случае

$$\varepsilon_1 = (1 - 0.8) / 0.8 \cdot 10^{-4} = 0.25 \cdot 10^{-4}.$$

Значения приближений в табл. 2.2 приводятся с пятью цифрами после десятичной точки.

Таблица 2.2

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\ X^{(k)} - X^{(k-1)}\ _{\infty}$
0	0.43750	-0.50000	0.03125	—
1	0.29688	-0.39375	0.07812	0.14062
2	0.30000	-0.39375	0.05625	0.02187
3	0.31094	-0.40625	0.05703	0.01250
4	0.30742	-0.40359	0.05820	$3.52 \cdot 10^{-3}$
5	0.30750	-0.40360	0.05766	$5.4 \cdot 10^{-4}$
6	0.30777	-0.40390	0.05768	$3 \cdot 10^{-4}$
7	0.30768	-0.40384	0.05770	$9 \cdot 10^{-5}$
8	0.30769	-0.40384	0.05769	$1 \cdot 10^{-5}$

При $k=8$ условие (2.21) выполняется и можно положить

$$x_1 = 0.3077 \pm 0.0001, x_2 = -0.4038 \pm 0.0001, x_3 = 0.0577 \pm 0.0001.$$

Пример 3. Используя метод Зейделя, найти решение системы (2.25) с точностью $\varepsilon = 10^{-4}$.

Решение. После приведения системы к виду (2.27) убеждаемся, что $\|C\|_{\infty} < 1$ (см. пример 2). В силу теоремы 5 метод Зейделя сходится.

Положим $\mathbf{X}^{(0)} = (0.43750, -0.50000, 0.03125)^T$ и будем вычислять последовательно приближения по формулам

$$x_i^{(k+1)} = z_i + \sum_{j=1}^{i-1} c_{ij} x_j^{(k+1)} + \sum_{j=i+1}^3 c_{ij} x_j^{(k)},$$

где c_{ij} и z_i находятся по формулам (2.26). Здесь

$$C_2 = \begin{pmatrix} 0 & 0.25 & -0.5 \\ 0 & 0 & 0.6 \\ 0 & 0 & 0 \end{pmatrix}$$

и $\|C_2\|_\infty = \max\{0.75, 0.6, 0\} = 0.75$. Будем вести итерации до выполнения критерия окончания (2.24), где

$$\varepsilon_2 = 10^{-4} (1 - 0.8) / 0.75 \approx 2.7 \cdot 10^{-4}.$$

Значения приближений с пятью цифрами после десятичной точки приведены в табл. 2.3.

Таблица 2.3

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\ X^{(k)} - X^{(k-1)}\ _\infty$
0	0.43750	-0.50000	0.03125	—
1	0.29688	-0.42187	0.05274	0.14062
2	0.30566	-0.40722	0.05676	0.01465
3	0.30732	-0.40448	0.05752	$2.74 \cdot 10^{-3}$
4	0.30762	-0.40396	0.05766	$5.2 \cdot 10^{-4}$
5	0.30768	-0.40387	0.05769	$6 \cdot 10^{-5}$
6	0.30769	-0.40385	0.05769	$2 \cdot 10^{-5}$

При $k=6$ критерий окончания выполняется и можно положить $x_1 = 0.3077 \pm 0.0001$, $x_2 = -0.4038 \pm 0.0001$, $x_3 = 0.0577 \pm 0.0001$.

Т е м а 3. Приближение функций

Литература: [1] гл. 2; [2] л. 5, 6; [3] ч. I, гл. 1; [4] гл. II, п. 1–3, 10–13, гл. 4 п. 11; [5] ч. 2, гл. 3, п. 1, 2, 4; [6] гл. 11, п. 11.1, 11.3–11.5, 11.7–11.9, 11.13; [7] гл. 6; [8] гл. V, п. 1, 2, 4.

Основные теоретические сведения

✓ **1. Постановка задачи о приближении (аппроксимации) функции:** данную функцию $f(x)$ требуется приближенно заменить (**аппроксимировать**) некоторой функцией $G(x)$ так, чтобы отклонение (в некотором смысле) $G(x)$ от $f(x)$ в заданной области было наименьшим. Функция $G(x)$ при этом называется **аппроксимирующей**.

✓ В процессе численной реализации этого подхода необходимо рассмотреть следующие четыре основных вопроса:

- 1) об имеющейся информации относительно функции $f(x)$, т.е. о виде, в котором задана функция $f(x)$;
- 2) о классе аппроксимирующих функций, т.е. о том, какими функциями $G(x)$ будет аппроксимирована функция $f(x)$;
- 3) о близости аппроксимируемой и аппроксимирующей функций, т.е. о выборе критерия согласия, которому должна удовлетворять функция $G(x)$;
- 4) о погрешности, т.е. об определении разности между точным и приближенным значениями.

Для практики весьма важен случай аппроксимации функции многочленом

$$G(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n. \quad (3.1)$$

При этом коэффициенты a_j будут подбираться так, чтобы достичь наименьшего отклонения многочлена от данной функции.

Если приближение строится на заданном дискретном множестве точек $\{x_i\}_0^n$, то аппроксимация называется **точечной**. К ней относятся интерполирование, среднеквадратичное приближение и др. При построении приближения на непрерывном множестве точек (например, на отрезке $[a; b]$) аппроксимация называется **непрерывной** (или **интегральной**).

✓2. **Постановка задачи интерполяции.** Пусть заданы точки $x_0, x_1, \dots, x_n$ и значения $f(x_0), f(x_1), \dots, f(x_n)$ функции $f(x)$ в этих точках. Соответствие

x_0	x_1	...	x_n
$f(x_0)$	$f(x_1)$	...	$f(x_n)$

будем называть *таблицей значений функции $f(x)$ в узлах $x_0, x_1, \dots, x_n$* и говорить, что функция $f(x)$ задана таблицей своих значений

$$y_i = f(x_i), (0 \leq i \leq n). \quad (3.2)$$

✓**Задача интерполяции** состоит в построении функции $G(x)$, удовлетворяющей условию:

$$G(x_i) = f(x_i) (0 \leq i \leq n). \quad (3.3)$$

Другими словами, ставится задача о построении функции $G(x)$, график которой проходит через заданные точки (x_i, y_i) . Указанный способ приближения функций принято называть *интерполяцией* (или *интерполированием*), а точки x_i — *узлами интерполяции*.

Выбор функций $G(x)$ неоднозначен, так как по заданной таблице можно построить бесконечно много интерполирующих функций.

✓3. **Экстраполяция.** Пусть $x_{\min}, x_{\max}$ — минимальный и максимальный из узлов интерполяции. В случае, когда интерполяция используется для вычисления приближенного значения функции $f(x)$ в точке x , не принадлежащей отрезку $[x_{\min}, x_{\max}]$ (*отрезку наблюдения*), принято говорить о том, что осуществляется *экстраполяция*. Этот метод приближения часто используют с целью прогнозирования характера протекания тех или иных процессов при значениях параметра x , выходящих за пределы отрезка наблюдения.

✓4. **Алгебраическим интерполяционным многочленом $G_n(x) = G_n(x, f, x_0, x_1, \dots, x_n)$** назовем многочлен

$$G_n(x) = a_0 + a_1x + \dots + a_nx^n = \sum_{k=0}^n a_k x^k \quad (3.4)$$

степени не выше n , который в узлах $x_0, x_1, \dots, x_n$ принимает значения $f(x_0), f(x_1), \dots, f(x_n)$,

$$G_n(x_i) = f(x_i) \quad (0 \leq i \leq n). \quad (3.5)$$

Теорема 1. Существует единственный интерполяционный многочлен степени n , удовлетворяющий условиям (3.5).

У5. Интерполяционный многочлен в форме Лагранжа. Приведем одну из форм записи интерполяционного многочлена – **многочлен Лагранжа**.

$$L_n(x) = \sum_{i=0}^n y_i \frac{(x-x_0)(x-x_1)\dots(x-x_{i-1})(x-x_{i+1})\dots(x-x_n)}{(x_i-x_0)(x_i-x_1)\dots(x_i-x_{i-1})(x_i-x_{i+1})\dots(x_i-x_n)}. \quad (3.6)$$

Если ввести в рассмотрение многочлен специального вида $(n+1)$ -й степени

$$\omega_{n+1}(x) = (x-x_0)(x-x_1)\dots(x-x_n), \quad (3.7)$$

тогда многочлен Лагранжа можно записать в виде

$$L_n(x) = \sum_{i=0}^n y_i \frac{\omega_{n+1}(x)}{\omega'_{n+1}(x_i)(x-x_i)}. \quad (3.8)$$

В инженерной практике наиболее часто используется интерполяция первой ($n=1$), второй ($n=2$) и третьей ($n=3$) степени (*линейная, квадратичная и кубическая интерполяция*).

Если узлы упорядочены по величине, т.е. $x_{i+1} > x_i$ для всех i , то величины $\{h_i = x_{i+1} - x_i; 0 \leq i \leq n\}$ называют **шагами интерполяции**. Если $h_i = h = \text{const}$, так что $x_i = x_0 + h$ ($0 \leq i \leq n$), то говорят об **интерполяции по равноотстоящим узлам**.

Употребительны и другие записи (единственного) интерполяционного многочлена $G_n(x)$. Особенно часто используют запись в форме Ньютона.

6. *Интерполяционный многочлен в форме Ньютона. Конечные разности.* Пусть функция $y = f(x)$ задана таблицей (3.2) своих значений. Величину $\Delta y_i = y_{i+1} - y_i$ принято называть *конечной разностью первого порядка* функции $y = f(x)$ в точке x_i (с шагом h). *Конечная разность второго порядка* определяется формулой $\Delta^2 y_i = \Delta y_{i+1} - \Delta y_i$. Аналогично определяются конечные разности третьего и более высоких порядков. Общее определение *конечной разности порядка k* таково:

$$\Delta^k y_i = \Delta^{k-1} y_{i+1} - \Delta^{k-1} y_i.$$

Здесь $k \geq 1$ и $\Delta^0 y_i = y_i$.

Если интерполируемая функция задана на таблице с постоянным шагом h (т.е. $x_i = x_0 + ih$, $0 \leq i \leq n$), то многочлен Ньютона можно записать в следующем виде:

$$P_n(x) = P_n(x_0 + ht) = y_0 + \frac{\Delta y_0}{1!}t + \frac{\Delta^2 y_0}{2!}t(t-1) + \frac{\Delta^3 y_0}{3!}t(t-1)(t-2) + \dots + \frac{\Delta^n y_0}{n!}t(t-1)\dots(t-n+1), \quad (3.9)$$

где $t = (x - x_0)/h$.

✓ Многочлен (3.9) называется *интерполяционным многочленом Ньютона с конечными разностями для интерполяции вперед*.

Если $t = (x - x_0)/h$, то можно записать многочлен в виде *интерполяционного многочлена Ньютона с конечными разностями для интерполяции назад*:

$$P_n(x) = P_n(x_n + ht) = y_n + \frac{\Delta y_{n-1}}{1!}t + \frac{\Delta^2 y_{n-2}}{2!}t(t+1) + \frac{\Delta^3 y_{n-3}}{3!}t(t+1)(t+2) + \dots + \frac{\Delta^n y_0}{n!}t(t+1)\dots(t+n-1). \quad (3.10)$$

✓ **7. Сравнение записей в форме Лагранжа и Ньютона.** Для вычисления значения функции $f(x)$ в точке x , не являющейся узлом интерполяции, можно положить $f(x) \approx G_n(x)$ (где $G_n(x)$ -либо $L_n(x)$, либо $P_n(x)$).

Пусть $G_n(x)$ уже найден, но мы решили для уточнения привлечь еще один узел x_{n+1} и значение $f(x_{n+1})$ в нем. Тогда для вычисления $G_{n+1}(x)$ с помощью формулы Лагранжа (3.8) нужно заново провести всю работу. Для вычисления по формуле Ньютона (3.9) или (3.10) нужно досчитать только поправку

$$\frac{\Delta^{n+1}y_0}{(n+1)!}t(t-1)\cdots(t-n) \text{ или } \frac{\Delta^{n+1}y_0}{(n+1)!}t(t+1)\cdots(t+n).$$

Кстати, сразу будет видно, насколько она велика.

Что касается многочлена Лагранжа, то с точки зрения вычислений он удобней (при правильной организации алгоритма), особенно, если требуется приближать различные функции, заданные табличными значениями в одних и тех же точках.

✓ **8. Погрешность интерполяции.** Ошибка приближения функции интерполяционным многочленом n -й степени в точке x — это разность

$$\rho_n(x) = f(x) - G_n(x).$$

Оценить значение погрешности позволяет следующая теорема.

Теорема 2. Пусть функция $f(x)$ дифференцируема $(n+1)$ раз на отрезке $[a; b]$, содержащем узлы интерполяции x_i , $0 \leq i \leq n$. Тогда

$$\rho_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(x), \quad (3.11)$$

где $\xi \in [a; b]$.

Из (3.11) следует оценка погрешности интерполяции

$$|\rho_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |\omega_{n+1}(x)|, \quad M_{n+1} = \max_{[a; b]} |f^{(n+1)}(x)|. \quad (3.12)$$

✓ **9. Интерполирование сплайнами.** Пусть функция $f(x)$ определена на отрезке $[a; b]$, который разбит точками $a = x_0 < x_1 < \dots < x_n = b$ Own

частичных отрезков $[x_{i-1}; x_i]$ Интерполяционным сплайном степени m функция $S_m(x)$, обладающая следующими свойствами:

- на каждом из частичных отрезков $[x_{i-1}; x_i]$ ($1 \leq i \leq n$) $S_m(x)$ является многочленом степени m ;
- функция $S_m(x)$ непрерывна на отрезке $[a; b]$ вместе со всеми своими производными до порядка $m-1$;
- $S_m(x_i) = f(x_i)$ ($1 \leq i \leq n$).

Если $m \geq 2$, то для единственности $S_m(x)$ следует задать еще $m-1$ условий, которые обычно задаются на концах отрезка $[a; b]$ либо произвольно, либо из дополнительной информации о поведении $f(x)$.

Разность между степенью сплайна и наивысшим порядком непрерывной на отрезке $[a; b]$ производной называется *дефектом сплайна*.

При $m=1$ имеем сплайн первой степени (метод ломаных) с дефектом, равным единице, так как непрерывна только сама функция (нулевая производная), а первая производная уже разрывная.

Наиболее широкое распространение на практике получили сплайны $S_3(x)$ третьей степени (*кубические сплайны*) с дефектом, равным 1 или 2.

✓ *Кубическим сплайном* называется функция $S_3(x)$, обладающая следующими свойствами:

- Функция $S_3(x)$ непрерывна вместе со своими производными до второго порядка включительно;
- на каждом из отрезков $[x_{i-1}; x_i]$ $S_3(x)$ является кубическим многочленом вида

$$S_3(x) = a_i + b_i(x - x_{i-1}) + c_i(x - x_{i-1})^2 + d_i(x - x_{i-1})^3, \quad (3.13)$$

$$x_{i-1} \leq x \leq x_i, \quad 1 \leq i \leq n;$$

- в узлах сетки $\{x_i\}_{i=1}^n$ выполняются равенства

$$S_3(x_i) = f(x_i), \quad 1 \leq i \leq n; \quad (3.14)$$

- $S_3''(x)$ удовлетворяет граничным условиям

$$S_3''(a) = S_3''(b) = 0. \quad (3.15)$$

10. Метод наименьших квадратов. Пусть функция $f(x)$ задана таблицей своих значений (3.2), т. е. $(x_i, y_i), 1 \leq i \leq n$. Будем искать приближение в виде многочлена степени m :

$$G_m(x) = a_0 x^m + a_1 x^{m-1} + \dots + a_{m-1} x + a_m \quad (a_0 \neq 0),$$

такого, который минимизирует сумму квадратов отклонений многочлена от заданных значений функции:

$$\delta(a_0, a_1, \dots, a_m) = \sum_{i=0}^n (G_m(x_i) - y_i)^2. \quad (3.16)$$

- $f \in \mathcal{A}_m \geq n$ существует бесконечное множество многочленов, для которых выполняется равенство $\delta = 0$.
- $f \in \mathcal{A}_m = n - 1$ решением поставленной задачи является интерполяционный многочлен, ибо на нем достигается абсолютный минимум: $\delta = 0$.
- $f \in \mathcal{A}_m < n - 1$ (в дальнейшем рассматриваем только этот случай), $\delta \leq 0$ при любых значениях коэффициентов многочлена $G_m(x)$; нужно так выбрать коэффициенты этого многочлена, чтобы величина δ стала минимальной.

В точке минимума функции δ ее производные $\partial \delta / \partial a_j$ обращаются в нуль. Дифференцируя δ и приравнявая к нулю производные, получим так называемую нормальную систему метода наименьших квадратов. Эта система линейных алгебраических уравнений (СЛАУ) относительно неизвестных $a_0, a_1, \dots, a_n$. Можно показать, что определитель этой системы отличен от нуля, т.е. решение системы существует и единственно. Однако на практике описанную методику применяют только для нахождения многочленов, степень которых не выше 4–5. При более высоких степенях нормальная система становится плохо обусловленной и погрешности определения коэффициентов велики.

Пример 1. Аппроксимировать многочленом второй степени по методу наименьших квадратов функцию, заданную таблицей:

x	-2.6	-2.1	-1.1	0	1.1	2.1
$y = f(x)$	2.5	3.8	5	4.1	0.6	-4.6

Вычислить среднеквадратическое отклонение по формуле

$$\sigma = \left(\sum_{i=1}^n (G_2(x_i) - y_i)^2 \right)^{1/2}.$$

Решение. В данном примере заданную табличную функцию ($n=6$) требуется аппроксимировать многочленом второй степени ($m=2$) $G_2(x) = a_0x^2 + a_1x + a_2$ ($a_0 \neq 0$). В этом случае нормальная система имеет вид

$$\begin{cases} a_0 \sum_{i=1}^n x_i^4 + a_1 \sum_{i=1}^n x_i^3 + a_2 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i x_i^2; \\ a_0 \sum_{i=1}^n x_i^3 + a_1 \sum_{i=1}^n x_i^2 + a_2 \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i; \\ a_0 \sum_{i=1}^n x_i^2 + a_1 \sum_{i=1}^n x_i + n \cdot a_2 = \sum_{i=1}^n y_i. \end{cases} \quad (3.17)$$

Исходные данные и результаты обработки представлены в таблице 3.1.

Таблица 3.1

i	x_i	x_i^2	x_i^3	x_i^4	y_i	$x_i y_i$	$x_i^2 y_i$
1	-2.6	6.76	-17.576	45.6976	2.5	-6.5	16.900
2	-2.1	4.41	-9.261	19.4481	3.8	-7.98	16.758
3	-1.1	1.21	-1.331	1.4661	5	-5.5	6.050
4	0	0	0	0	4.1	0	0
5	1.1	1.21	1.331	1.4661	0.6	0.66	0.726
6	2.1	4.41	9.261	19.4481	-4.6	-9.66	-20.286
Σ	-2.6	18.00	-17.576	87.5260	11.4	-28.98	20.148

Система уравнений (3.17) в данном случае принимает вид

$$\begin{cases} 87.526a_0 - 17.576a_1 + 18.000a_2 = 20.148; \\ -17.576a_0 + 18.000a_1 - 2.600a_2 = -28.980; \\ 18.000a_0 - 2.600a_1 + 6.000a_2 = 11.4. \end{cases}$$

Решая эту систему методом Гаусса (см. пример 1 темы 2), находим:

$$a_0 = -1.00424, a_1 = -2.00657, a_2 = 4.04321.$$

Следовательно,

$$G_2(x) = -1.00424x^2 - 2.00657x + 4.04321.$$

Т е м а 4. Численное дифференцирование и интегрирование

Литература: [1] гл. 3; [2] л. 6; [3] ч. 1, гл. 3; [4] гл. 2, п. 15, 16, гл. 3, п. 1-6, 8; [5] ч. 2, гл. 4; [6] гл. 12, 13; [7] гл. 7, 8; [8] гл. 6, 7.

Основные теоретические сведения

1. Численное дифференцирование. Пусть функция $y = f(x)$ задана таблицей своих значений

$$y_i = f(x_i), \quad (0 \leq i \leq n). \quad (4.1)$$

Простейший способ построения формул численного дифференцирования состоит в следующем. По табличным данным приближаем функцию интерполяционным многочленом с остаточным членом R_n , так что

$$f(x) = G_n(x) + R_n(x) \quad (4.2)$$

Дифференцируя равенство (4.2) необходимое число раз, можно найти значения производных $f'(x), f''(x), \dots, f^{(k)}(x)$ ($k \leq n$):

$$\begin{aligned} f'(x) &= G'_n(x) + R'_n(x), \quad f''(x) = G''_n(x) + R''_n(x), \dots, \\ f^{(k)}(x) &= G^{(k)}_n(x) + R^{(k)}_n(x). \end{aligned} \quad (4.3)$$

В качестве приближенного значения производной порядка k функции $f(x)$ можно принять соответствующее значение производной многочлена $G_n(x)$, т. е. $f^{(k)}(x) \approx G^{(k)}_n(x)$.

Величина

$$R^{(k)}_n(x) = f^{(k)}(x) - G^{(k)}_n(x),$$

характеризующая отклонение приближенного значения производной от ее истинного значения, называется *погрешностью аппроксимации* производной.

При вычислениях на ЭВМ, когда мы заинтересованы, скорее, не в малом количестве арифметических операций, а в экономии элементов памяти и в простоте программирования, достаточно выгодно использовать представление интерполяционного многочлена $G_n(x)$ в форме Лагранжа. При счете на микрокалькуляторах более рационально, по-видимому, воспользоваться многочленом $G_n(x)$ в форме Ньютона. Наиболее простой вид принимают формулы численного дифференцирования при использовании таблиц (4.1) с постоянным шагом ($x_n = x_0 + nh$) или ($x_{i+1} = x_i + h; 0 \leq i \leq n$). Если, например, исходить из правила Ньютона для интерполирования в начале таблицы, получатся следующие выражения для производных:

$$\begin{aligned} y'(x_0 + th) &= \frac{dG_n}{dt} \cdot \frac{dt}{dx} = \frac{1}{h} \cdot \frac{dG_n}{dx} = \\ &= \frac{1}{h} \left(\Delta y_0 + \frac{2t-1}{2!} \Delta^2 y_0 + \frac{3t^2-6t+2}{3!} \Delta^3 y_0 + \dots \right) \\ y''(x_0 + th) &= \frac{1}{h^2} (\Delta^2 y_0 + (t-1) \Delta^3 y_0 + \dots), \end{aligned} \quad (4.4)$$

$$t = (x - x_0)/h, \quad h = x_1 - x_0 = \text{const.}$$

При $x = x_0$ и $t = 0$ будет

$$\begin{aligned} y'(x_0) &= \frac{1}{h} \left(\Delta y_0 - \frac{1}{2} \Delta^2 y_0 + \frac{1}{3} \Delta^3 y_0 - \dots + (-1)^{n-1} \frac{1}{n} \Delta^n y_0 \right) = \\ &= \frac{1}{h} \sum_{j=1}^n \frac{(-1)^{j-1}}{j} \Delta^j y_0, \end{aligned} \quad (4.5)$$

$$y''(x_0) = \frac{1}{h^2} \left(\Delta^2 y_0 - \Delta^3 y_0 + \frac{11}{12} \Delta^4 y_0 - \frac{5}{6} \Delta^5 y_0 + \dots \right). \quad (4.6)$$

Если для интерполирования в конце таблицы воспользоваться многочленом Ньютона, то получим следующие выражения для производных:

$$y'(x_n + th) = \frac{1}{h} \left(\Delta y_{n-1} + \frac{2t+1}{2!} \Delta^2 y_{n-2} + \frac{3t^2+6t+2}{3!} \Delta^3 y_{n-3} + \dots \right),$$

$$y''(x_n + th) = \frac{1}{h^2} (\Delta^2 y_{n-2} + (t+1) \Delta^3 y_{n-3} + \dots),$$

$$t = (x - x_n) / h.$$

Очень простые формулы численного дифференцирования для функции (4.1) можно получить из формулы Тейлора, которую запишем для $f(x)$, полагая ее достаточно гладкой в окрестности точки x :

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \frac{f'''(x)}{3!}h^3 + \dots + \frac{f^{(n)}(x)}{n!}h^n + O(h^{n+1}) \quad (4.7)$$

Если ограничиться первыми двумя слагаемыми в правой части (4.7), т.е. аппроксимировать $f(x)$ в окрестности точки x линейной функцией, то получим формулу **правой разностной производной**

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}. \quad (4.8)$$

Погрешность такой аппроксимации является величиной $O(h)$.

Заменив в последней формуле h на $-h$, получим **левую разностную производную**

$$f'(x) \approx \frac{f(x) - f(x-h)}{h}. \quad (4.9)$$

Вычитая из приближенного равенства $f(x+h) \approx f(x) + f'(x)h$ приближенное равенство $f(x-h) \approx f(x) - f'(x)h$, получим **центральную разностную производную**

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}, \quad (4.10)$$

порядок погрешности которой $O(h^2)$.

Если в формуле Тейлора оставить на одно слагаемое больше, т.е. аппроксимировать функцию многочленом второй степени по h , то

получим центральную разностную производную второго порядка, погрешность которой является величиной $O(h^2)$:

$$f''(x) \approx \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}. \quad (4.11)$$

Одним из источников погрешности численного дифференцирования является погрешность аппроксимации. Она определяется величиной остаточного члена. Погрешности, возникающие при численном дифференцировании, определяются также неточными значениями функции $f(x_i)$ в узлах (неустраняемая погрешность) и погрешностями округлений при проведении расчетов. Погрешность метода численного дифференцирования тем меньше, чем меньше шаг h , а вычислительная погрешность с уменьшением шага увеличивается. С учетом этого обстоятельства выбирают *оптимальный шаг дифференцирования*, при котором суммарная погрешность была бы минимальной.

2. Численное интегрирование. Рассмотрим способы приближенного вычисления определенных интегралов вида

$$\int_a^b F(x)dx = \int_a^b p(x)f(x)dx, \quad (4.12)$$

где $[a, b]$ – любой конечный или бесконечный отрезок числовой оси; $f(x)$ – вещественная произвольная функция некоторого класса, а $p(x)$ – некоторая фиксированная функция, которую называют *весовой*. Если $F(x)$ имеет какие-нибудь особенности, то их выделение обычно делается при помощи разложения $F(x)$ на два сомножителя $F(x) = p(x) \cdot f(x)$, где $p(x)$ имеет особенности того же типа, что и $F(x)$, а $f(x)$ есть гладкая функция.

Наиболее часто приближенное значение интеграла (4.12) ищут в виде линейной комбинации значений функции $f(x)$ на отрезке $[a, b]$:

$$\int_a^b p(x)f(x)dx \approx \sum_{k=0}^n A_k f(x_k). \quad (4.13)$$

Приближенное равенство (4.13) называют **квадратурной формулой**. Выражение в правой части (4.13) называют **квадратурной суммой**, точки x_k — **узлами квадратурной формулы**, а числа A_k — **коэффициентами квадратурной формулы**. Разность

$$r_n = \int_a^b p(x)f(x)dx - \sum_{k=0}^n A_k f(x_k)$$

называется **погрешностью квадратурной формулы**. Погрешность зависит как от расположения узлов, так и от выбора коэффициентов. Будем говорить, что квадратурная формула (4.13) **точна для многочленов степени m** , если для любого многочлена степени не выше m эта формула дает точное значение интеграла, т. е.

$$\int_a^b \varpi_m(x)dx = \sum_{k=0}^n A_k \varpi_m(x_k).$$

✓ **3. Интерполяционное квадратурное правило.** Пусть узлы x_k избраны каким-либо образом и фиксированы. По заданной системе узлов x_k и таблице значений $f(x_k)$ ($0 \leq k \leq n$) построим интерполяционный многочлен Лагранжа и положим в интеграле (4.12)

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

Получим правило приближенного вычисления интеграла, которое характеризуется определенным законом выбора коэффициентов A_k . В связи со способом получения оно названо интерполяционным квадратурным правилом:

$$\int_a^b p(x)f(x)dx \approx \sum_{k=0}^n A_k f(x_k), \quad (4.14)$$

$$A_k = \int_a^b p(x) \frac{\omega(x)}{(x-x_k)\omega'(x_k)} dx.$$

✓ **4. Правила Ньютона–Котеса.** Отрезок интегрирования $[a, b]$ разделим на n одинаковых частей длины $h = (b-a)/n$ и точки деления

$a + kh$ ($0 \leq k \leq n$) примем за узлы интерполяционного квадратурного правила. Квадратурные формулы с равноотстоящими узлами будем применять для вычисления интеграла $\int_a^b f(x)dx$ с постоянной весовой функцией ($p(x) \equiv 1$) и конечным отрезком интегрирования. Для этого случая интерполяционное квадратурное правило (4.14) может быть записано в виде

$$\int_a^b f(x)dx \approx (b-a) \sum_{k=0}^n B_k f(a+kh),$$

$$B_k = \frac{(-1)^{n-k}}{n \cdot k!(n-k)!} \int_0^n \frac{t(t-1) \cdots (t-n)}{t-k} dt. \quad (4.15)$$

Квадратурные формулы (4.15) называют формулами Ньютона–Котеса.

Положим $n=1$. Тогда приближенное равенство (4.15) преобразуется в известную элементарную формулу трапеций

$$\int_a^b f(x)dx \approx \frac{b-a}{2} (f(a) + f(b)), \quad (4.16)$$

которая получается путем замены подынтегральной функции $f(x)$ интерполяционным многочленом первой степени.

При $n=2$ квадратурная формула (4.15) принимает вид

$$\int_a^b f(x)dx \approx \frac{b-a}{6} \left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right), \quad (4.17)$$

который получается путем замены подынтегральной функции $f(x)$ интерполяционным многочленом второй степени (параболой). Эту формулу называют элементарной квадратурной формулой Симпсона или формулой парабол.

Для уменьшения погрешности предварительно разбивают отрезок $[a, b]$ на достаточно большое число интервалов и к каждому из них применяют простейшую квадратурную формулу.

Так, разбивая отрезок $[a, b]$ на m равных частей длиной $h = (b-a)/m$ и применяя к частичному отрезку $[a+kh, a+(k+1)h]$

формулу (4.16), после суммирования по всем частичным отрезкам получаем так называемую *составную* формулу трапеций

$$\int_a^b f(x) dx \approx \sum_{k=0}^n \frac{f(x_k) + f(x_{k-1})}{2} h = h(f_0/2 + f_1 + \dots + f_{m-1} + f_m/2), \quad (4.18)$$

где $f_k = f(x_k)$, $0 \leq k \leq m$.

Если $f''(x)$ непрерывна на $[a, b]$, то оценка погрешности формулы трапеций

$$|r_n| \dots \leq \frac{(b-a)^3}{12m^2} M_2, \quad M_2 = \max_{a \leq x \leq b} |f''(x)|, \quad (4.19)$$

Пусть m – четное число. Возьмем удвоенный частичный отрезок $[a + (k-1)h, a + (k+1)h]$ и применим к нему формулу Симпсона; в результате после суммирования по всем удвоенным частичным отрезкам найдем

$$\int_a^b f(x) dx \approx h/3(f_0 + 2(f_2 + \dots + f_{2l-2}) + 4(f_1 + \dots + f_{2l-1}) + f_{2l}), \quad (4.20)$$

где $m = 2l$.

Приближенную формулу (4.20) называют *составной формулой Симпсона*. Для оценки погрешности имеем неравенство

$$|r_n| \leq \frac{(b-a)^5}{180m^4} M_4, \quad M_4 = \max_{a \leq x \leq b} |f^{(4)}(x)|. \quad (4.21)$$

Для практической оценки погрешности квадратурной формулы можно использовать правило Рунге. Для этого проводят вычисления на сетках с шагом h и $h/2$, получают приближенные значения интеграла J_h и за окончательные значения интеграла принимают величины:

$J_{h/2} + (J_{h/2} - J_h)/3$ – для формулы трапеций;

$J_{h/2} + (J_{h/2} - J_h)/15$ – для формулы Симпсона.

При этом за погрешность приближенного значения интеграла принимают величину:

$|J_{h/2} - J_h|/3$ – для формулы трапеций;

$|J_{h/2} - J_h|/15$ – для формулы Симпсона.

5. Квадратурные формулы Гаусса. В квадратурных формулах типа формулы Гаусса

$$\int_a^b p(x)f(x)dx \approx \sum_{k=1}^n A_k f(x_k) \quad (4.22)$$

коэффициенты $A_1, \dots, A_n$ и узлы $x_1, \dots, x_n$ (обратим внимание, что здесь для удобства изложения нумерация узлов начинается с $k=1$) подбирают так, чтобы приближенное равенство (4.22) было точным для всех многочленов возможной наивысшей степени N . При $p(x) \leq 0$, $a \leq x \leq b$, $N = 2n - 1$ установлено ([4]), что коэффициенты и узлы определяются однозначно: квадратурная формула (4.22) должна быть интерполяционной, а многочлен $\omega_n(x)$ должен быть ортогонален по весу $p(x)$ на $[a, b]$ ко всем многочленам степени меньше n .

Квадратурная формула Гаусса

$$\int_{-1}^1 f(t)dt \approx \sum_{k=1}^n A_k f(t_k) \quad (4.23)$$

имеет своими узлами корни многочлена Лежандра

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n (x^2 - 1)}{dx^n}.$$

При вычислении интеграла $\int_a^b f(x)dx$ следует сделать замену переменной интегрирования

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

Тогда

$$\int_a^b p(x)f(x)dx \approx \frac{b-a}{2} \sum_{k=1}^n A_k f(x_k) \quad (4.24)$$

где

$$x_k = \frac{b+a}{2} + \frac{b-a}{2}t_k. \quad (4.25)$$

Пример 1. При $m = 10$ вычислить интеграл $\int_2^3 \frac{\cos x}{x} dx$ по составной формуле Симпсона.

Решение. Вычисляем

$$f^{(4)}(x) = \frac{24 \cos x}{x^5} + \frac{24 \sin x}{x^4} - \frac{12 \cos x}{x^3} - \frac{4 \sin x}{x^2} + \frac{\cos x}{x}.$$

Поскольку $\cos x < 0$, $\sin x > 0$ при $x \in [2, 3]$, для производной $f^{(4)}(x)$ получаем $-5/4 \leq f^{(4)}(x) \leq 3$, т.е. $|f^{(4)}(x)| \leq 3$. Из неравенства (4.21) при $m = 10$ следует

$$|r_{10}| \leq \frac{1}{180 \cdot 10^4} \cdot 3 < 1/67 \cdot 10^{-6}.$$

Это означает, что если вести вычисления с 6 цифрами после десятичной точки, то результат будет иметь пять верных цифр после десятичной точки. Вычисления можно организовать следующим образом: $a = 2$, $b = 3$, $h = (3 - 2)/10 = 0.1$. Составим табл. 4.1 (Обратите внимание, что аргумент a функции $\cos x$ задан в радианах).

Таблица 4.1

k	$x_k = a + kh$	$\cos x_k$	$f_k = f(x_k)$		
			$k = 0, 10$	$k = 1(2)9$	$k = 2(2)8$
0	2.0	-0.416147	-0.208074		
1	2.1	-0.504846		-0.240403	
2	2.2	-0.588501			-0.267500
3	2.3	-0.666276		-0.289685	
4	2.4	-0.737394			-0.307248
5	2.5	-0.801144		-0.320458	
6	2.6	-0.856889			-0.329573
7	2.7	-0.904072		-0.334841	
8	2.8	-0.942222			-0.336508
9	2.9	-0.970958		-0.334813	
10	3.0	-0.989992	-0.329997		
Суммы			-0.538071	-1.520200	-1.240829

По составной формуле Симпсона находим

$$\int_2^3 \frac{\cos x}{x} \approx \frac{1}{3 \cdot 10} (-0.538071 + 2(-1.240829) + 4(-1.520200)) = -0.30335.$$

Пример 2. При $n = 5$ вычислить интеграл примера 1 по формуле Гаусса.

Решение. Сделаем замену переменной

$$x = \frac{3-2}{2}t + \frac{3+2}{2} = \frac{1}{2}t + \frac{5}{2}.$$

Получим интеграл

$$\int_2^3 \frac{\cos x}{x} dx \approx \frac{1}{2} \int_{-1}^1 \frac{\cos(t/2 + 5/2)}{t/2 + 5/2} dt.$$

Составляем таблицу значений подынтегральной функции (табл. 4.2).

Таблица 4.2

k	t_k	x_k	f_k	A_k	$A_k \cdot f_k$
1	-0.906180	2.046910	-0.223912	0.236927	-0.053051
2	-0.538469	2.230766	-0.274835	0.478629	-0.131544
3	0	2.5	-0.320457	0.568889	-0.182304
4	0.538469	2.769234	-0.336364	0.478629	-0.160994
5	0.906180	2.953090	-0.332630	0.236927	-0.078809
Сумма					-0.606702

По формуле Гаусса при $n = 5$ находим

$$\int_2^3 \frac{\cos x}{x} dx \approx \frac{1}{2} (-0.606702) = -0.303351.$$

Т е м а 5. Численные методы решения задачи Коши для обыкновенных дифференциальных уравнений

Литература: [1] гл. 7; [2] л. 7; [4] гл. 8, п.1–5; [5] ч. 2, гл. 6, п.1, 2; [6] гл. 14, п.14.1–14.10; [7] гл. 9, п.1–6; [8] гл. 8, п.1–8.

Основные теоретические сведения

✓ 1. **Постановка задачи.** Пусть требуется найти численные методы решения задачи Коши

$$y' = f(x, y), x \in [a, b] \quad (5.1)$$

$$y(a) = y_0, \quad (5.2)$$

где y_0 задано.

✓ Численные методы – это алгоритмы вычисления приближенных значений искомой функции на некотором конечном множестве точек. Решение получается при этом в виде таблицы. Численные методы не позволяют найти общее решение уравнения (5.1). С их помощью можно определить лишь частное решение задачи (5.1, 5.2). Необходимо, чтобы задача была хорошо обусловлена (устойчива), т.е. малые изменения в задании исходных данных приводили к достаточно малым изменениям искомого решения.

✓ 2. **Построение численных алгоритмов** опирается на дискретизацию задачи. Введем в область расчета $x \in [a, b]$ дискретный набор точек

$$x_k = a + kh, 0 \leq k \leq n-1, x_0 = a, x_n = b, h = (b-a)/n,$$

в которых будет вычисляться приближенное решение. Точки x_k называются **узлами интегрирования**, или **узлами сетки**, расстояние h между ними – **шагом интегрирования**, или **шагом сетки**, а совокупность узлов $\{x_k\}_{k=0}^n$ – **сеточной областью**, или **сеткой узлов**. Совокупность величин $\{y_k\}_{k=0}^n$, отнесенных к узлам сетки, будем называть **сеточными функциями**.

3. **Использование формулы Тейлора.** Наиболее простым способом построения решения в точке x_{k+1} , если оно известно в точке x_k , является способ, основанный на разложении в ряд Тейлора (в предположении надлежащей дифференцируемости решения):

$$y_{k+1} = y(x_k + h) = y(x_k) + \frac{y'(x_k)}{1!}h + \frac{y''(x_k)}{2!}h^2 + \dots + \frac{y^{(m)}(x_k)}{m!}h^m + R_{m+1}. \quad (5.3)$$

Если заменить $y(x_k)$ приближенным значением y_k , отбросить остаточный член R_{m+1} , то получим явный одношаговый метод, т.е. такой метод, который позволяет найти приближенное значение решения заданной задачи в узле x_{k+1} по информации об этом решении лишь в одной предыдущей узловой точке x_k :

$$y_{k+1} \cong \sum_{p=0}^m \frac{1}{p!} f^{(p)}(x_k, y_k) h^p \quad (5.4)$$

где $y'_k = f(x_k, y_k)$.

Для $m = 1$ и $m = 2$ имеем формулы

$$y_{k+1} = y_k + h f(x_k, y_k) \text{ (метод Эйлера),}$$

$$y_{k+1} = y_k + h \left(f(x_k, y_k) + \frac{h}{2} (f_x(x_k, y_k) + f_y(x_k, y_k) f(x_k, y_k)) \right),$$

по каждой из которых при заданном y_0 можно последовательно получить приближенное решение $\{y_k\}_{k=0}^n$.

4. Линейные многошаговые методы. После того как y_{k+1} получено каким-либо методом, приближение y_k уже не используется в дальнейших расчетах. Один из способов построения методов, использующих некоторые предварительные вычисленные значения $y_k, y_{k-1}, y_{k-2}, \dots$, заключается в следующем. Интегрируя (5.1), получаем тождество

$$y(x+h) - y(x) = \int_x^{x+h} f(t, y(t)) dt, \quad (5.5)$$

где x и $x+h$ — некоторые точки отрезка $[a, b]$. Теперь заменим $f(t, y(t))$ интерполяционным многочленом, принимающим значения $f_k = f(x_k, y_k)$ на множестве точек x_k , в которых y_k уже вычислены или тут же вычисляются. Проинтегрировав этот многочлен, из соотношения (5.5) можно получить различные формулы,

которые определяются положением точек x и $x+h$ относительно узлов интерполяции $x_k, x_{k-1}, \dots, x_{k-m}$ (Адамса, Адамса-Бэшворта).

5. Метод Эйлера. Простейшим и исторически первым численным методом решения задачи Коши (5.1), (5.2) является метод Эйлера. Его можно получить, если в приближенном равенстве (5.4) оставить только два первых слагаемых (т.е. взять $p=1$). Тогда формула (5.4) примет вид

$$y_{k+1} = y_k + hf(x_k, y_k). \quad (5.6)$$

Геометрическая интерпретация одного шага метода Эйлера заключается в аппроксимации решения на отрезке $[x_k, x_{k+1}]$ касательной $y = y_k + y'(x_k)(x - x_k)$, проведенной в точке (x_k, y_k) к интегральной кривой, проходящей через эту точку. После выполнения n шагов неизвестная интегральная кривая заменяется ломаной линией (*ломаная Эйлера*), для которой угловой коэффициент очередного k -го звена равен значению $f(x_k, y_k)$. Для метода Эйлера погрешность аппроксимации имеет вид $R_{k+1} = \frac{h^2}{2} y''(\xi_k)$, где $x_k \leq \xi_k \leq x_{k+1}$.

6. Метод Рунге-Кутты четвертого порядка точности, который является одним из самых распространенных методов решения задач (5.1), (5.2), описывается следующими шестью соотношениями:

$$y_{k+1} = y_k + \Delta y_k; \quad (5.7)$$

$$\Delta y_k = \frac{1}{6}(K_1^{(k)} + 2K_2^{(k)} + 2K_3^{(k)} + K_4^{(k)}); \quad (5.8)$$

$$K_1^{(k)} = hf(x_k, y_k),$$

$$K_2^{(k)} = hf(x_k + h/2, y_k + K_1^{(k)}/2); \quad (5.9)$$

$$K_3^{(k)} = hf(x_k + h/2, y_k + K_2^{(k)}/2);$$

$$K_4^{(k)} = hf(x_k + h, y_k + K_3^{(k)}).$$

При однократном использовании метода значения функции $f(x, y)$ необходимо вычислять четыре раза с аргументами:

$$x_k \text{ и } y_k, x_k + h/2 \text{ и } y_k + K_1^{(k)}/2, \\ x_k + h/2 \text{ и } y_k + K_2^{(k)}/2, x_k + h \text{ и } y_k + K_3^{(k)}.$$

Шаг сетки можно менять при переходе от одной точки к другой. Для контроля правильности выбора шага h вычисляют дробь

$$\Theta = \left| \frac{K_2^{(k)} - K_3^{(k)}}{K_1^{(k)} - K_2^{(k)}} \right|.$$

Величина Θ не должна превышать нескольких сотых, в противном случае шаг следует уменьшить. В практике для контроля вычислений применяют двойной пересчет, т.е. сначала вычисляют решение с шагом h , затем с шагом $h/2$. В заданных точках приближенное решение должно совпадать в пределах заданной точности. Погрешность метода на каждом шаге имеет порядок h^5 . Методы Эйлера и Рунге-Кутты легко распространяются на системы дифференциальных уравнений первого порядка.

Пример 1. Применяя метод Эйлера, найти на отрезке $[0; 0.5]$ решение дифференциального уравнения $y' = y + x^2$ с начальным условием $y(0) = 1$, выбрав шаг $h = 0.1$.

Решение.

1-й шаг. По начальным данным в табл. 5.1 заполняем вторую и третью графы первой строки ($k = 0$).

2-й шаг. Из уравнения $y'_k = y_k + x_k^2$ вычисляем $f_k = f(x_k, y_k)$ и вписываем значение в четвертую графу.

3-й шаг. Значение четвертой графы умножаем на величину h (вычисляем $\Delta y_k = h f_k$) и полученное значение записываем в пятую графу этой же строки.

4-й шаг. К значению третьей графы прибавляем значение пятой графы этой же строки (вычисляем $y_{k+1} = y_k + \Delta y_k$) и результат записываем в третью графу следующей строки. Определяем $x_{k+1} = x_0 + k h$ и результат записываем во вторую графу. Затем шаги 2, 3, 4 повторяем до тех пор, пока не будет пройден весь отрезок $[0; 0.5]$.

Таблица 5.1

k	x_k	y_k	$f_k = y_k + x_k^2$	$\Delta y_k = h \cdot f_k$
0	0	1.000000	1.000000	0.100000
1	0.1	1.100000	1.110000	0.111000
2	0.2	1.211000	1.251000	0.125100
3	0.3	1.336100	1.426100	0.142610
4	0.4	1.478710	1.638710	0.163871
5	0.5	1.642581		

З а м е ч а н и е. Точное решение уравнения

$$y' = y + x^2,$$

удовлетворяющее условию $y(0) = 1$, выражается формулой

$$y = 3e^x - x^2 - 2x - 2,$$

следовательно,

$$y(0.5) = 3e^{0.5} - (0.5)^2 - 2 \cdot 0.5 - 2 \cong 1.696164.$$

Т е м а 6. Разностные схемы для уравнений с частными производными

Литература: [1] гл. 8; [2] л. 9, 10; [3] ч. 3, гл. 9, п.1, 2, 5; [4] гл. 10, п.1, 2, 5, 6; [5] ч. 3, гл. 1, п.1–4, 6; [6] гл. 15, п.15.1–15.3; [7] гл. 11; [8] гл. 10, п.1–7.

Основные теоретические сведения

- 1. Основные понятия.** Для простейших краевых задач, сформулированных в различных курсах математической физики, приводятся некоторые точные решения. В то же время большинство даже линейных уравнений, описывающих реальные процессы, таковы, что не допускают построения точного решения с помощью элементарных функций. В таких случаях прибегают к приближенным методам. Обычно рассматривают два вида приближенных решений: аналитические и численные. Рассмотрим численные методы, основанные на разностной аппроксимации производных. Такой подход называют *разностным методом, методом конечных разностей или методом сеток*.

Пусть требуется найти решение дифференциальной задачи

$$Lu(x, y) = f(x, y), \quad (6.1)$$

поставленной в некоторой области Ω с контуром $\partial\Omega$.

Разностный метод решения краевой задачи (6.1) можно представить в виде двух этапов:

- 1) построение разностной схемы, аппроксимирующей данную непрерывную задачу;
- 2) получение решения разностной задачи и оценка погрешности этого решения.

При построении разностной схемы *первым шагом* является замена области $\bar{\Omega} = \Omega + \partial\Omega$ непрерывного изменения аргументов областью дискретного их изменения – *сеточной областью* $\bar{\Omega}_h$ (или просто *сеткой*), т.е. множеством точек (x_i, y_j) называемых *узлами сетки*. Например, для квадрата $\bar{\Omega}^0$ сеточную область можно построить следующим образом. Проведем прямые

$$x_i = ih, y_j = jh \quad (h = 1/n; 0 \leq i, j \leq n).$$

Множество точек пересечения (x_i, y_j) этих прямых и составит сеточную область $\bar{\Omega}_h^0$, а сами точки образуют узлы сетки. Вместо функций непрерывного аргумента рассматривают функции, определенные только в узлах сетки, – *сеточные функции*. *Второй шаг* в построении разностной схемы состоит в аппроксимации производных, которые входят в дифференциальное уравнение и краевые условия, их разностными аналогами – линейными комбинациями значений сеточных функций в некоторых узлах сетки, называемых *шаблоном*. В результате краевую задачу заменяют *дискретной краевой задачей (разностной схемой)*, представляющей собой систему конечного числа линейных алгебраических уравнений. Решение разностной схемы (предполагается, что оно существует) принимают за приближенное решение краевой задачи. Следует иметь в виду, что для одной краевой задачи можно построить большое число различных разностных схем, среди которых далеко не все пригодны для использования на практике.

Пусть u_h^* – точное решение задачи (6.1) в узлах сетки $\bar{\Omega}_h$. Как правило, вычислить u_h^* не удастся. Поэтому находят сеточную функцию $u_h \approx u_h^*$ как решение разностной задачи

$$L_h u_h = f_h. \quad (6.2)$$

Обозначим линейное нормированное пространство, образованное функциями u_h , через U_h , а пространство, образованное функциями f_h , через F_h . Будем считать, что в пространствах U_h , F_h введены нормы $\|\cdot\|_{U_h}$, $\|\cdot\|_{F_h}$. Если

$$\|u_h^* - u_h\|_{U_h} \rightarrow 0 \text{ при } h \rightarrow 0, \quad (6.3)$$

то говорят, что разностная схема является сходящейся.

Если для всех $h < h_0$ имеет место неравенство

$$\|u_h^* - u_h\|_{U_h} \leq Ch^k \quad (C = \text{const}), \quad (6.4)$$

то говорят, что имеет место сходимость порядка k относительно h .

Исследование сходимости разностной схемы проводят в два этапа:

- 1) устанавливают аппроксимацию задачи (6.1) разностной схемы (6.2);
- 2) проверяют устойчивость разностной схемы.

Говорят, что разностная схема (6.2) аппроксимирует задачу (6.1), если

$$\begin{aligned} L_h u_h^* &= f_h + \delta f_h, \\ \|\delta f_h\|_{F_h} &\rightarrow 0 \text{ при } h \rightarrow 0. \end{aligned} \quad (6.5)$$

Величину δf_h называют погрешностью аппроксимации. Если $\|\delta f_h\|_{F_h} \leq Mh^s$, то говорят, что разностная схема (6.2) аппроксимирует задачу (6.1) на решении u_h^* с порядком s относительно h .

Разностная схема (6.2) называется устойчивой, если существует такое $h_0 > 0$, что при всех $h < h_0$ и любых $f_h \in F_h$:

- 1) разностная схема (6.2) имеет единственное решение;
- 2) $\|u_h\|_{U_h} \leq K\|f_h\|_{F_h}$, где K — постоянная относительно h и f_h .

Если разностная схема (6.2) аппроксимирует задачу (6.1) с порядком s относительно h и устойчива, то будет выполняться (6.4) при $k = s$.

2. Метод сеток решения смешанной задачи для уравнения теплопроводности. Пусть требуется найти функцию $u(x, t)$, удовлетворяющую в области $\bar{\Omega} = \{(x, t) | a \leq x \leq b, 0 \leq t \leq T\}$, уравнению

$$Lu \equiv u_t - a^2 u_{xx} = f, \quad (a^2 = \text{const} > 0), \quad (6.5)$$

начальному условию

$$u(x, 0) = \varphi(x) \quad (a \leq x \leq b) \quad (6.6)$$

и краевым условиям первого рода

$$u(a, t) = \psi_a(t) \quad (0 \leq t \leq T), \quad (6.7)$$

$$u(b, t) = \psi_b(t) \quad (0 \leq t \leq T).$$

Здесь $f(x, t)$, $\varphi(x)$, $\psi_a(t)$, $\psi_b(t)$ — известные функции, причем

$$\psi_a(0) = \varphi(a), \quad \psi_b(0) = \varphi(b) \quad (\text{условия согласования}).$$

Будем считать, что задача (6.5)–(6.7) имеет единственное, достаточно гладкое решение. Построим в области

$$\bar{\Omega} = \{(x, t) \mid a \leq x \leq b, 0 \leq t \leq T\}$$

равномерную прямоугольную сетку $\Omega_h = \{x_i, t_k\}$ с шагом τ по оси t и шагом $h = (b - a)/n$ по оси x :

$$x_i = x_0 + ih, \quad x_0 = a, \quad x_n = b, \quad (0 \leq i \leq n);$$

$$t_k = k\tau, \quad (0 \leq k \leq m).$$

Покажем, как несущественные с первого взгляда различия в аппроксимации дифференциального уравнения приводят к разностным схемам с весьма различными свойствами. С этой целью приведем две разностные схемы для численного интегрирования (6.5)–(6.7), которые разнятся лишь аппроксимацией производной по времени. Применяя последовательно формулы

$$\frac{\partial u(x, t)}{\partial t} \approx \frac{u_h(x, t + \tau) - u_h(x, t)}{\tau}, \quad (6.8)$$

$$\frac{\partial u(x, t)}{\partial t} \approx \frac{u_h(x, t) - u_h(x, t - \tau)}{\tau} \quad (6.9)$$

для аппроксимации $\partial u / \partial t$ разностными отношениями и формулу

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{u_h(x + h, t) - 2u_h(x, t) + u_h(x - h, t)}{h^2} \quad (6.10)$$

для аппроксимации $\partial^2 u / \partial x^2$, получаем два варианта разностных уравнений:

вариант 1 для расчетной точки $(x, t) = (x_i, t_k)$

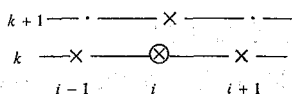
$$L_h u_h \equiv \frac{u_i^{k+1} - u_i^k}{\tau} - a^2 \frac{u_{i+1}^k - 2u_i^k + u_{i-1}^k}{h^2} = f_i^k; \quad (6.11)$$

вариант 2 для расчетной точки (x_i, t_{k+1})

$$L_h u_h \equiv \frac{u_i^{k+1} - u_i^k}{\tau} - a^2 \frac{u_{i+1}^{k+1} - 2u_i^{k+1} + u_{i-1}^{k+1}}{h^2} = f_i^{k+1}, \quad (6.12)$$

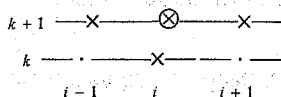
где введены обозначения $u_i^k = u(x_i, t_k)$, $f_i^k = f(x_i, t_k)$

Вариантам 1, 2 отвечают следующие шаблоны, в которых кружком обозначена расчетная точка. Оба шаблона являются четырехточечными двухслойными.



вариант 1

(шаблон явной схемы)



вариант 2

(шаблон неявной схемы)

Начальные и граничные условия для первой краевой задачи аппроксимируются точно:

$$u_i^0 = \varphi(x_i) \quad (0 \leq i \leq n), \quad (6.13)$$

$$u_0^k = \psi_a(t_k) \quad (0 \leq k \leq m), \quad (6.14)$$

$$u_n^k = \psi_b(t_k) \quad (0 \leq k \leq m).$$

В случае второй и третьей краевых задач граничные условия аппроксимируются на основе формул, аналогичных соотношениям (6.8), (6.9). В силу условий согласования $u_0^0 = \varphi(a)$ и $u_0^0 = \psi_a(0)$, а также $u_n^0 = \varphi(b)$ и $u_n^0 = \psi_b(0)$ совместны.

Учитывая (6.13), (6.14), мы будем иметь $(n-1)m$ неизвестных значений u_i^k ($1 \leq i \leq n-1$, $0 \leq k \leq m-1$), так что число уравнений оказывается равным числу неизвестных.

Рассмотрим вопрос о том, как решать системы линейных алгебраических уравнений, получившихся в обоих вариантах.

В двухслойном шаблоне для варианта 1 в верхнем слое расположена всего лишь одна точка. Следовательно, если мы уже знаем сечение u^k , то из уравнений (6.11) сможем найти u^{k+1} . Для этого достаточно преобразовать уравнение (6.11) к виду

$$u_i^{k+1} = cu_{i-1}^k + (1-2c)u_i^k + cu_{i+1}^k + \tau f_i^k \quad (1 \leq i \leq n-1), \quad (6.15)$$

где $c = a^2 \tau / h^2$.

При счете по формуле (6.15) для $i = 1$ и $i = n-1$ будут использованы значения u_0^k и u_n^k , полученные в (6.14) из граничных условий. Поскольку нулевое сечение u_i^0 известно, счет по формулам (6.15) позволяет определить u_i^k для всех нужных пар (i, k) .

Таким образом, использование варианта 1 полностью освобождает от необходимости решать системы линейных уравнений, сводя дело к счету по рекуррентным формулам (6.15). В связи с тем, что в уравнениях (6.15) составляющие u^{k+1} явно выражаются через значения составляющих u^k , схему варианта 1 называют *явной*.

Уравнения (6.12) при первом взгляде мало отличаются от уравнений (6.11). Однако шаблон варианта 2 весьма существенно разнится от шаблона варианта 1. В шаблоне варианта 2 в верхнем слое расположены три точки, а не одна, как в шаблоне варианта 1. В связи с этим из формул (6.12) не удастся извлечь явного выражения для составляющих u^{k+1} через составляющие u^k . Преобразуем уравнения (6.12) к виду

$$-cu_{i-1}^{k+1} + (1+2c)u_i^{k+1} - cu_{i+1}^{k+1} = u_i^k + \tau f_i^{k+1} \quad (1 \leq i \leq n-1). \quad (6.16)$$

С учетом того, что значения u_0^{k+1} и u_n^{k+1} известны из (6.14), мы имеем дело с системой $n-1$ линейных уравнений с $n-1$ неизвестными u_i^{k+1} ($1 \leq i \leq n-1$). Для того чтобы от k -го слоя продвинуться к $(k+1)$ -му, следует решить эту систему. В связи с этим схему варианта 1 называют *неявной*. Зная из (6.13) значения u^0 и на каждом слое u_0^k и u_n^k , последовательно находим $u^1, u^2, \dots, u^m$. Система (6.16) имеет трехдиагональную матрицу, а для решения этой системы удобно использовать *метод разностной прогонки*.

Погрешность аппроксимации задачи (6.5)–(6.7) разностной схемой вариантов 1 и 2 оценивается величиной

$$\frac{\tau}{2} M_t^{(2)} + a^2 \frac{h^2}{12} M_x^{(4)},$$

где

$$M_t^{(2)} = \max_{x,t} \left| \frac{\partial^2 u}{\partial t^2} \right|, \quad M_x^{(4)} = \max_{x,t} \left| \frac{\partial^4 u}{\partial x^4} \right|.$$

Явная разностная схема устойчива при $c \leq 1/2$ и в этом случае

$$|u^k - u(x_i, t_k)| \leq T \left(\frac{\tau}{2} M_t^{(2)} + a^2 \frac{h^2}{12} M_x^{(4)} \right).$$

Условие $c \leq 1/2$, необходимое для устойчивости явной разностной схемы, может вызвать неоправданно большое число шагов по t . Неявная разностная схема устойчива при любых конечных значениях c .

3. Метод прогонки. Перепишем систему (6.14), (6.16) следующим образом (для одного временного слоя, при этом опуская верхний индекс):

$$\begin{aligned} u_0 &= \gamma_0, \\ A_i u_{i-1} + B_i u_i + C_i u_{i+1} &= F_i \quad (1 \leq i \leq n-1), \\ u_n &= \gamma_n, \end{aligned} \quad (6.17)$$

где

$$\begin{aligned} \gamma_0 &= \psi_a(t_k), \quad \gamma_n = \psi_b(t_k), \\ A_i &= -c, \quad B_i = 1 + 2c, \quad C_i = -c, \quad F_i = u_i^k + \tau f_i^{k+1}. \end{aligned}$$

Решение (6.17) по методу правой прогонки, состоит в следующем:

- принимаем $L_0 = 0, K_0 = \gamma_0$;
- по рекуррентным формулам

$$L_i = -C_i / (A_i L_{i-1} + B_i), \quad K_i = (F_i - A_i K_{i-1}) / (A_i L_{i-1} + B_i) \quad (6.18)$$

последовательно вычисляем $L_1, K_1; L_2, K_2; \dots; L_{n-1}, K_{n-1}$;

- зная $u_n = \gamma_n$, по формуле

$$u_i = L_i u_{i+1} + K_i \quad (6.19)$$

вычисляем $u_{n-1}, u_{n-2}, \dots, u_0$.

Первые два шага называют **прямой прогонкой**, последний – **обратной прогонкой**. Они выполнены при условии, что

$$A_i L_{i-1} + B_i \neq 0.$$

Пример 1. Найти решение уравнения

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad 0 \leq x \leq 1, \quad 0 \leq t \leq 0.05,$$

удовлетворяющее условиям

$$u(x, 0) = x(1+x), \quad 0 \leq x \leq 1,$$

$$u(0, t) = 2t, \quad u(1, t) = t + 2, \quad 0 \leq t \leq 0.05,$$

по явной разностной схеме, взяв $h = 0.2$, $\tau = 0.01$.

Решение. Вычислим $c = \tau/h^2 = 0.25$ ($a^2 = 1$). Явная разностная схема записывается в виде

$$u_i^{k+1} = 0.25(u_{i-1}^k + u_{i+1}^k) + 0.5u_i^k,$$

$$u_i^0 = x_i(1+x_i),$$

$$u_0^k = 2t_k, \quad u_5^k = t_k + 2, \quad 1 \leq i \leq 4, \quad 0 \leq k \leq 5.$$

Получаем начальные условия

$$u_1^0 = 0.24; \quad u_2^0 = 0.56; \quad u_3^0 = 0.96; \quad u_4^0 = 1.44$$

и граничные условия $u_0^0 = 0.00; \quad u_5^0 = 2.00$.

На первом шаге имеем $u_0^1 = 0.02; \quad u_5^1 = 2.01$;

$$u_1^1 = 0.25(u_0^0 + u_2^0) + 0.5u_1^0 = 0.26; \quad u_2^1 = 0.25(u_1^0 + u_3^0) + 0.5u_2^0 = 0.58;$$

$$u_3^1 = 0.25(u_2^0 + u_4^0) + 0.5u_3^0 = 0.98; \quad u_4^1 = 0.25(u_3^0 + u_5^0) + 0.5u_4^0 = 1.46.$$

Аналогичным образом проводятся вычисления и на последующих временных сечениях (слоях). Результаты вычислений приведены в табл. 6.1 с тремя десятичными знаками.

Пример 1. Найти решение задачи примера 1 по неявной разностной схеме, взяв $h = 0.2$, $\tau = 0.05$.

Решение. Вычислим $c = \tau/h^2 = 1.25$. При решении задачи по неявной разностной схеме получаем систему уравнений

$$-1.25u_{i-1}^1 + 3.5u_i^1 - 1.25u_{i+1}^1 = u_i^0,$$

$$u_i^0 = x_i(1+x_i),$$

$$u_0^1 = 0.1, \quad u_5^1 = 2.05, \quad 1 \leq i \leq 4.$$

Начальные условия те же, что в примере 1. Систему решаем методом прогонки. Здесь $A_i = C_i = -1.25$, $B_i = 3.5$, $L_0 = 0$, $K_0 = 0.1$. По формулам (6.18) последовательно вычисляем прогоночные коэффициенты:

$$L_1 = 0.357, K_1 = 0.104; L_2 = 0.409, K_2 = 0.226;$$

$$L_3 = 0.418, K_3 = 0.416; L_4 = 0.420, K_4 = 0.658.$$

Зная $u_3^1 = 2.05$, по формулам (6.19) находим решение системы:

$$u_0^1 = 0.1, u_1^1 = 0.338, u_2^1 = 0.656, u_3^1 = 1.051, u_4^1 = 1.519.$$

Таблица 6.1

$t_k \backslash x_i$		0.0	0.2	0.4	0.6	0.8	1.0
	$k \backslash i$	0	1	2	3	4	5
0.00	0	0.00	0.24	0.56	0.96	1.44	2.00
0.01	1	0.02	0.26	0.58	0.98	1.46	2.01
0.02	2	0.04	0.28	0.60	1.00	1.478	2.02
0.03	3	0.06	0.30	0.62	1.020	1.494	2.03
0.04	4	0.08	0.32	0.64	1.038	1.510	2.04
0/05	5	0.10	0.34	0.660	1.056	1.524	2.05

СОДЕРЖАНИЕ

	Стр.
Предисловие.....	3
Литература.....	4
Методические указания по разделам.....	5
Тема 1. Элементы теории погрешностей.....	5
Тема 2. Численные методы решения СЛАУ.....	8
Тема 3. Приближение функций.....	20
Тема 4. Численное дифференцирование и интегрирование.....	28
Тема 5. Численные методы решения задачи Коши для обыкновенных дифференциальных уравнений.....	38
Тема 6. Разностные схемы для уравнений с частными производными.....	42

Учебное издание

УЧЕБНО-МЕТОДИЧЕСКОЕ ПОСОБИЕ ДЛЯ ВЫПОЛНЕНИЯ КОНТРОЛЬНОЙ РАБОТЫ

по дисциплине

«МАТЕМАТИКА»

Раздел – Вычислительная математика

Составитель: Валентин Николаевич Веретенников

Редактор И. Г. Максимова

ЛР №020309 от 30.19.96

