

Черкасов М.А.
Численные методы. Решение задач: Учеб. пособие. – М.: Изд-во МАИ, 2007. – 92с.: ил.

В работе представлено содержание курса лекций и практических занятий, основную часть которых автор читает студентам дневного и вечернего отделений факультетов «Авиационная техника», «Двигатели летательных аппаратов», «Прикладная механика» МАИ. Часть материала преподаётся на спецкурсе по программированию для школьников 11 класса и бакалавриате факультета «Прикладная математика и физика». Курс ориентирован на один семестр обучения.

Работа может быть полезна студентам при выполнении практических, лабораторных, курсовых работ и для самообразования.

Рецензенты:

кафедра общей и прикладной математики МГИУ (зав. кафедры д-р физ.-мат. наук проф. Е.А.Пушкарь);
д-р физ.-мат. наук проф. Э.М.Карташов

ISBN № 978-5-7035-1851-9 © Московский авиационный институт
(государственный технический университет), 2007

Предисловие. О погрешностях

Зачастую в вычислительной практике мы имеем дело не с точными величинами, которые получить иногда и невозможно, а с их приближёнными значениями.

Пусть A – точное значение некоторой величины;
 a – приближённое значение той же величины.

Определение 1. Будем называть *абсолютной погрешностью* приближённого числа a число Δ_a (A – точное значение):

$$\Delta_a = |A - a| \quad (1)$$

Пример 1. Длина комнаты $a = 5,43$ м (с точностью 1 см), ширина комнаты $b = 3,82$ м. Оценить абсолютную погрешность Δ_S площади комнаты $S_0 = a \cdot b = 20,7426$ м².

Решение. $\Delta_a = 0,01$ м, $\Delta_b = 0,01$ м Тогда

$$S_{\max} = (a + \Delta_a) \cdot (b + \Delta_b) = 20,8352 \text{ м}^2;$$

$$S_{\min} = (a - \Delta_a) \cdot (b - \Delta_b) = 20,6502 \text{ м}^2.$$

Сравнивая их с S_0 , получаем $|S - S_0| < 0,0926$, следовательно, $\Delta_S = 0,0926$ м². Тогда после округления в большую сторону $\Delta_S = 0,1$ м².

Определение 2. Относительной погрешностью δ_a приближённого числа a ($a \neq 0$) будем называть

$$\delta_a = \frac{\Delta_a}{|a|} \quad (2)$$

Относительную погрешность обычно выражают в процентах.

Пример 2. Для предыдущего примера $\delta_S = \Delta_S / S_0 = 926 / 207426 = 0,0045 = 0,45\%$.

Относительной погрешностью принято характеризовать точность вычисленных результатов во многих технических задачах. Относительная погрешность приближённо вычисленного числа связана с *количеством верных знаков* в его записи. Количество верных знаков числа отсчитывается от первой (левой) значащей цифры до первой значащей цифры его абсолютной погрешности. Например, $S = 20,7426$ с $\Delta_S = 0,0926$ имеет три верных знака ($S = 20,7$), остаток сомнителен.

Один верный знак в числе соответствует относительной погрешности приблизительно в 10%; два верных знака – 1%; три – 0,1%.

В табличных данных обычно приводятся величины уже округлённые до верных знаков. Например, $e = 2,718$, следовательно $\Delta_e \leq 0,5 \cdot 10^{-3}$ (половина единицы последнего не записанного знака).

В расчётах оставляют 2-4 сомнительных знака, а в окончательном ответе – один сомнительный знак.

Пример 3. Определить, какое равенство точнее:

$$\text{а) } a_1 = \frac{9}{11} = 0,818 \quad \text{или б) } b_1 = \sqrt{18} = 4,24 ?$$

Решение. 1. Находим более точные значения $a_0=0,81818$, $b_0=4,2426$;
2. вычисляем абсолютные погрешности, округляя их с избытком:

$$\Delta_a = |0,81818 - 0,818| \leq 0,00019,$$

$$\Delta_b = |4,2426 - 4,24| \leq 0,0027.$$

3. вычисляем относительные погрешности:

$$\delta_a = \Delta_a / a_1 = 0,00019 / 0,818 = 0,00024 = 0,024 \% ,$$

$$\delta_b = \Delta_b / b_1 = 0,0027 / 4,24 = 0,00064 = 0,064 \% .$$

Так как $\delta_a \leq \delta_b$, следовательно, $\frac{9}{11} = 0,818$ – более точное равенство.

Пример 4. Округлить сомнительные цифры числа, оставив верные знаки:
а) 72,353 ($\pm 0,026$) ; б) 2,3544 , $\delta = 0,2 \%$.

Решение а. Пусть $a = 72,353 \pm 0,026$, следовательно $\Delta_a = 0,026 < 0,05$, т.е. в числе 72,353 верными являются цифры 7, 2, 3. Ответ: $a = 72,4$.
б. Пусть $a = 2,3544$, $\delta_a = 0,2 \%$, следовательно $\Delta_a = a \cdot \delta_a = 0,00471$, т.е. в числе 2,3544 верными являются три первые цифры. Ответ: $a = 2,35$.

Рассмотрим функцию $z=f(x;y)$. Найдём погрешность Δ_z вычисления z по известным погрешностям для x и y . Используем выражение для полного дифференциала от z . Введём приращение Δ : $\Delta x = x - x_0$, $\Delta y = y - y_0$, а для Δz :

$$\Delta z = \frac{\partial f(x_0; y_0)}{\partial x} \Delta x + \frac{\partial f(x_0; y_0)}{\partial y} \Delta y . \quad (3)$$

Тогда абсолютную погрешность вычисления z можно найти как

$$\Delta_z = \left| \frac{\partial f(x_0; y_0)}{\partial x} \right| \Delta x + \left| \frac{\partial f(x_0; y_0)}{\partial y} \right| \Delta y \quad (4)$$

После несложных преобразований получим выражение для δ_z :

$$\delta_z = \left| \frac{\partial f(x_0; y_0)}{\partial x} \cdot \frac{x}{f(x_0; y_0)} \right| \delta_x + \left| \frac{\partial f(x_0; y_0)}{\partial y} \cdot \frac{y}{f(x_0; y_0)} \right| \delta_y , \quad (5)$$

$$\delta_z = \left| \frac{\partial \ln f(x_0; y_0)}{\partial x} \right| \Delta x + \left| \frac{\partial \ln f(x_0; y_0)}{\partial y} \right| \Delta y \quad (6)$$

Пример 5. Пусть $Z=X+Y$, тогда $\Delta_z = \Delta_x + \Delta_y$, $\delta_z = \frac{\Delta_x + \Delta_y}{x + y}$.

Пример 6. Пусть $Z=X-Y$, тогда $\Delta_z = \Delta_x + \Delta_y$, $\delta_z = \frac{\Delta_x + \Delta_y}{x - y}$.

Если $x \approx y$, то погрешность возрастает.

Пример 7. Пусть $Z=X \cdot Y$, тогда $\Delta_z = Y \cdot \Delta_x + X \cdot \Delta_y$, $\delta_z = \delta_x + \delta_y$

Пример 8. Пусть $Z = \frac{x}{y}$, тогда $\Delta_z = \frac{y \cdot \Delta_x + x \cdot \Delta_y}{y^2}$, $\delta_z = y \cdot \delta_x + x \cdot \delta_y$.

Пример 9. Дана длина ребра куба l и абсолютная погрешность её вычисления Δ_l . Найти объём куба V и точность его вычисления (Δ_v и δ_v).

Решение. $V=l^3$, $V'=3l^2$, $\Delta_v = 3l^2 \Delta_l$, $\delta_v = 3\delta_l$.

Пример 10. По известным катетам a и b (и их Δ_a и Δ_b) вычислить гипотенузу c .

Решение. $c = \sqrt{a^2 + b^2}$, $\Delta_c = \frac{a \cdot \Delta_a + b \cdot \Delta_b}{\sqrt{a^2 + b^2}}$, $\delta_c = \frac{a^2}{a^2 + b^2} \delta_a + \frac{b^2}{a^2 + b^2} \delta_b$.

Обычно выделяют погрешности трёх видов:

- погрешность исходных данных;
- погрешность метода;
- погрешность вычислений.

Погрешность исходных данных возникает при неточном задании исходных данных и связана с погрешностями приборов при измерении результатов экспериментов.

Погрешность вычислений связана с количеством, типом и даже порядком выполняемых арифметических операций. Простой пример: попробуйте выполнить на калькуляторе вычисление двух похожих выражений $A=3 \cdot 10^{20} + 2 \cdot 10^{-20} - 3 \cdot 10^{20}$ и $B=3 \cdot 10^{20} - 3 \cdot 10^{20} + 2 \cdot 10^{-20}$, и Вы увидите, что результаты будут разные. С погрешностью вычислений связано и количество цифр в представлении чисел, т.е. с разрядность вычислительной техники. Например, вычисления, выполненные на Фортране с переменными типа *REAL* *8, будут точнее, чем те же вычисления с переменными типа *REAL* *4.

Проиллюстрировать погрешность методов можно на примере.

$$\text{Возьмём разложение } \sin x = \sum_{n=0}^{\infty} \frac{(-1)^n \cdot x^{2n+1}}{(2 \cdot n + 1)!} = x - \frac{x^3}{6} + \frac{x^5}{120} - \dots$$

Для вычисления $\sin x$ рассмотрим две формулы $A_1 = x$ и $A_2 = x - x^3/6$. Формула A_2 будет точнее формулы A_1 . Так для $x=0,1$ $A_1=0,1$, $A_2=0,0998(3)$ и $\sin x = 0,099833417\dots$ Для близкого к нулю x погрешность формулы (метода) A_1 равна $O(x^3)$, а погрешность формулы (метода) A_2 равна $O(x^5)$. Говорят, что формула A_1 третьего порядка точности, а формула A_2 пятого порядка точности.

1. Прямые методы решения первой основной задачи линейной алгебры

Прямыми называют методы, которые при решении поставленной задачи требуют конечное (иногда заранее известное) число арифметических операций. К таким методам при решении первой основной задачи линейной алгебры (вычисление определителя, обратной матрицы, решения систем линейных алгебраических уравнений) относится *метод Гаусса* и его модификация для трёхдиагональной системы – *метод прогонки*. Известны и широко используются другие прямые методы для решения этой задачи: методы Краута, Дулитла (факторизация матриц общего вида), Халецкого (для симметричных матриц) и др.

1.1. Метод Гаусса решения СЛАУ

Рассмотрим схему единственного деления метода Гаусса. Пусть дана система n линейных алгебраических уравнений с n неизвестными (СЛАУ):

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots\dots\dots \dots\dots\dots \dots\dots\dots \dots\dots\dots \dots\dots\dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_n = b_n \end{cases} \quad (1.1.1)$$

В матричной форме эта система выглядит как $A\bar{x}=\bar{b}$, $A=\{a_{ij}\}$, $(i,j=1,\dots,n)$, $\bar{x}=(x_1,x_2,\dots,x_n)^T$, $\bar{b}=(b_1,b_2,\dots,b_n)^T$. Задача (1.1.1) имеет единственное решение, если определитель (детерминант) матрицы системы не равен нулю ($|A|\neq 0$ или $\det A \neq 0$).

Метод Гаусса заключается в исключении из (1.1.1) тех слагаемых, которые лежат в матрице A ниже главной диагонали ($a_{ij}, i > j$). Исключать слагаемые разрешается только с помощью трёх допустимых преобразований:

- 1) любую строку (уравнение) можно умножить (разделить) на любое число, кроме нуля;
- 2) любую строку можно прибавить к другой строке;
- 3) можно переставить любые две строки.

При каждом применении третьего преобразования определитель будет менять свой знак.

Перепишем систему (1.1.1) в виде расширенной матрицы, которую будем преобразовывать к верхней треугольной форме с единицами на главной диагонали и с нулями под главной диагональю:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} & b_2 \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} & b_3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} & b_n \end{array} \right) \Rightarrow \left(\begin{array}{cccc|c} 1 & c_{12} & c_{13} & \dots & c_{1n} & d_1 \\ 0 & 1 & c_{23} & \dots & c_{2n} & d_2 \\ 0 & 0 & 1 & \dots & c_{3n} & d_3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & d_n \end{array} \right). \quad (1.1.2)$$

С помощью допустимых преобразований обнуляются элементы матрицы, лежащие ниже главной диагонали. Это преобразование расширенной матрицы системы (матрица системы с добавленным столбцом правых частей) называют *прямым ходом* метода Гаусса. Правая из (1.1.2) матрица соответствует системе:

[illegible]

При *обратном ходе* вычисляются неизвестные $x_n, x_{n-1}, x_{n-2}, \dots, x_1$. Из последнего уравнения находим x_n , подставляем его в предпоследнее уравнение и получаем x_{n-1} . Так, в обратном порядке находим все неизвестные x .

Пример: решить СЛАУ:

$$\begin{cases} 2x + 5y + z = 1 \\ -x + 2y - 2z = 2 \\ 6x + 2y + z = 3 \end{cases} \quad (1.1.4)$$

Составим из коэффициентов системы расширенную матрицу :

$$\left(\begin{array}{ccc|c} \mathbf{2} & 5 & 1 & 1 \\ -1 & 2 & -2 & 2 \\ 6 & 2 & 1 & 3 \end{array} \right)_{\text{TT}}.$$

Делим всю первую строку на **2** (на a_{11}):

$$\left(\begin{array}{ccc|c} 1 & 5/2 & 1/2 & 1/2 \\ -1 & 2 & -2 & 2 \\ 6 & 2 & 1 & 3 \end{array} \right).$$

Чтобы обнулить **-1** (первый элемент во второй строке), прибавим ко второй строке первую строку

$$\left(\begin{array}{ccc|c} 1 & 5/2 & 1/2 & 1/2 \\ -\mathbf{1}+1 & 2+5/2 & -2+1/2 & 2+1/2 \\ 6 & 2 & 1 & 3 \end{array} \right) \Rightarrow \left(\begin{array}{ccc|c} 1 & 5/2 & 1/2 & 1/2 \\ 0 & 9/2 & -3/2 & 5/2 \\ \mathbf{6} & 2 & 1 & 3 \end{array} \right)$$

Чтобы обнулить **6** ($=a_{31}$ - первый элемент в третьей строке), прибавим к третьей строке первую строку, умноженную на **-6** :

$$\left(\begin{array}{ccc|c} 1 & 5/2 & 1/2 & 1/2 \\ 0 & 9/2 & -3/2 & 5/2 \\ 6+1 \cdot (-6) & 2+5/2 \cdot (-6) & 1+1/2 \cdot (-6) & 3+1/2 \cdot (-6) \end{array} \right) \Rightarrow$$

$$\left(\begin{array}{ccc|c} 1 & 5/2 & 1/2 & 1/2 \\ 0 & 9/2 & -3/2 & 5/2 \\ 0 & -13 & -2 & 0 \end{array} \right)$$

Делим всю вторую строку на **9/2** (на a_{22}) :

$$\left(\begin{array}{ccc|c} 1 & 5/2 & 1/2 & 1/2 \\ 0 & 1 & -1/3 & 5/9 \\ 0 & -13 & -2 & 0 \end{array} \right)$$

Чтобы обнулить **-13** (второй элемент в третьей строке), прибавим к третьей строке вторую строку, домноженную на 13:

$$\left(\begin{array}{ccc|c} 1 & 5/2 & 1/2 & 1/2 \\ 0 & 1 & -1/3 & 5/9 \\ 0 & -13+1 \cdot 13 & -2+(-1/3) \cdot 13 & 0+5/9 \cdot 13 \end{array} \right) \Rightarrow$$

$$\left(\begin{array}{ccc|c} 1 & 5/2 & 1/2 & 1/2 \\ 0 & 1 & -1/3 & 5/9 \\ 0 & 0 & -19/3 & 65/9 \end{array} \right)$$

Делим всю третью строку на **-19/3** (на a_{33}) :

$$\left(\begin{array}{ccc|c} 1 & 5/2 & 1/2 & 1/2 \\ 0 & 1 & -1/3 & 5/9 \\ 0 & 0 & 1 & -65/57 \end{array} \right)$$

Закончился прямой ход метода Гаусса. Эта матрица соответствует системе

$$\begin{cases} x + 5/2 y + 1/2 z = 1/2 \\ y - 1/3 z = 5/9 \\ z = -65/57 \end{cases} \quad (1.1.5)$$

Определитель системы равен произведению тех элементов главной диагонали, на которые мы делили строки: $|A| = 2 \cdot 9/2 \cdot (-19/3) = -57$. Последнее уравнение (1.1.5) сразу даёт $z = -65/57$. Подставляем найденное z во второе уравнение и находим $y = 5/9 + 1/3 \cdot (-65/57) = 10/57$. Теперь в первое уравнение подставим найденные z и y , чтобы найти $x = 1/2 - 1/2 \cdot (-65/57) - 5/2 \cdot 10/57 = 12/19$. Окончательно решение СЛАУ (1.1.4) можно записать:

$$\begin{cases} x = 12/19 \approx 0,6315789 \\ y = 10/57 \approx 0,1754386 \\ z = -65/57 \approx -1,1403509 \end{cases} \quad (1.1.6)$$

Замечание. Если в процессе преобразований на главной диагонали элемент a_{ii} , на который нужно будет делить i -ю строку, равен нулю, то следует поменять местами i -ю строку с нижней строкой j , у которой на месте a_{ji} стоит не ноль. Например, пусть возникла похожая ситуация :

$$\left(\begin{array}{ccc|c} 1 & 5/2 & 1/2 & 1/2 \\ 0 & 0 & 4 & 9 \\ 0 & -3 & 2 & 5 \end{array} \right)$$

Поменяем местами вторую строку с третьей строкой:

$$\left(\begin{array}{ccc|c} 1 & 5/2 & 1/2 & 1/2 \\ 0 & -3 & 2 & 5 \\ 0 & 0 & 4 & 9 \end{array} \right)$$

Теперь вторую строку можно делить на **-3** и продолжать процесс.

Проверка. Чтобы проверить, действительно ли x, y, z из (1.1.6) являются решением исходной системы надо подставить их в исходную систему (1.1-4) во все уравнения и убедиться, что действительно левая часть каждого уравнения почти равна его правой части. Их разность $\bar{b} - A \cdot \bar{x} = \bar{r}$ в общем случае для (1.1-1) называют *невязкой*. Если она не равна нулю, то найденное решение \bar{x} отличается от точного решения x^* на некоторую величину $\bar{e} = x - x^*$, которая называется *погрешностью*. В математике используется величина $\nu(A) = \|A\| \cdot \|A^{-1}\|$, называемая *мерой обусловленности*. Чем больше $\nu(A)$, тем больше \bar{e} зависит от неточности задания матрицы A , вектора b и от погрешности вычисления \bar{x} .

Если определитель $|A|$ матрицы системы близок к нулю, то погрешности вычислений будут сильно влиять на увеличение $\|\bar{e}\|$. Для уменьшения погрешности с помощью перестановок строк на главную диагональ перемещают элемент максимальный по модулю (ведущий) в своём столбце. Погрешность может быть ещё меньше, если искать ведущий элемент не в столбце, а по всей матрице (но в этом случае следует запоминать перестановку номеров столбцов).

В методе Гаусса–Жордана в обратном ходе обнуляются все элементы над главной диагональю. В результате преобразований получается система с диагональной матрицей.

Если у матрицы большой порядок n ($n > 50$), то количество арифметических операций метода Гаусса будет примерно равно $2n^3/3$.

1.2. Вычисление обратной матрицы методом Гаусса

Метод вычисления обратной матрицы во многом схож с вычислением решения СЛАУ. Начинаем с построения расширенной матрицы. Справа от исходной матрицы дописываем единичную, схематично получим $(A|E)$. Далее с помощью допустимых преобразований (умножение строки на любое число, сложение двух строк) преобразуем расширенную матрицу к виду $(E|C)$, т.е. на месте исходной матрицы должна стоять единичная матрица. Тогда справа от неё будет стоять матрица C , которая и будет обратной к матрице A . Если перемножить матрицы A и C , то получится единичная матрица: $A \cdot C = E$, $C \cdot A = E$. Если возникает необходимость применения третьего преобразования (обмен местами двух строк), то оно допустимо с оговоркой, что если поменять в исходной матрице строки n и k , то в обратной матрице поменяются местами столбцы с теми же номерами n и k .

Пример. Найти обратную матрицу для матрицы

$$A = \begin{pmatrix} 2 & 3 & 6 \\ 3 & 6 & 2 \\ 6 & 2 & 8 \end{pmatrix}.$$

Составим расширенную матрицу, присоединив к исходной справа единичную:

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 6 & 1 & 0 & 0 \\ 3 & 6 & 2 & 0 & 1 & 0 \\ 6 & 2 & 8 & 0 & 0 & 1 \end{array} \right).$$

Разделим первую строку на **2** ($a_{11} = 2$):

$$\left(\begin{array}{ccc|ccc} 1 & 3/2 & 3 & 1/2 & 0 & 0 \\ 3 & 6 & 2 & 0 & 1 & 0 \\ 6 & 2 & 8 & 0 & 0 & 1 \end{array} \right).$$

Ко второй строке прибавим первую, умноженную на $-a_{21} = -3$ и получим:

$$\left(\begin{array}{ccc|ccc} 1 & 3/2 & 3 & 1/2 & 0 & 0 \\ 0 & 3/2 & -7 & -3/2 & 1 & 0 \\ 6 & 2 & 8 & 0 & 0 & 1 \end{array} \right).$$

К третьей строке прибавим первую, умноженную на $-a_{31} = -6$:

$$\left(\begin{array}{ccc|ccc} 1 & 3/2 & 3 & 1/2 & 0 & 0 \\ 0 & 3/2 & -7 & -3/2 & 1 & 0 \\ 0 & -7 & -10 & -3 & 0 & 1 \end{array} \right).$$

Разделим вторую строку на **3/2** ($a_{22} = 3/2$):

$$\left(\begin{array}{ccc|ccc} 1 & 3/2 & 3 & 1/2 & 0 & 0 \\ 0 & 1 & -14/3 & -1 & 2/3 & 0 \\ 0 & -7 & -10 & -3 & 0 & 1 \end{array} \right).$$

К третьей строке прибавим вторую, умноженную на $-a_{32} = 7$:

$$\left(\begin{array}{ccc|ccc} 1 & 3/2 & 3 & 1/2 & 0 & 0 \\ 0 & 1 & -14/3 & -1 & 2/3 & 0 \\ 0 & 0 & -128/3 & -10 & 14/3 & 1 \end{array} \right).$$

Разделим вторую строку на **-128/3** ($a_{33} = -128/3$):

$$\left(\begin{array}{ccc|ccc} 1 & 3/2 & 3 & 1/2 & 0 & 0 \\ 0 & 1 & -14/3 & -1 & 2/3 & 0 \\ 0 & 0 & 1 & 15/64 & -7/64 & -3/128 \end{array} \right).$$

Закончился прямой ход метода Гаусса.

Начинаем обратный ход, в котором обнулим все элементы, лежащие выше главной диагонали в левой подматрице. Начнём с обнуления элемента $a_{23} = -14/3$. Для этого ко второй строке прибавим третью, умноженную на $14/3$. В записи правой подматрицы перейдём к десятичным дробям:

$$\left(\begin{array}{ccc|ccc} 1 & 3/2 & 3 & 0,5 & 0 & 0 \\ 0 & 1 & 0 & 0,09375 & 0,15625 & -0,109375 \\ 0 & 0 & 1 & 0,234375 & -0,109375 & -0,0234375 \end{array} \right).$$

Теперь избавимся от $a_{13} = 3$, прибавив к первой строке третью, умноженную на -3 :

$$\left(\begin{array}{ccc|ccc} 1 & 3/2 & 0 & -0,203125 & 0,328125 & 0,0703125 \\ 0 & 1 & 0 & 0,09375 & 0,15625 & -0,109375 \\ 0 & 0 & 1 & 0,234375 & -0,109375 & -0,0234375 \end{array} \right).$$

Осталось избавиться от $a_{12} = 3/2$, прибавив к первой строке вторую, умноженную на $-3/2$:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -0,34375 & 0,09375 & 0,234375 \\ 0 & 1 & 0 & 0,09375 & 0,15625 & -0,109375 \\ 0 & 0 & 1 & 0,234375 & -0,109375 & -0,0234375 \end{array} \right).$$

Процесс преобразования (и обратный ход метода Гаусса тоже) закончен. Та матрица, которая стоит в правой части расширенной матрицы и есть искомая обратная матрица:

$$A^{-1} = \begin{pmatrix} -11/32 & 3/32 & 15/64 \\ 3/32 & 5/32 & -7/64 \\ 15/64 & -7/64 & -3/128 \end{pmatrix} = \begin{pmatrix} -0,34375 & 0,09375 & 0,234375 \\ 0,09375 & 0,15625 & -0,109375 \\ 0,234375 & -0,109375 & -0,0234375 \end{pmatrix}$$

Проверка. Если перемножить полученную матрицу с исходной: $A^{-1} \cdot A$ или $A \cdot A^{-1}$, то результатом этих произведений должна стать единичная матрица E .

Вычисление определителя для матрицы порядка N потребует $2N^3/3 - 1,5N^2 + 5N/6$ арифметических операций. Для матрицы порядка $N \times (N+1)$ прямой ход метода Гаусса для решения СЛАУ потребует $2N^3/3 + N^2/2 - 7N/6$ арифметических операций, причём операций умножения (деления) на N больше, чем операций сложения (вычитания). Обратному ходу при решении СЛАУ потребуется N^2 операций. Для вычисления обратной матрицы потребуется почти $2N^3 - 5N^2/2 + 5N/2$ операций.

Итак, метод Гаусса за конечное число арифметических действий позволяет находить определитель матрицы, решение системы, обратную матрицу. Может случиться, что в процессе вычислений элемент главной диагонали a_{kk} , на который нужно делить его строку k , равен нулю. Тогда следует искать ненулевой элемент под ним (в его столбце, но где-нибудь ниже строкой). Если таковой a_{mk} ($m > k$) найдётся, то меняем местами строки k и m . Если не найдётся, то СЛАУ не имеет единственного решения.

При решении СЛАУ больших порядков ($n > 50$) возрастает вероятность того, что на главной диагонали окажется элемент не равный нулю, но ничтожный по модулю. Поэтому в программы, реализующие метод Гаусса, добавляют блок, который под элементом a_{kk} на главной диагонали находит максимальный по модулю элемент a_{mk} ($m \geq k$): $|a_{mk}| = \max(|a_{ik}|)$, $i=k, k+1, \dots, n$. Затем переставляются строки k и m . Такая перестановка строк называется *выбором ведущего элемента* и позволяет увеличить точность вычислений.

Для решения систем с матрицей общего вида существуют и другие методы решения поставленных задач, например методы факторизации (LU разложения) Краута, Дулитла. Для симметричных матриц – метод Холецкого.

1.3. UL разложение Дулитла

Рассматриваемый в этом параграфе метод факторизации (разложения исходной матрицы в произведение двух треугольных) является модификацией метода Гаусса (схема единственного деления, см. 1.1) и по количеству операций аналогичен прямому ходу метода Гаусса при вычислении детерминанта.

Целью алгоритма является получение матриц U и L , таких, что

$$\begin{pmatrix} u_{11} & 0 & 0 & \dots & 0 \\ u_{21} & u_{22} & 0 & \dots & 0 \\ u_{31} & u_{32} & u_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ u_{n1} & u_{n2} & u_{n3} & \dots & u_{nn} \end{pmatrix} \times \begin{pmatrix} 1 & l_{12} & l_{13} & \dots & l_{1n} \\ 0 & 1 & l_{23} & \dots & l_{2n} \\ 0 & 0 & 1 & \dots & l_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} = A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}$$

Используемые в методе Гаусса преобразования с матрицами можно представить с помощью матричных операций. Так умножение k -ой строки матрицы на число f можно записать как произведение изменённой единичной матрицы, у которой на пересечении k -ой строки и k -ого столбца стоит f вместо 1, например:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & f & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ f \cdot a_{21} & f \cdot a_{22} & f \cdot a_{23} & f \cdot a_{24} & f \cdot a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{pmatrix}$$

Чтобы поменять местами строки k и m нужно умножить матрицу на единичную, в которой изменены четыре элемента: $e_{kk}=0$, $e_{mm}=0$, $e_{km}=1$, $e_{mk}=1$. Например, такое произведение меняет первую и четвёртую строки:

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{pmatrix} = \begin{pmatrix} a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{pmatrix}$$

Чтобы к строке k прибавить строку m , умноженную на g , нужно умножить матрицу на единичную, в которой изменен элемент: $e_{km}=g$.

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & g & 0 & 1 \end{pmatrix} \times \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ ga_{21} + a_{41} & ga_{22} + a_{42} & ga_{23} + a_{43} & ga_{24} + a_{44} \end{pmatrix}$$

В качестве примера рассмотрим матрицу (1.1.4) из параграфа 1.1:

$$\begin{pmatrix} 2 & 5 & 1 \\ -1 & 2 & -2 \\ 6 & 2 & 1 \end{pmatrix} = (\text{первая строка делится на 2}) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 5/2 & 1/2 \\ -1 & 2 & -2 \\ 6 & 2 & 1 \end{pmatrix} =$$

ко второй строке прибавляем первую строку умноженную на -1

$$= \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 5/2 & 1/2 \\ 0 & 9/2 & -3/2 \\ 6 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 5/2 & 1/2 \\ 0 & 9/2 & -3/2 \\ 6 & 2 & 1 \end{pmatrix} =$$

К третьей строке прибавим первую, умноженную на -6

$$= \begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 6 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 5/2 & 1/2 \\ 0 & 9/2 & -3/2 \\ 0 & -13 & -2 \end{pmatrix} \text{ разделим вторую строку на } 9/2$$

$$= \begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 6 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 5/2 & 1/2 \\ 0 & 9/2 & -3/2 \\ 0 & -13 & -2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 6 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 9/2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 5/2 & 1/2 \\ 0 & 1 & -1/3 \\ 0 & -13 & -2 \end{pmatrix} =$$

$$\begin{pmatrix} 2 & 0 & 0 \\ -1 & 9/2 & 0 \\ 6 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 5/2 & 1/2 \\ 0 & 1 & -1/3 \\ 0 & -13 & -2 \end{pmatrix} = \text{К третьей строке прибавим вторую,}$$

умноженную на $-(-13)$:

$$\begin{pmatrix} 2 & 0 & 0 \\ -1 & 9/2 & 0 \\ 6 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -13 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 5/2 & 1/2 \\ 0 & 1 & -1/3 \\ 0 & 0 & -19/3 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 9/2 & 0 \\ 6 & -13 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 5/2 & 1/2 \\ 0 & 1 & -1/3 \\ 0 & 0 & -19/3 \end{pmatrix} =$$

разделим третью строку на $-19/3$:

$$\begin{pmatrix} 2 & 0 & 0 \\ -1 & 9/2 & 0 \\ 6 & -13 & 1 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -19/3 \end{pmatrix} \times \begin{pmatrix} 1 & 5/2 & 1/2 \\ 0 & 1 & -1/3 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 9/2 & 0 \\ 6 & -13 & -19/3 \end{pmatrix} \times \begin{pmatrix} 1 & 5/2 & 1/2 \\ 0 & 1 & -1/3 \\ 0 & 0 & 1 \end{pmatrix} =$$

$U \times L$.

Теперь слау $A \bar{x} = \bar{b}$ будет иметь вид $U \cdot (L \cdot \bar{x}) = \bar{b}$ решение системы

$$\begin{cases} 2x_1 + 5x_2 + x_3 = 1 \\ -x_1 + 2x_2 - 2x_3 = 2 \\ 6x_1 + 2x_2 + x_3 = 3 \end{cases}$$

Сведётся к последовательному решению двух систем: $U \cdot \bar{y} = \bar{b}$ и $L \cdot \bar{x} = \bar{y}$:

$$\begin{cases} 2y_1 = 1 \\ -y_1 + 9y_2/2 = 2 \\ 6y_1 - 13y_2 - 19y_3/3 = 3 \end{cases} \text{ и } \begin{cases} x_1 + 5x_2/2 + x_3/2 = y_1 \\ x_2 - x_3/3 = y_2 \\ x_3 = y_3 \end{cases}$$

Из первой системы: $y_1=1/2$; $y_2=5/9$; $y_3=-65/57$. Из второй: $x_3=-65/57$; $x_2=10/57$; $x_1=12/19$.

Решение слау с помощью факторизации матрицы системы незначительно (приблизительно на N^2 операций больше) метода Гаусса. Существенное преимущество факторизации скажется при многократном решении слау с одной матрицей системы и разными правыми частями.

В случае необходимости менять местами строки в разложение добавится матрица перестановок P : $A=U \cdot L \cdot P$.

1.4. LU разложение Краута .

Рассматриваемый в этом параграфе метод факторизации является незначительно (приблизительно на N^2 операций) быстрее UL разложения Дулитла. Целью алгоритма является получение матриц L и U , таких, что

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ l_{21} & 1 & 0 & \dots & 0 \\ l_{31} & l_{32} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ l_{n1} & l_{n2} & l_{n3} & \dots & 1 \end{pmatrix} \times \begin{pmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & u_{22} & u_{23} & \dots & u_{2n} \\ 0 & 0 & u_{33} & \dots & u_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & u_{nn} \end{pmatrix} = A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}$$

Основное отличие этой модификации метода Гаусса в том, что мы не будем делить строки матрицы, пытаясь получить единичную главную диагональ.

В качестве примера рассмотрим ту же матрицу:

$$\begin{pmatrix} 2 & 5 & 1 \\ -1 & 2 & -2 \\ 6 & 2 & 1 \end{pmatrix} =$$

из второй строки вычитаем первую строку, умноженную на $-1/2$:

$$= \begin{pmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 2 & 5 & 1 \\ 0 & 9/2 & -3/2 \\ 6 & 2 & 1 \end{pmatrix} =$$

из третьей строки вычитаем первую строку, умноженную на $6/2$ (на 3)

$$= \begin{pmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 2 & 5 & 1 \\ 0 & 9/2 & -3/2 \\ 0 & -13 & -2 \end{pmatrix} =$$

из третьей строки вычитаем вторую строку, умноженную на $(-26/9)$:

$$= \begin{pmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 3 & -26/9 & 1 \end{pmatrix} \times \begin{pmatrix} 2 & 5 & 1 \\ 0 & 9/2 & -3/2 \\ 0 & 0 & -19/3 \end{pmatrix} = L \times U.$$

Теперь слау $A \bar{x} = \bar{b}$ будет иметь вид $L \cdot (U \cdot \bar{x}) = \bar{b}$ решение системы

$$\begin{cases} 2x_1 + 5x_2 + x_3 = 1 \\ -x_1 + 2x_2 - 2x_3 = 2 \\ 6x_1 + 2x_2 + x_3 = 3 \end{cases}$$

Сведётся к последовательному решению двух систем: $L \cdot \bar{z} = \bar{b}$ и $U \cdot \bar{x} = \bar{z}$:

$$\begin{cases} z_1 = 1 \\ -z_1/2 + z_2 = 2 \\ 3z_1 - 26z_2/9 + z_3 = 3 \end{cases} \text{ и } \begin{cases} 2x_1 + 5x_2 + x_3 = z_1 \\ 9x_2/2 - 3x_3/2 = z_2 \\ -19x_3/3 = z_3 \end{cases}$$

Решение первой системы $z_1=1$; $z_2=5/2$; $z_3=65/9$.

Решение второй системы $x_3=-65/57$; $x_2=10/57$; $x_1=12/19$.

1.5. Метод Холецкого (квадратного корня) для симметричных матриц.

Симметричные матрицы иногда можно представить в виде произведения двух треугольных: $A=U \cdot U^*$:

$$A = \begin{pmatrix} a_{11} & a_{21} & a_{31} & \dots & a_{n1} \\ a_{21} & a_{22} & a_{32} & \dots & a_{n2} \\ a_{31} & a_{32} & a_{33} & \dots & a_{n3} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} u_{11} & 0 & 0 & \dots & 0 \\ u_{21} & u_{22} & 0 & \dots & 0 \\ u_{31} & u_{32} & u_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ u_{n1} & u_{n2} & u_{n3} & \dots & u_{nn} \end{pmatrix} \times \begin{pmatrix} u_{11} & u_{21} & u_{31} & \dots & u_{n1} \\ 0 & u_{22} & u_{32} & \dots & u_{n2} \\ 0 & 0 & u_{33} & \dots & u_{n3} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & u_{nn} \end{pmatrix}$$

Для этого симметричная матрица должна иметь одно из равносильных свойств:

- положительно определена;
- все главные миноры должны быть положительными;
- иметь диагональное преобладание и на главной диагонали иметь положительные значения;
- все её собственные числа должны быть положительны.

Алгоритм вычисления элементов матрицы U следующий:

1. $u_{11} = \sqrt{a_{11}}$,
2. $u_{j1} = a_{j1}/u_{11}$ для $j=2,3,\dots,N$;
3. для $j=2,3,\dots,N$: $u_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} u_{jk}^2}$;
4. для $j=2,3,\dots,N-1$ и $i=j+1,\dots,N$: $u_{ij} = \left(a_{ij} - \sum_{k=1}^{j-1} u_{jk} \cdot u_{ik} \right) / u_{jj}$;

Пример. Решить СЛАУ методом квадратного корня Холецкого.

$$\begin{cases} 9x_1 + 3x_2 - 3x_3 = -3 \\ 3x_1 + 10x_2 - 5x_3 = -40 \\ -3x_1 - 4x_2 + 11x_3 = 50 \end{cases} \quad A = \begin{pmatrix} 9 & 3 & -3 \\ 3 & 10 & -4 \\ -3 & -4 & 11 \end{pmatrix}$$

Вычислим элементы матрицы U :

$$U_{11} = \sqrt{a_{11}} = \sqrt{9} = 3, \quad U_{21} = \frac{a_{21}}{U_{11}} = 3/3 = 1; \quad U_{31} = \frac{a_{31}}{U_{11}} = -3/3 = -1;$$

$$U_{22} = \sqrt{a_{22} - U_{21}^2} = \sqrt{10 - 1} = 3, \quad U_{32} = \frac{a_{32} - U_{31} \cdot U_{21}}{U_{22}} = (-4 - (-1) \cdot 1) / 3 = -1;$$

$$U_{33} = \sqrt{a_{33} - U_{31}^2 - U_{32}^2} = \sqrt{11 - (-1)^2 - (-1)^2} = 3.$$

Разложение будет выглядеть следующим образом.

$$A = U \cdot U^T = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ -1 & -1 & 3 \end{pmatrix} \times \begin{pmatrix} 3 & 1 & -1 \\ 0 & 3 & -1 \\ 0 & 0 & 3 \end{pmatrix}$$

Вместо слау $A \cdot \bar{x} = \bar{b}$ будем решать $U \cdot (U^T \cdot \bar{x}) = \bar{b}$, сделав замену $\bar{y} = U^T \cdot \bar{x}$. А именно, сначала решим систему $U \cdot \bar{y} = \bar{b}$:

$$\begin{cases} 3y_1 = -3 \\ y_1 + 3y_2 = -40 \\ -y_1 - y_2 + 3y_3 = 50 \end{cases} \quad \begin{cases} \text{Получим: } y_1 = -1, \quad y_2 = -13, \quad y_3 = 12. \\ \text{Затем решаем систему } U^T \cdot \bar{x} = \bar{y}: \\ \begin{cases} 3x_1 + x_2 - x_3 = -1 \\ 3x_2 - x_3 = -13 \\ 3x_3 = 12 \end{cases} \end{cases}$$

Метод квадратных корней Холецкого быстрее метода Гаусса (по количеству арифметических операций) почти в два раза. Но применим он не для любых симметричных матриц, а для положительно определённых.

1.6. Решение трёхдиагональной СЛАУ методом прогонки

СЛАУ с трёхдиагональной матрицей встречаются, например, при построении кубических интерполяционных сплайнов, при численном решении дифференциальных уравнений.

Пусть дана СЛАУ специального вида

$$a_i x_{i-1} + b_i x_i + c_i x_{i+1} = d_i, \quad i=1,\dots,n. \quad (1.6.1)$$

Запишем расширенную матрицу для системы (1.3-1):

$$\begin{pmatrix} b_1 & c_1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & d_1 \\ a_2 & b_2 & c_2 & 0 & 0 & \dots & 0 & 0 & 0 & d_2 \\ 0 & a_3 & b_3 & c_3 & 0 & \dots & 0 & 0 & 0 & d_3 \\ 0 & 0 & a_4 & b_4 & c_4 & \dots & 0 & 0 & 0 & d_4 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & b_{n-2} & c_{n-2} & 0 & d_{n-2} \\ 0 & 0 & 0 & 0 & 0 & \dots & a_{n-1} & b_{n-1} & c_{n-1} & d_{n-1} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & a_n & b_n & d_n \end{pmatrix} \quad (1.6.2)$$

Заметим, что в первом уравнении нет коэффициента a_1 , а в последнем уравнении отсутствует c_n , т.е. $a_1 = 0, c_n = 0$.

При проведении прямого хода метода прогонки вычисляются прогоночные коэффициенты P_i и Q_i .

$P_0 = 0$, $Q_0 = 0$, остальные коэффициенты P_i и Q_i ($i = 1, 2, 3, \dots, n$) вычисляем по формулам:

$$P_i = \frac{-c_i}{b_i + a_i \cdot P_{i-1}}, \quad Q_i = \frac{d_i - a_i \cdot Q_{i-1}}{b_i + a_i \cdot P_{i-1}} \quad \text{для } i = 1, 2, \dots, n. \quad (1.6.3)$$

После вычисления всех прогоночных коэффициентов можно в обратном ходе прогонки вычислить все неизвестные: $x_n, x_{n-1}, x_{n-2}, \dots, x_1$:

$$\begin{aligned} x_n &= Q_n, \\ x_i &= Q_i + P_i \cdot x_{i+1}, \quad i = n-1, n-2, \dots, 2, 1. \end{aligned} \quad (1.6.4)$$

Пример. Решить СЛАУ порядка 5 с трёхдиагональной расширенной матрицей :

$$\left(\begin{array}{ccccc|c} -34 & -26 & 0 & 0 & 0 & 34 \\ 64 & -124 & -56 & 0 & 0 & 38 \\ 0 & 94 & -274 & -86 & 0 & 42 \\ 0 & 0 & 124 & -484 & -116 & 46 \\ 0 & 0 & 0 & 154 & -754 & 50 \end{array} \right)$$

Первая строка даёт $a_1 = 0$, $b_1 = -34$, $c_1 = -26$, $d_1 = 34$. Из второй строки $a_2 = 64$, $b_2 = -124$, $c_2 = -56$, $d_2 = 38$. Из последней строки $a_5 = 154$, $b_5 = -754$, $c_5 = 0$, $d_5 = 50$. Проведём прямой ход прогонки. $P_0 = 0$, $Q_0 = 0$;

$$P_1 = \frac{-c_1}{b_1 + a_1 P_0} = \frac{-c_1}{b_1} = \frac{-26}{-34} = -0,764705882 ;$$

$$Q_1 = \frac{d_1 - a_1 \cdot Q_0}{b_1 + a_1 P_0} = \frac{d_1}{b_1} = \frac{-34}{-34} = -1 ;$$

$$P_2 = \frac{-c_2}{b_2 + a_2 P_1} = \frac{56}{-124 + 64 \cdot (-0,764705882)} = -0,3238095 ;$$

$$Q_2 = \frac{d_2 - a_2 \cdot Q_1}{b_2 + a_2 P_1} = \frac{64 \cdot (-1) - 38}{-124 - 64 \cdot (-0,764705882)} = -0,58979559 ;$$

$$P_3 = \frac{-c_3}{b_3 + a_3 P_2} = \frac{86}{-274 + 94 \cdot (-0,3238095)} = -0,2824876 ;$$

$$Q_3 = \frac{d_3 - a_3 \cdot Q_2}{b_3 + a_3 P_2} = \frac{42 - 94 \cdot (-0,58979559)}{-274 + 94 \cdot (-0,3238095)} = -0,3200677 ;$$

$$P_4 = \frac{-c_4}{b_4 + a_4 P_3} = \frac{-116}{484 - 124 \cdot (-0,2824876)} = -0,2234945 ;$$

$$Q_4 = \frac{d_4 - a_4 \cdot Q_3}{b_4 + a_4 P_3} = \frac{124 \cdot (-0,3200677) - 46}{484 - 124 \cdot (-0,2824876)} = -0,1650938 ;$$

$$P_5 = \frac{-c_5}{b_5 + a_5 P_4} = \frac{0}{754 - 154 \cdot (-0,2234945)} = 0 ;$$

$$Q_5 = \frac{d_5 - a_5 \cdot Q_4}{b_5 + a_5 P_4} = \frac{154 \cdot (-0,1650938) - 50}{754 - 154 \cdot (-0,2234945)} = -0,0956655 .$$

Закончился прямой ход прогонки, в обратном ходе найдём все x :

$$x_5 = Q_5 = -0,0956655 ;$$

$$x_4 = Q_4 + P_4 \cdot x_5 = -0,1650938 + (-0,2234945) \cdot (-0,0956655) = -0,1437131 ;$$

$$x_3 = Q_3 + P_3 \cdot x_4 = -0,3200677 + (-0,2824876) \cdot (-0,1437131) = -0,2794706 ;$$

$$x_2 = Q_2 + P_2 \cdot x_3 = -0,5897959 + (-0,3238095) \cdot (-0,2794706) = -0,4993007 ;$$

$$x_1 = Q_1 + P_1 \cdot x_2 = -1 + (-0,764705882) \cdot (-0,4993007) = -0,6181818 .$$

Запишем ответ в привычном виде:

$$\begin{cases} x_1 = -0,6181818 ; \\ x_2 = -0,4993007 ; \\ x_3 = -0,2794706 ; \\ x_4 = -0,1437131 ; \\ x_5 = -0,0956655 . \end{cases}$$

Проверка. Если подставить найденные неизвестные в исходную систему, то в каждой строке должно быть тождество (точное равенство).

Для матриц больших порядков (более 500) возникает вопрос об устойчивости метода. При неустойчивом процессе ошибки вычислений могут быть настолько велики, что вычисленное решение не будет иметь ничего общего с тем, которое существует на самом деле. Исследование устойчивости не входит в цели, которые ставил перед собой автор при написании этой работы. Но можно утверждать, что если у матрицы есть диагональное преобладание (т.е. у матрицы модули элементов главной диагонали превосходят сумму абсолютных величин остальных элементов в своей строке и своём столбце), то этого достаточно для устойчивости метода.

При программировании метода прогонки достаточно использовать два последовательных цикла отдельно для прямого и обратного хода. Необходимо использование дополнительно двух массивов для хранения P и Q , но общий объём памяти всё равно меньше, чем при программировании решения для матрицы общего вида.

Кроме трёхдиагональных, существуют и другие типы разреженных матриц с большим количеством нулей. Для решения систем с такими матрицами существуют свои специфические методы.

2. Итерационные методы решения СЛАУ

Поскольку из-за погрешностей при выполнении вычислений ошибка решения может возрасти до неприемлемых величин, иногда возникает необходимость уточнить решение и приблизить его к точному решению. Эту задачу могут решить итерационные методы.

Определение. Нормой $\|\cdot\|$ на множестве X называют любую функцию, удовлетворяющую трём соотношениям:

- 1) $\|x\| \geq 0$ для любого $x \in X$. И $\|x\| = 0$ тогда и только тогда, когда $x = 0$,
- 2) $\|a \cdot x\| = |a| \cdot \|x\|$ (a -число и $x \in X$),
- 3) $\|x+y\| \leq \|x\| + \|y\|$ для любых $x, y \in X$.

Говорят, что нормы матрицы B и вектора \bar{a} согласованы, если $\|B \cdot \bar{a}\| \leq \|B\| \cdot \|\bar{a}\|$.

В качестве примера из бесконечного числа существующих норм рассмотрим три нормы векторов размерности n и согласованные с ними нормы матриц размерности $n \times k$:

$$\|\bar{x}\|_1 = \max_{1 \leq i \leq n} |x_i| \quad (= \text{максимальная по абсолютной величине компонента вектора}),$$

$$\|\bar{x}\|_2 = \sum_{i=1}^n |x_i| \quad (= \text{сумма абсолютных величин компонент вектора}),$$

$$\|\bar{x}\|_3 = \sqrt{\sum_{i=1}^n |x_i|^2} \quad (= \text{корень из суммы квадратов абсолютных величин компонент вектора}), \text{ она же евклидова норма,}$$

$$\|A\|_1 = \max_{1 \leq i \leq n} \sum_{j=1}^k |a_{ij}| \quad (\text{в каждой строке вычисляется сумма абсолютных величин компонент строки и из этих сумм выбирают максимальную}),$$

$$\|A\|_2 = \max_{1 \leq j \leq k} \sum_{i=1}^n |a_{ij}| \quad (\text{в каждом столбце вычисляется сумма абсолютных величин компонент столбца и из этих сумм выбирают максимальную}),$$

$$\|A\|_3 = \sqrt{\sum_{i=1}^n \sum_{j=1}^k |a_{ij}|^2} \quad (= \text{корень из суммы квадратов абсолютных величин компонент матрицы}), \text{ она же евклидова норма,}$$

$\|A\|_4 = n \cdot \max |a_{ij}|$ (= размер квадратной матрицы $n \times n$ умножается на абсолютную величину максимального по модулю элемента матрицы).

В приведённых примерах нормы матриц и векторов с одинаковыми номерами согласованы. Норма $\|A\|_4$ согласована со всеми приведёнными нормами векторов.

Пример. Найти значения приведённых выше норм для матрицы A и вектора \bar{b} :

$$\bar{b} = \begin{pmatrix} 10 \\ -11 \\ 12 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & -2 & 3 \\ -4 & 5 & -6 \\ 7 & -8 & 9 \end{pmatrix}$$

$$\|\bar{b}\|_1 = \max \{|10|, |-11|, |12|\} = 12 \quad \|\bar{b}\|_2 = |10| + |-11| + |12| = 33,$$

$$\|\bar{b}\|_3 = \sqrt{|10|^2 + |-11|^2 + |12|^2} = \sqrt{365};$$

$$\|A\|_1 = \max \{(|1| + |-2| + |3|), (|-4| + |5| + |-6|), (|7| + |-8| + |9|)\} = \max \{6, 15, 24\} = 24;$$

$$\|A\|_2 = \max \{(|1| + |-4| + |7|), (|-2| + |5| + |-8|), (|3| + |-6| + |9|)\} = \max \{12, 15, 18\} = 18;$$

$$\|A\|_3 = \sqrt{|1|^2 + |-2|^2 + |3|^2 + |-4|^2 + |5|^2 + |-6|^2 + |7|^2 + |-8|^2 + |9|^2} = \sqrt{285}.$$

2.1 Метод простых итераций решения СЛАУ

Для решения системы $A\bar{x} = \bar{b}$ каким-либо образом преобразуем эту систему к виду (схеме) $\bar{x} = B\bar{x} + \bar{\beta}$. По этой схеме можно построить итерационный процесс:

$$\bar{x}^{(n+1)} = B\bar{x}^{(n)} + \bar{\beta}. \quad (2.1.1)$$

В левой части стоит новый вектор неизвестных $\bar{x}^{(n+1)}$, а в правой части – старый вектор неизвестных $\bar{x}^{(n)}$. После вычисления нового вектора он превращается в старый и вычисляем следующий новый вектор. За начальный вектор $\bar{x}^{(1)}$ можно взять вектор $\bar{\beta}$. Если хотя бы какая-нибудь норма матрицы B окажется меньше 1, то последовательность векторов $\bar{x}^{(n)}$ из (2.1.1) будет сходиться к точному решению $\bar{\mu}$. Сходимость будет тем быстрее, чем меньше норма у матрицы B . В процессе итераций точность (близость вектора $\bar{x}^{(n)}$ к $\bar{\mu}$) можно контролировать с помощью несложного соотношения

$$\|\bar{x}^{(n)} - \bar{\mu}\| < \frac{\|B\|^n}{1 - \|B\|} \cdot \|\bar{\beta}\| \quad (2.1.2)$$

В этом соотношении нормы $\|B\|$ и $\|\bar{\beta}\|$ должны быть согласованы (см. разд. 2). Поэтому желательно выбрать норму матрицы B с наименьшим значением и уже для неё искать ей согласованную норму вектора $\bar{\beta}$. Итак, на последних шагах итерационного процесса вектора $\bar{x}^{(n)}$ и $\bar{x}^{(n+1)}$ будут мало отличаться друг от друга, т.е. их компоненты попарно будут почти одинаковыми. Вычисления прекращают, если правая часть выражения (2.1.2) стала меньше некоторого заранее заданного числа ε ($\varepsilon > 0$). Оценка (2.1.2) значительно завышает число итераций, поэтому иногда для оценки точности используют другое соотношение:

$$\|\bar{x}^{(n)} - \bar{\mu}\| \leq \frac{\|B\|}{1 - \|B\|} \cdot \|\bar{x}^{(n)} - \bar{x}^{(n-1)}\|$$

Пример. Решить

$$\begin{cases} 20x + 4y - 8z = 1 \\ -3x + 15y + 5z = -2 \\ 6x + 3y - 18z = 3 \end{cases} \quad (2.1.3)$$

Воспользуемся тем обстоятельством, что матрица системы обладает свойством диагонального преобладания (модуль каждого элемента, стоящего на главной диагонали, больше сумм модулей остальных элементов в своей строке, столбце). Оставим диагональные элементы слева от знака равно, остальные перенесём вправо:

$$\begin{cases} 20x = 0x - 4y + 8z + 1 \\ 15y = 3x + 0y - 5z - 2 \\ -18z = -6x - 3y + 0z + 3 \end{cases}$$

Разделим строки на коэффициенты неизвестных, стоящих слева от знака равно: первую строку – на 20, вторую – на 15, третью – на -18.

$$\begin{cases} x = 0x - \frac{1}{5}y + \frac{2}{5}z + \frac{1}{20} \\ y = \frac{1}{5}x + 0y - \frac{1}{3}z - \frac{2}{15} \\ z = \frac{1}{3}x + \frac{1}{6}y + 0z - \frac{1}{6} \end{cases} \quad (2.1.4)$$

Или в матричном виде $\bar{x} = B\bar{x} + \bar{\beta}$:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{5} & \frac{2}{5} \\ \frac{1}{5} & 0 & -\frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} \frac{1}{20} \\ -\frac{2}{15} \\ -\frac{1}{6} \end{pmatrix}$$

Вычислим нормы матрицы B и вектора $\bar{\beta}$:

$$\|B\|_1 = \max \{ (0 + | -1/5 | + | 2/5 |), (| 1/5 | + 0 + | -1/3 |), (| 1/3 | + | 1/6 | + 0) \} = \max \{ 3/5, 8/15, 1/2 \} = 3/5 = 0,6$$

$$\|B\|_2 = \max \{ (0 + | 1/5 | + | 1/3 |), (| -1/5 | + 0 + | 1/6 |), (| 2/5 | + | -1/3 | + 0) \} = \max \{ 8/15, 11/30, 11/15 \} = 11/15 \approx 0,73333333$$

Минимальная из них $\|B\|_1$, следовательно, вычисляем $\|\bar{\beta}\|_1 = \max (1/20, 2/15, 1/6) = 1/6 \approx 0,16666666$. Так как $\|B\| < 1$, то итерационный процесс

$$\begin{cases} x^{(n+1)} = 0x^{(n)} - \frac{1}{5}y^{(n)} + \frac{2}{5}z^{(n)} + \frac{1}{20} \\ y^{(n+1)} = \frac{1}{5}x^{(n)} + 0y^{(n)} - \frac{1}{3}z^{(n)} - \frac{2}{15} \\ z^{(n+1)} = \frac{1}{3}x^{(n)} + \frac{1}{6}y^{(n)} + 0z^{(n)} - \frac{1}{6} \end{cases} \quad (2.1.5)$$

будет сходиться к точному решению системы (2.1.3).

За начальный вектор $\bar{x}^{(0)}$ возьмём вектор $\bar{\beta}$, тогда

$$\begin{aligned} x^{(0)} &= 1/20 = 0,05; \\ y^{(0)} &= -2/15 \approx -0,13333333; \\ z^{(0)} &= -1/6 \approx -0,16666666, \end{aligned}$$

а новые $x^{(1)}, y^{(1)}, z^{(1)}$ (уточнённые) вычислим, подставив числовые значения $x^{(0)}, y^{(0)}, z^{(0)}$ в систему (2.1.5):

$$\begin{cases} x^{(1)} = 0x^{(0)} - \frac{1}{5}y^{(0)} + \frac{2}{5}z^{(0)} + \frac{1}{20} = -\frac{1}{5} \cdot \left(-\frac{2}{15}\right) + \frac{2}{5} \cdot \left(-\frac{1}{6}\right) + \frac{1}{20} = \frac{1}{100} = 0,01 \\ y^{(1)} = \frac{1}{5}x^{(0)} + 0y^{(0)} - \frac{1}{3}z^{(0)} - \frac{2}{15} = \frac{1}{5} \cdot \frac{1}{20} + 0 - \frac{1}{3} \cdot \left(-\frac{1}{6}\right) - \frac{2}{15} = -\frac{61}{900} \approx -0,067 \\ z^{(1)} = \frac{1}{3}x^{(0)} + \frac{1}{6}y^{(0)} + 0z^{(0)} - \frac{1}{6} = \frac{1}{3} \cdot \frac{1}{20} + \frac{1}{6} \cdot \left(-\frac{2}{15}\right) - \frac{1}{6} = -\frac{155}{900} \approx -0,172 \end{cases}$$

Оценка точности (погрешности) найденных $x^{(1)}, y^{(1)}, z^{(1)}$ вычисляется по формуле (2.1.2)

$$\varepsilon_1 = \frac{\|B\|}{1-\|B\|} \|\bar{\beta}\| = \frac{0,6}{1-0,6} \cdot \frac{1}{6} = 0,25.$$

Кроме этого, на каждом шаге будем вычислять первую норму от разности векторов- неизвестных старого и нового:

$$\delta_1 = \|\bar{x}^{(1)} - \bar{x}^{(0)}\|_1 = \max\{|0,05 - 0,01|; |-2/15 + 61/900|; |1/6 - 155/900|\} \approx 0,06555555;$$

Сделаем ещё один шаг итерации, вычислим $x^{(2)}$, $y^{(2)}$, $z^{(2)}$:

$$x^{(2)} = 0 \cdot x^{(1)} - \frac{1}{5} \cdot y^{(1)} + \frac{2}{5} \cdot z^{(1)} + \frac{1}{20} = 0 - \frac{1}{5} \cdot (-0,06777) + \frac{2}{5} \cdot (-0,17222) + \frac{1}{20} \approx -0,0053333$$

$$y^{(2)} = \frac{1}{5} \cdot x^{(1)} + 0 \cdot y^{(1)} - \frac{1}{3} \cdot z^{(1)} - \frac{2}{15} = \frac{1}{5} \cdot 0,01 + 0 - \frac{1}{3} \cdot (-0,17222) - \frac{2}{15} \approx -0,0739259$$

$$z^{(2)} = \frac{1}{3} \cdot x^{(1)} + \frac{1}{6} \cdot y^{(1)} + 0 \cdot z^{(1)} - \frac{1}{6} = \frac{1}{3} \cdot 0,01 + \frac{1}{6} \cdot (-0,06777) + 0 - \frac{1}{6} \approx -0,1746296.$$

Оценка точности ε_2 найденных $x^{(2)}$, $y^{(2)}$, $z^{(2)}$ и δ_2 вычисляются аналогично:

$$\varepsilon_2 = \frac{\|B\|^2}{1-\|B\|} \|\bar{\beta}\| = \frac{0,6^2}{1-0,6} \cdot \frac{1}{6} = 0,15,$$

$$\delta_2 = \|\bar{x}^{(2)} - \bar{x}^{(1)}\|_1 \approx 0,01533333.$$

Поведём ещё несколько итераций и их результаты запишем в таблицу:

№	x	y	z	ε	δ
0	.0500000	-.1333333	-.1666667		
1	.0100000	-.0677778	-.1722222	.25000000	.06555556
2	-.0053333	-.0739259	-.1746296	.15000000	.01533333
3	-.0050667	-.0761901	-.1807654	.09000000	.00613580
4	-.0070681	-.0740915	-.1810539	.05400000	.00209860
5	-.0076033	-.0743957	-.1813713	.03240000	.00053511
6	-.0076694	-.0743969	-.1816004	.01944000	.00022906
7	-.0077608	-.0743338	-.1816226	.01166400	.00009138

Проверка. В простейшем случае (для хорошо обусловленных систем, которые мы и рассматриваем) если подставить найденные неизвестные в исходную систему, то в каждой строке должно быть почти точное равенство (с точностью ε).

2.2. Метод Зейделя решения СЛАУ

Вернёмся к решению примера (2.1-3). Исследование сходимости итерационного процесса (2.1-5) уже проведено, найдены нормы матрицы B и вектора $\bar{\beta}$. За начальный вектор $\bar{x}^{(0)}$ возьмём вектор $\bar{\beta}$, т.е.

$$x^{(0)} = 1/20 = 0,05; \quad y^{(0)} = -2/15 \approx -0,133333333; \quad z^{(0)} = -1/6 \approx -0,166666666.$$

Тогда новый $x^{(1)}$ вычислим, подставив числовые значения $x^{(0)}$, $y^{(0)}$, $z^{(0)}$ в первое уравнение (2.1-4):

$$x^{(1)} = 0x^{(0)} - \frac{1}{5}y^{(0)} + \frac{2}{5}z^{(0)} + \frac{1}{20} = -\frac{1}{5} \cdot \left(-\frac{2}{15}\right) + \frac{2}{5} \cdot \left(-\frac{1}{6}\right) + \frac{1}{20} = 0,01.$$

Теперь в правую часть второго уравнения (2.1-4) вместо x , y , z подставим только что вычисленное $x^{(1)}$ и старые $y^{(0)}$, $z^{(0)}$:

$$y^{(1)} = \frac{1}{5}x^{(1)} + 0y^{(0)} - \frac{1}{3}z^{(0)} - \frac{2}{15} = \frac{1}{5} \cdot 0,01 + \frac{1}{3} \cdot \frac{1}{6} - \frac{2}{15} = \frac{341}{4500} \approx -0,075777777.$$

В правую часть третьего уравнения (2.1-4) вместо x , y , z подставим только что вычисленные $x^{(1)}$, $y^{(1)}$ и старое $z^{(0)}$:

$$z^{(1)} = \frac{1}{3}x^{(1)} + \frac{1}{6}y^{(1)} + 0z^{(0)} - \frac{1}{6} = \frac{1}{3} \cdot 0,01 + \frac{1}{6} \cdot (-0,07577777) - \frac{1}{6} \approx -0,175963.$$

При вычислении $x^{(2)}$ используем значения $x^{(1)}$, $y^{(1)}$, $z^{(1)}$. При вычислении $y^{(2)}$ используем значения только что вычисленного $x^{(2)}$ и старых $y^{(1)}$, $z^{(1)}$. При вычислении $z^{(2)}$ используем значения только что вычисленных $x^{(2)}$ и $y^{(2)}$ и старого $z^{(1)}$.

Этот итерационный процесс называют *методом Зейделя* и записывают:

$$\begin{cases} x^{(n+1)} = 0x^{(n)} - \frac{1}{5}y^{(n)} + \frac{2}{5}z^{(n)} + \frac{1}{20} \\ y^{(n+1)} = \frac{1}{5}x^{(n+1)} + 0y^{(n)} - \frac{1}{3}z^{(n)} - \frac{2}{15} \\ z^{(n+1)} = \frac{1}{3}x^{(n+1)} + \frac{1}{6}y^{(n+1)} + 0z^{(n)} - \frac{1}{6} \end{cases} \quad (2.1.6)$$

Такой способ вычислений позволяет иногда серьёзно увеличить скорость сходимости к точному решению. По сравнению с методом простых итераций количество итераций в методе Зейделя может быть меньше в разы.

Поведём ещё несколько итераций по методу Зейделя и их результаты запишем в таблицу:

№	x	y	z	ε	δ
0	.0500000	-.1333333	-.1666667		
1	.0100000	-.0757778	-.1759630	.25	.058
2	-.0052296	-.0757249	-.1810307	.15	.015

3	-.0072673	-.0744432	-.1814963	.09	.002
4	-.0077099	-.0743765	-.1816327	.054	.0004
5	-.0077778	-.0743447	-.1816500	.032	.000068

Проверка. Как и в предыдущем примере, если подставить найденные неизвестные в исходную систему, то в каждой строке должно быть почти точное равенство (с точностью не хуже ε).

Замечание. Критерием сходимости рассмотренных итерационных процессов (метод простой итерации и метод Зейделя) является $\|B\|_5 = \max |\lambda_i| < 1$. Здесь λ_i – собственные числа матрицы B . Поиску собственных чисел посвящена следующая глава.

3. Решение второй задачи линейной алгебры

Пусть дана квадратная матрица A размерности $n \times n$. Если найдутся такое число λ и такой вектор \bar{v} , для которых будет справедливо соотношение

$$A \times \bar{v} = \lambda \times \bar{v}, \quad (3.1)$$

то такие λ и \bar{v} называются *собственным* (*характеристическим*) числом и (соответствующим ему) *собственным вектором* матрицы A .

Собственные числа у матрицы могут быть одинаковыми (кратными) или различными, большими ($|\lambda_i| \gg 1$) или маленькими ($|\lambda_i| \ll 1$). Наибольшее по модулю собственное число называют *спектральным радиусом*. У кратного собственного числа (с алгебраической кратностью k) может быть целое пространство (размерность пространства – это геометрическая кратность m) соответствующих ему собственных векторов. Но $k \geq m$.

Нахождение (вычисление) для заданной матрицы её собственных чисел и векторов называют второй основной задачей линейной алгебры. Нахождение всех собственных чисел и векторов матрицы называют *полной* задачей (проблемой), а вычисление лишь некоторых из них – *частичной* проблемой.

3.1. Степенной метод вычисления спектрального радиуса

Пусть дана матрица A у которой собственные числа (спектр) удовлетворяют соотношению $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$. Пусть λ_1 – действительное (не комплексное) число.

Построим последовательность векторов \bar{v} , $\bar{\omega}$ и чисел ρ :

$$\bar{v}^{(n+1)} = A \cdot \bar{\omega}^{(n)}, \quad \rho_{n+1} = (\bar{v}^{(n+1)}, \bar{\omega}^{(n)}) \quad , \quad \bar{\omega}^{(n+1)} = \frac{\bar{v}^{(n+1)}}{\sqrt{(\bar{v}^{(n+1)}, \bar{v}^{(n+1)})}}. \quad (3.1.1)$$

За начальный вектор $\bar{\omega}^{(0)}$ можно взять, например, единичный вектор $\bar{e} = \{1; 1; \dots; 1\}^T$. Так можно найти спектральный радиус $\rho = \lambda_1$ как предел последовательности ρ_n и соответствующий ему собственный вектор $\bar{\omega}$ как предел последовательности векторов $\bar{\omega}^{(n)}$.

Пример. Рассмотрим матрицу A

$$\begin{pmatrix} -12 & 4 & 8 \\ 4 & 11 & -6 \\ 8 & -6 & 2 \end{pmatrix}$$

Возьмём нормированный вектор-столбец $\bar{\omega}^{(0)} = \{ 1/\sqrt{3}; 1/\sqrt{3}; 1/\sqrt{3} \}^T$, чтобы скалярное произведение $(\bar{\omega}^{(0)}, \bar{\omega}^{(0)}) = 1$ (т.е. $\|\bar{\omega}^{(0)}\|_E = 1$). Умножив матрицу A на $\bar{\omega}^{(0)}$, получим вектор $\bar{v}^{(1)} = A \cdot \bar{\omega}^{(0)}$:

$$\bar{v}^{(1)} = \begin{pmatrix} -12 & 4 & 8 \\ 4 & 11 & -6 \\ 8 & -6 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} = \begin{pmatrix} -12 \cdot \frac{1}{\sqrt{3}} + 4 \cdot \frac{1}{\sqrt{3}} + 8 \cdot \frac{1}{\sqrt{3}} \\ 4 \cdot \frac{1}{\sqrt{3}} + 11 \cdot \frac{1}{\sqrt{3}} - 6 \cdot \frac{1}{\sqrt{3}} \\ 8 \cdot \frac{1}{\sqrt{3}} - 6 \cdot \frac{1}{\sqrt{3}} + 2 \cdot \frac{1}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} 0 \\ 5,19615242 \\ 2,30940108 \end{pmatrix}$$

Вычислим $\rho_1 = (\bar{v}^{(1)}, \bar{\omega}^{(0)}) = 0 \cdot \left(\frac{1}{\sqrt{3}}\right) + 5,196152 \cdot \left(\frac{1}{\sqrt{3}}\right) + 2,309401 \cdot \left(\frac{1}{\sqrt{3}}\right) = 4,33333333$. Нормируем вектор $\bar{v}^{(1)}$, используя евклидову норму.

$$\|\bar{v}^{(1)}\|_E = \sqrt{0^2 + (5,19615242)^2 + (2,30940108)^2} = 5,68624069.$$

$$\bar{\omega}^{(1)} = \left\{ \frac{0}{5,68624069}; \frac{5,19615242}{5,68624069}; \frac{2,30940108}{5,68624069} \right\}^T = \{0; 0,91381155; 0,40613847\}^T.$$

Закончилась первая итерация. Сделаем вторую итерацию: вычислим $\bar{v}^{(2)}$, ρ_2 и $\bar{\omega}^{(2)}$:

$$\bar{v}^{(2)} = \begin{pmatrix} -12 & 4 & 8 \\ 4 & 11 & -6 \\ 8 & -6 & 2 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0,91381155 \\ 0,40613847 \end{pmatrix} = \begin{pmatrix} -12 \cdot 0 + 4 \cdot 0,91381155 + 8 \cdot 0,40613847 \\ 4 \cdot 0 + 11 \cdot 0,91381155 - 6 \cdot 0,40613847 \\ 8 \cdot 0 - 6 \cdot 0,91381155 + 2 \cdot 0,40613847 \end{pmatrix} = \{6,90435392; 7,61509624; -4,67059236\}^T$$

$$\rho_2 = (\bar{v}^{(2)}, \bar{\omega}^{(1)}) = 6,90435392 \cdot 0 + 7,61509624 \cdot 0,91381155 - 4,67059236 \cdot 0,40613847 = 5,06185567.$$

Вычисляем

$$\|\bar{v}^{(2)}\|_E = \sqrt{6,90435392^2 + 7,61509624^2 + (-4,67059236)^2} = 11,29044848$$

$$\bar{\omega}^{(2)} = \left\{ \frac{6,90435392}{11,29044848}; \frac{7,61509624}{11,29044848}; \frac{-4,67059236}{11,29044848} \right\}^T =$$

$$\begin{pmatrix} 0,61152167 \\ 0,67447243 \\ -0,41367642 \end{pmatrix}$$

Закончилась вторая итерация. Продолжив итерации, получим $\rho_3 = 3,45903761$, $\rho_4 = 0,59546447$, $\rho_5 = -2,41285986$, ..., $\rho_{10} = -13,49472047$, ..., $\rho_{20} = -16,91885178$, т.е. наблюдается сходимость к точному значению спектрального радиуса, равного -17 и его вектора $\bar{\omega} = \{-0,872872; 0,218218; 0,436436\}^T$, нормированное евклидовой нормой (или $\{1; -0,25; -0,5\}$, нормированное первой нормой). Напомним, что собственные вектора определяются с точностью нормирующего множителя.

вектор	$v_I (w_I)$	$v_2 (w_2)$	$v_3 (w_3)$	ρ
$\bar{\omega}^{(0)}$	0.577350	0.577350	0.577350	
$\bar{v}^{(1)}$	0.0	5.196152	2.309401	4.33333333
$\bar{\omega}^{(1)}$	0.0	0.913812	0.406138	
$\bar{v}^{(2)}$	6.904354	7.615096	-4.670592	5.06185567
$\bar{\omega}^{(2)}$	0.611522	0.674472	-0.413676	
$\bar{v}^{(3)}$	-7.949782	12.347342	0.017986	3.45903761
$\bar{\omega}^{(3)}$	-0.541345	0.840800	0.001225	
$\bar{v}^{(4)}$	9.869138	7.076067	9.373109	0.59546447
$\bar{\omega}^{(4)}$	0.643346	0.461272	0.611011	
$\bar{v}^{(5)}$	-10.763148	11.313441	1.157112	-2.41285986
$\bar{\omega}^{(5)}$	-0.687382	0.722527	0.073898	
$\bar{v}^{(6)}$	11.729881	4.754872	-9.686421	-5.34320208
$\bar{\omega}^{(6)}$	0.735961	0.298332	-0.607749	
$\bar{v}^{(7)}$	-12.500199	9.871994	2.882196	-8.00617858
$\bar{\omega}^{(7)}$	-0.772238	0.609873	0.178056	
...
$\bar{v}^{(20)}$	14.819382	-3.064180	-7.730024	-16.91885178
$\bar{\omega}^{(20)}$	0.872096	-0.180322	-0.454899	

Модификация степенного метода.

Построим последовательность векторов \bar{v} , $\bar{\omega}$ и чисел ρ :

$$\bar{v}^{(n+1)} = A \cdot \bar{\omega}^{(n)}, \quad \rho_{n+1} = v_i^{(n+1)} : |v_i^{(n+1)}| = \max_{j=1..N} (|v_j^{(n+1)}|), \quad \bar{\omega}^{(n+1)} = \frac{\bar{v}^{(n+1)}}{\rho_{n+1}}, \quad (3.1.2)$$

т.е. ρ_{n+1} - максимальная по абсолютной величине компонента вектора $\bar{v}^{(n+1)}$. Например, если $\bar{v} = (2; -4; 3)^T$, то $\rho = -4$, а $\bar{\omega} = (-0,5; 1; -0,75)^T$.

За начальный вектор $\bar{\omega}^{(0)}$ можно взять, например, единичный вектор $\bar{e} = (1; 1; \dots; 1)^T$. Под нормой в (3.1-2) будем понимать первую норму из предыдущей главы, т.е. чтобы вычислить вектор $\bar{\omega}$, нужно в векторе \bar{v} найти максимальную по модулю компоненту и разделить все компоненты вектора \bar{v} на неё. Эта компонента станет единицей, а остальные станут по модулю меньше единицы. Такой алгоритм позволяет найти спектральный

радиус $\rho = |\lambda_1|$ как предел последовательности максимальных по модулю компонент векторов $\bar{v}^{(n)}$. Пределом последовательности векторов $\bar{\omega}$ будет собственный вектор, соответствующий λ_1 .

Пример. Рассмотрим матрицу A

$$A = \begin{pmatrix} 17 & 1 & 1 \\ 1 & 17 & 2 \\ 1 & 2 & 4 \end{pmatrix}.$$

Возьмём вектор-столбец $\bar{\omega}^{(0)} = \{1; 1; 1\}^T$. Умножив матрицу A на $\bar{\omega}^{(0)}$, получим вектор $\bar{v}^{(1)}$: $\bar{v}^{(1)} = A \cdot \bar{\omega}^{(0)}$,

$$\bar{v}^{(1)} = \begin{pmatrix} 17 & 1 & 1 \\ 1 & 17 & 2 \\ 1 & 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 17 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 \\ 1 \cdot 1 + 17 \cdot 1 + 2 \cdot 1 \\ 1 \cdot 1 + 2 \cdot 1 + 4 \cdot 1 \end{pmatrix} = \begin{pmatrix} 19 \\ 20 \\ 7 \end{pmatrix}.$$

Так как $\|\bar{v}^{(1)}\|_1 = 20$, то

$$\bar{\omega}^{(1)} = \begin{pmatrix} 19/20 \\ 20/20 \\ 7/20 \end{pmatrix} = \begin{pmatrix} 0,95 \\ 1 \\ 0,35 \end{pmatrix}.$$

Вектор $\bar{v}^{(2)}$ получим, умножив матрицу A на нормированный вектор $\bar{\omega}^{(1)}$: $\bar{v}^{(2)} = A \cdot \bar{\omega}^{(1)}$,

$$\bar{v}^{(2)} = \begin{pmatrix} 17 & 1 & 1 \\ 1 & 17 & 2 \\ 1 & 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} 0,95 \\ 1 \\ 0,35 \end{pmatrix} = \begin{pmatrix} 17 \cdot 0,95 + 1 \cdot 1 + 1 \cdot 0,35 \\ 1 \cdot 0,95 + 17 \cdot 1 + 2 \cdot 0,35 \\ 1 \cdot 0,95 + 2 \cdot 1 + 4 \cdot 0,35 \end{pmatrix} = \begin{pmatrix} 17,5 \\ 18,65 \\ 4,35 \end{pmatrix}.$$

Теперь вычислим $\bar{\omega}^{(2)}$, нормируя $\bar{v}^{(2)}$: $\|\bar{v}^{(2)}\|_1 = 18,65$,

$$\bar{\omega}^{(2)} = \begin{pmatrix} 17,5/18,65 \\ 18,65/18,65 \\ 4,35/18,65 \end{pmatrix} = \begin{pmatrix} 0,9383378 \\ 1 \\ 0,233244 \end{pmatrix}.$$

Процесс продолжаем до тех пор, пока нормы $\|\bar{v}^{(n)}\|_1$ и $\|\bar{v}^{(n-1)}\|_1$ векторов $\bar{v}^{(n)}$ и $\bar{v}^{(n-1)}$ соседних итераций почти перестанут меняться, т.е. итерации прекращаются, если для заранее заданного ϵ ($\epsilon > 0$) станет справедливо

$$|\|\bar{v}^{(n)}\|_1 - \|\bar{v}^{(n-1)}\|_1| < \epsilon \quad (3.1.3)$$

Продолжим вычисления для $\epsilon = 0,01$ и занесём их в таблицу:

вектор	$v_1 (w_1)$	$v_2 (w_2)$	$v_3 (w_3)$	$\ \bar{v}^{(n)}\ _1$
$\bar{\omega}^{(0)}$	1.00000000	1.00000000	1.00000000	

$\bar{v}^{(1)}$	19.0000000	20.0000000	7.0000000	20.
$\bar{\omega}^{(1)}$	0.9500000	1.0000000	0.3500000	
$\bar{v}^{(2)}$	17.5000000	18.6500000	4.3500000	18.65
$\bar{\omega}^{(2)}$	0.9383378	1.0000000	0.2332440	
$\bar{v}^{(3)}$	17.1849866	18.4048257	3.8713137	18.40482574
$\bar{\omega}^{(3)}$	0.9337218	1.0000000	0.2103423	
$\bar{v}^{(4)}$	17.0836125	18.3544064	3.7750910	18.35440641
$\bar{\omega}^{(4)}$	0.9307636	1.0000000	0.2056776	
$\bar{v}^{(5)}$	17.0286580	18.3421188	3.7534741	18.34211884
$\bar{\omega}^{(5)}$	0.9283910	1.0000000	0.2046369	
$\bar{v}^{(6)}$	16.9872839	18.3376648	3.7469386	18.33766478
$\bar{\omega}^{(6)}$	0.9263603	1.0000000	0.2043302	
$\bar{v}^{(7)}$	16.9524546	18.3350206	3.7436810	18.33502064
$\bar{\omega}^{(7)}$	0.9245942	1.0000000	0.2041820	

Закончим итерации, так как $|18.33502064 - 18.33766478| < 0,01$.

Теперь можем утверждать, что спектральный радиус исходной матрицы равен $\rho \approx 18,33$, а соответствующий ему вектор $\bar{\omega}^T \approx \{0.924; 1.00; 0.204\}$.

Замечание 1. Если заранее не известен порядок ответа, то вместо условия для абсолютной погрешности (3.1.3) используют условие для относительной погрешности:

$$\left| \frac{\|\bar{v}^{(n-1)}\|_1 - \|\bar{v}^{(n)}\|_1}{\|\bar{v}^{(n)}\|_1} \right| < \epsilon.$$

Замечание 2. Сходимость степенного метода ухудшается, если, например, у матрицы есть не одно, а несколько вещественных максимальных по модулю одинаковых или даже близких друг к другу собственных чисел. Если геометрическая кратность m у спектрального радиуса ρ больше 1, то найденный собственный вектор $\bar{\omega}$ будет принадлежать пространству размерности m собственных векторов для ρ .

Замечание 3. Если процесс итераций сошёлся за 1 шаг ($\bar{\omega}^{(0)} = \bar{\omega}^{(1)}$), это может означать, что нулевое приближение $\bar{\omega}^{(0)}$ случайно совпало с каким-то собственным вектором. А этот вектор не обязательно совпадает с собственным вектором спектрального радиуса $\rho(A)$. Т.е. полученное $\rho_1 \neq |\rho(A)|$.

Так у матрицы

$$A = \begin{pmatrix} 4 & -4 & -20 \\ -4 & -2 & -14 \\ -20 & -14 & 14 \end{pmatrix}$$

собственными парами являются пары собственных чисел и векторов:

$$\lambda_1 = 4, \quad \bar{\omega}_1 = \{-0,8; 1; -0,2\}^T,$$

$$\lambda_2 = -20, \quad \bar{\omega}_2 = \{1; 1; 1\}^T,$$

$$\lambda_3 = 32, \quad \bar{\omega}_3 = \{-2/3; -1/3; 1\}^T.$$

Выбор в качестве $\bar{\omega}^{(0)}$ вектора $\{1; 1; 1\}^T$ приведёт к результату **-20**, хотя на самом деле $\rho(A) = 32$.

Рекомендуется сделать несколько расчётов с разными начальными векторами $\bar{\omega}^{(0)}$. Конечные результаты хотя бы двух расчётов для разных $\bar{\omega}^{(0)}$ должны совпадать, тогда Вы нашли то, что хотели.

Замечание 4. Метод вычисляет и знак максимального по модулю собственного числа, если у матрицы нет другого, равного ему собственного числа, но противоположного знака.

3.2. Метод вращений Якоби для симметричной матрицы

Рассмотрим задачу нахождения всех собственных чисел и векторов для вещественной симметричной матрицы порядка n . Будем применять к ней преобразование подобия, не изменяющее спектра (собственных чисел). Выберем наибольший по модулю элемент a_{km} матрицы, лежащий выше главной диагонали, и преобразуем матрицу так, чтобы он стал нулём. Это преобразование представляет собой поворот двухмерной плоскости, проходящей через k -ю и m -ю оси координат на специально подобранный угол φ . Преобразование осуществляется перемножением трёх матриц.

Преобразование *подобия* вида $A_1 = H^{-1} \cdot A \cdot H$ не меняет собственные числа, т.е. они для матриц A и A_1 совпадают.

Остальные элементы тоже как-то изменятся. Опять выберем наибольший по модулю элемент матрицы, не лежащий на главной диагонали и опять преобразуем (повернём) матрицу так, чтобы он стал нулём. Там, где до поворота стоял ноль, его может уже и не быть, но максимальные внедиагональные элементы будут по модулю быстро уменьшаться. Через некоторое число преобразований (поворотов-вращений) все внедиагональные элементы будут равняться почти нулю, тогда стоящие на главной диагонали числа и будут собственными числами исходной матрицы. Собственные векторы получают, перемножив все матрицы поворотов. Вычисленная таким образом матрица будет иметь своими столбцами собственные векторы.

Пример. Вычислить все собственные числа и векторы у матрицы

$$A = \begin{pmatrix} 17 & 1 & 1 \\ 1 & 17 & 2 \\ 1 & 2 & 4 \end{pmatrix}. \quad (3.3.1)$$

Ищем максимальный по модулю элемент a_{km} , лежащий выше главной диагонали ($k < m$). В этой матрице это элемент $a_{23} = 2$, $k=2$, $m=3$.

Вычислим угол поворота по формуле

$$\varphi = \frac{1}{2} \arctan \left(\frac{2 \cdot a_{km}}{a_{kk} - a_{mm}} \right). \quad (3.3.2)$$

Подставив в эту формулу $a_{23} = 2$, $a_{22} = 17$, $a_{33} = 4$, получим $\varphi = 0,1492494658$ радиан. Матрицу поворота H строим, изменяя в единичной матрице E только элементы, стоящие на пересечении строк и столбцов с номерами $k=2$ $m=3$:

$$\begin{aligned} h_{kk} &= h_{22} = \cos \varphi, & h_{km} &= h_{23} = -\sin \varphi, \\ h_{mk} &= h_{32} = \sin \varphi, & h_{mm} &= h_{33} = \cos \varphi; \end{aligned} \quad (3.3.3)$$

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad H = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix}.$$

Матрица первого поворота

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0,9888830 & -0,1486960 \\ 0 & 0,1486960 & 0,9888830 \end{pmatrix}.$$

Матрица поворота ортогональна, т.е. её обратная равна её транспонированной:

$$H_1^{-1} = H_1^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0,9888830 & 0,1486960 \\ 0 & -0,1486960 & 0,9888830 \end{pmatrix}.$$

Поворот осуществляется перемножением трёх матриц: $H_1^{-1} \cdot A \cdot H_1$. Умножение матриц не коммутативно (нельзя менять левый и правый сомножители местами), но ассоциативно, т.е. $H_1^{-1} \cdot A \cdot H_1 = (H_1^{-1} \cdot A) \cdot H_1 = H_1^{-1} \cdot (A \cdot H_1)$.

Напомним, что если вычисляется произведение $C = A \cdot B$ двух квадратных матриц размерности $n \times n$, то элемент c_{ij} , стоящий в матрице C на пересечении строки i и столбца j , вычисляется как скалярное произведение i -й строки матрицы A и j -го столбца матрицы B :

$$c_{ij} = \sum_{p=1}^n a_{ip} \cdot b_{pj}.$$

Вычислим вначале $H_1^{-1} \cdot A =$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0,988883 & 0,148696 \\ 0 & -0,148696 & 0,988883 \end{pmatrix} \times \begin{pmatrix} 17 & 1 & 1 \\ 1 & 17 & 2 \\ 1 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 17 & 1 & 1 \\ 1,137579 & 17,1084 & 2,57255 \\ 0,840187 & -0,5500658 & 3,65814 \end{pmatrix}$$

Теперь полученную матрицу справа умножим на H_1 :

$$\begin{pmatrix} 17 & 1 & 1 \\ 1,137579 & 17,1084 & 2,57255 \\ 0,840187 & -0,5500658 & 3,65814 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0,988883 & -0,148696 \\ 0 & 0,148696 & 0,988883 \end{pmatrix} =$$

$$= A_1 = \begin{pmatrix} 17 & 1,137579 & 0,840187 \\ 1,137579 & 17,30074 & 0 \\ 0,840187 & 0 & 3,699265 \end{pmatrix}.$$

Закончился первый поворот. В новой матрице A_1 на месте a_{23} и a_{32} стоят точные нули. Новая матрица A_1 тоже симметрична. Сделаем ещё один поворот. Максимальный по модулю элемент над главной диагональю равен 1,137579 и находится на пересечении первой строки и второго столбца ($a_{km} = 1,137579$, $k = 1$, $m = 2$). Вычислим угол поворота

$$\varphi = \frac{1}{2} \arctan\left(\frac{2 \cdot 1,137579}{17 - 17,30074}\right) = -0,7196880213.$$

Теперь можно вычислить матрицу второго поворота

$$H_2 = \begin{pmatrix} 0,7520114 & 0,6591501 & 0 \\ -0,6591501 & 0,7520114 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Запишем обратную для неё и выполним необходимые перемножения матриц:

$$H_2^{-1} = \begin{pmatrix} 0,7520114 & -0,6591501 & 0 \\ 0,6591501 & 0,7520114 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$H_2^{-1} A_1 = \begin{pmatrix} 0,7520114 & -0,6591501 & 0 \\ 0,6591501 & 0,7520114 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \begin{pmatrix} 17 & 1,137579 & 0,840187 \\ 1,137579 & 17,30074 & 0 \\ 0,840187 & 0 & 3,699265 \end{pmatrix} =$$

$$= \begin{pmatrix} 12,03436 & -10,54831 & 0,6318302 \\ 12,06102 & 13,76019 & 0,5538093 \\ 0,840187 & 0 & 3,699265 \end{pmatrix};$$

$$(H_2^{-1} A_1) H_2 = \begin{pmatrix} 12,03436 & -10,54831 & 0,6318302 \\ 12,06102 & 13,76019 & 0,5538093 \\ 0,840187 & 0 & 3,699265 \end{pmatrix} \times \begin{pmatrix} 0,7520114 & 0,6591501 & 0 \\ -0,6591501 & 0,7520114 & 0 \\ 0 & 0 & 1 \end{pmatrix} =$$

$$= A_2 = \begin{pmatrix} 16,00289 & 0 & 0,6318302 \\ 0 & 18,29784 & 0,5538093 \\ 0,6318302 & 0,5538093 & 3,699265 \end{pmatrix}.$$

Как видно из матрицы A_2 , она симметрична и имеет нулевые элементы там, где у матрицы A_1 были максимальные по модулю.

Так после каждого вращения элементы матриц A_n (кроме стоящих на главной диагонали) будут по модулю уменьшаться к нулю (свойство ортогональных матриц H_i). После седьмого поворота

$$A_7 = \begin{pmatrix} 16,03490 & 10^{-7} & 10^{-8} \\ 10^{-7} & 18,31907 & -10^{-7} \\ 10^{-8} & -10^{-7} & 3,646025 \end{pmatrix}.$$

Элементы, стоящие на главной диагонали, и есть собственные числа исходной матрицы, вычисленные с точностью 10^{-7} :

$$\lambda_1 = 16,03490, \lambda_2 = 18,31907, \lambda_3 = 3,646025.$$

Для получения собственных векторов надо перемножить все матрицы вращения H и получить матрицу V :

$$H_1 \cdot H_2 \cdot \dots \cdot H_7 = V = \begin{pmatrix} 0,7427887 & -0,6693896 & 0,01351136 \\ 0,6524570 & 0,7282323 & 0,2097083 \\ -0,1502160 & -0,1469534 & 0,9776706 \end{pmatrix}.$$

После нормировки столбцов матрицы V получим, что первый столбец преобразованной матрицы V - собственный вектор для λ_1 . Второй столбец - собственный вектор для λ_2 , третий - для λ_3 :

$$\lambda_1 = 16,03490, \bar{v}_1 = \begin{pmatrix} 1 \\ 0,8783884 \\ -0,2022325 \end{pmatrix}; \lambda_2 = 18,31907, \bar{v}_2 = \begin{pmatrix} -0,9191978 \\ 1 \\ -0,2017947 \end{pmatrix},$$

$$\lambda_3 = 3,646025, \bar{v}_3 = \begin{pmatrix} 0,01381995 \\ 0,2144980 \\ 1 \end{pmatrix}.$$

Замечание 1. Если при вычислении угла φ происходит «деление на ноль» (в формуле (3.3.2) $a_{kk} = a_{mm}$) примите значение угла φ равное $\pi/4$ или $-\pi/4$ в зависимости от знака a_{km} .

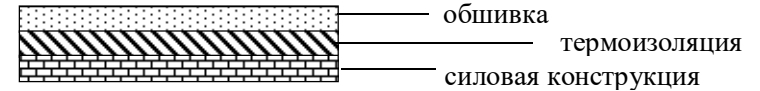
Замечание 2. У всех матриц A_i сохраняется симметричность и след (число $\sum_{p=1}^n a_{pp}$ одинаковое для всех матриц A_i).

Замечание 3. Метод вращений Якоби для симметричных матриц не является самым быстрым, но по сравнению со своими быстрыми конкурентами (например, QL алгоритмом) он проще для программирования.

Замечание 4. Матрицы общего вида (несимметричные) сначала приводят к какому-либо специальному виду (например, к верхней почти треугольной форме Хессенберга), а затем применяют специальные алгоритмы (например, QR алгоритм Кублановской – Френсиса). Есть специальные алгоритмы для трёхдиагональных, разреженных и других типов матриц.

4. Методы решения трансцендентных уравнений

В качестве примера постановки задачи рассмотрим задачу расчёта установившейся температуры поверхности летательного аппарата. Пусть силовая конструкция имеет два внешних слоя: теплоизоляция и тонкая обшивка :



Из уравнения теплового баланса на поверхности летательного аппарата (конвективный тепловой поток от нагретого газа к поверхности тела + тепловой поток от солнечной радиации = тепловой поток излучения с поверхности тела) получается уравнение для нахождения температуры поверхности летательного аппарата T_ω :

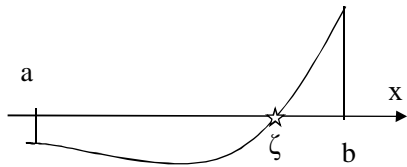
$$\delta \varepsilon T_\omega^4 - C_p q V_\delta \rho_\delta St_\delta [T_\delta (1 + 0,2 r M_\delta^2) - T_\omega] - A_s \beta_s G_s \sin \theta = 0, \quad (4.1)$$

где C_p – теплоёмкость воздуха; ρ_δ , V_δ , T_δ , M_δ – параметры потока на внешней границе пограничного слоя; St_δ – число Стантона; r – коэффициент восстановления; A_s – коэффициент затенения; β_s , ε – параметры материала обшивки; G_s – интенсивность солнечного излучения; θ – угол наклона поверхности к солнечным лучам; δ – постоянная Стефана–Больцмана.

Уравнение (4.1) является уравнением четвёртой степени относительно T_ω .

Рассмотрим некоторые методы решения уравнений $f(x) = 0$. Оставим в стороне проблему отделения корней, т.е. нахождение отрезка $[a; b]$, который содержит корень ζ уравнения. Нахождение такого (таких) отрезков может быть задачей более сложной, чем вычисление корня. Но пусть такой отрезок уже известен.

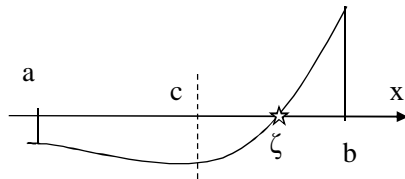
4.1. Метод половинного деления (дихотомии, бисекций)



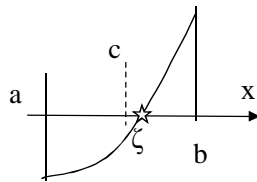
Пусть известен отрезок $[a; b]$, на котором непрерывная функция $y=f(x)$ меняет знак, т.е. в точках a и b функция имеет значения разных знаков: $f(a) \cdot f(b) \leq 0$. Тогда существует точка $x=\zeta$ такая, что $f(\zeta) = 0$.

Разделим отрезок $[a; b]$ точкой c пополам: $c = (a+b)/2$.

Точка c разделила отрезок $[a; b]$ на два отрезка: $[a; c]$ и $[c; b]$ (левый и правый). Но только на одном из них функция будет менять знак: на $[a; c]$ (если $f(a) \times f(c) < 0$) или на $[c; b]$ (если $f(c) \times f(b) < 0$).

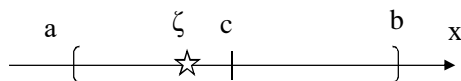


Для данного конкретного рисунка выбираем правый отрезок. Теперь точка c стала новой точкой a , точка b осталась на месте.



Новый отрезок $[a; b]$ делится новой точкой $c = (a+b)/2$ на два отрезка: $[a; c]$ и $[c; b]$. И вновь из двух отрезков выбираем тот, на котором функция будет менять знак. Для данного конкретного рисунка видно, что это опять правый отрезок.

Но при следующем делении может быть выбран уже левый отрезок. С каждым делением длина отрезка $[a; b]$ уменьшается в два раза и скоро станет меньше $2 \cdot \varepsilon$ ($\varepsilon > 0$ и задаётся заранее) .



Тогда расстояние от ζ (ζ = неизвестное точное значение) до c не больше расстояния от c до a , т.е.

$$|c - \zeta| \leq |(b - a) / 2| \leq \varepsilon. \quad (4.1.1)$$

Пример. Найти положительный корень уравнения

$$x^2 - 2 = 0. \quad (4.1.2)$$

Пусть $a=0$, тогда $f(a)=-2$. Рассмотрим точку $b=8$, тогда $f(b)=62$. Так как $f(a) \times f(b) < 0$, то на отрезке $[0; 8]$ есть корень. Середина этого отрезка – $c=(8+0)/2=4$. $f(4) = 14 > 0$, следовательно, из двух отрезков $[0;4]$ и $[4;8]$ выбираем тот, на котором функция принимает значения разных знаков: $f(a) \times f(c) = -2 \times 14 = -28 < 0$, следовательно, новая точка $b=4$. Теперь новый отрезок $[0;4]$. Его делим пополам точкой $c=(4+0)/2 = 2$. Получили опять два отрезка: $[0;2]$ и $[2;4]$; $f(2) = 2$. Опять ($f(a) \times f(c) = -2 \times 2 = -4 < 0$) из двух выбираем левый отрезок, т.е. точка a осталась той же, а точка b стала равна 2. Продолжим вычисления, записав результаты в таблицу:

a	c	b	$f(a)$	$f(c)$	$f(b)$
0.000000	4.000000	8.000000	-2.000000	14.000000	62.000000
0.000000	2.000000	4.000000	-2.000000	2.000000	14.000000
0.000000	1.000000	2.000000	-2.000000	-1.000000	2.000000
1.000000	1.500000	2.000000	-1.000000	0.250000	2.000000
1.000000	1.250000	1.500000	-1.000000	-0.437500	0.250000
1.250000	1.375000	1.500000	-0.437500	-0.109375	0.250000
1.375000	1.437500	1.500000	-0.109375	0.066406	0.250000
1.375000	1.406250	1.437500	-0.109375	-0.022461	0.066406
1.406250	1.421875	1.437500	-0.022461	0.021729	0.066406
1.406250	1.414062	1.421875	-0.022461	-0.000427	0.021729
1.414062	1.417969	1.421875	-0.000427	0.010635	0.021729
1.414062	1.416016	1.417969	-0.000427	0.005100	0.010635

Последний отрезок $[1.414062; 1.417969]$ имеет длину 0,0039, следовательно его середина 1.416016 может считаться корнем с точностью не хуже 0,002. Итак, корень $\approx 1,41$.

Сходимость метода не высокая, но метод неприхотлив к поведению функции около корня. Главное, чтобы на границах отрезка функция принимала значения разных знаков. Это и есть условие применимости метода (ещё, конечно, условие непрерывности функции).

Замечание. Выполнение условия $f(a) \cdot f(b) < 0$ не означает, что существующий на отрезке $[a; b]$ корень единственен. А означает, что количество корней на этом отрезке (с учётом их кратности) нечётно, и метод дихотомии найдёт один из них.

4.2. Метод хорд (секущих)

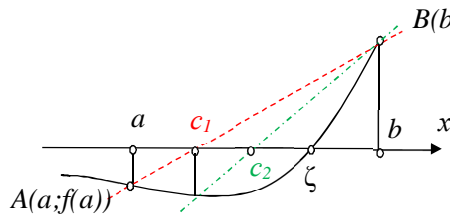
Пусть функция $f(x)$ (уравнения $f(x) = 0$) непрерывна на отрезке $[a; b]$ и имеет значения разных знаков на границе отрезка: $(f(a) \times f(b) < 0)$.

Через точки $A(a; f(a))$ и $B(b; f(b))$ проведём прямую линию, которая пересечёт ось ox в некоторой точке c . Точка c находится из соотношения

$$c = \frac{a \cdot f(b) - b \cdot f(a)}{f(b) - f(a)}.$$

Точка c разделила отрезок $[a; b]$ на два отрезка, но разные знаки значения функции будут иметь только на одном из них: на $[a; c]$ (если $f(a) \times f(c) < 0$) или на $[c; b]$ (если $f(c) \times f(b) < 0$). В ситуации,

которая представлена на рисунке, мы выбираем правый отрезок. В новом отрезке $[a; b]$ через точки $A(a; f(a))$ и $B(b; f(b))$ проведём прямую линию, которая пересечёт ось ox в некоторой новой точке c_2



(4.2.1). И так далее.

Особенность метода в том, что раз выбрав, например, правый отрезок, на каждом шаге итераций выбирать придётся тоже правый отрезок.

Все точки c приближаются к корню с одной стороны (имеет место односторонняя сходимость). Итерации можно остановить, когда точки c на соседних двух шагах почти одинаковы и значения функции $f(x)$ в точке c почти ноль ($\varepsilon > 0$, $\varepsilon \approx 0$):

$$\begin{cases} |c_n - c_{n-1}| < \varepsilon \\ |f(c_n)| < \varepsilon \end{cases} \quad (4.2.2)$$

Пример. Найти положительный корень уравнения $x^2 - 2 = 0$

на отрезке $[0; 8]$.

Так как $f(0) = -2$, $f(8) = 62$, то $f(0) \cdot f(8) < 0$. По точкам $A(0; -2)$ и $B(8; 62)$ находим точку c :

$c = \frac{0 \cdot 62 - 8 \cdot (-2)}{62 - (-2)} = \frac{16}{64} = 0,25$ $f(c) = -1,9375$. Теперь новая точка $A(0,25; -1,9375)$, так как $f(c) \cdot f(b) < 0$. Точка $B(8; 62)$ не изменилась. По координатам новых точек A и B вычисляем новую точку c :

$$c = \frac{0,25 \cdot 62 - 8 \cdot (-1,9375)}{62 - (-1,9375)} = 0,4848484848.$$

Продолжим процесс итераций, записывая результаты в таблицу:

a	c	b	$f(a)$	$f(c)$	$f(b)$
0.0		8.	-2.		62.
	0.25			-1.93750000	
0.25		8.	-1.93750000		62.
	0.48484848			-1.76492195	
0.48484848		8.	-1.76492195		62.
	0.69285714			-1.51994898	
0.69285714		8.	-1.51994898		62.
	0.86770748			-1.24708373	
0.86770748		8.	-1.24708373		62.
	1.00833951			-0.98325143	
1.00833951		8.	-0.98325143		62.
	1.11748853			-0.75121938	
1.11748853		8.	-0.75121938		62.
	1.19988177			-0.56028375	
1.19988177		8.	-0.56028375		62.
	1.26078296			-0.41042634	
1.26078296		8.	-0.41042634		62.
	1.30510171			-0.29670953	
1.30510171		8.	-0.29670953		62.
	1.33698847			-0.21246184	
1.33698847		8.	-0.21246184		62.
	1.35974332			-0.15109809	
1.35974332		8.	-0.15109809		62.
	1.37588673			-0.10693572	
1.37588673		8.	-0.10693572		62.
	1.38729212			-0.07542056	

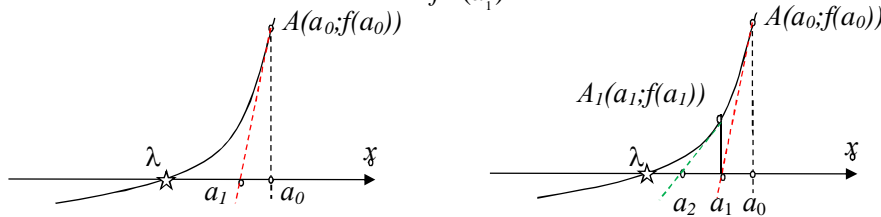
Согласно соотношению (4.2.2) мы нашли корень **1.38** с точностью 0,1.

Этот метод как и метод дихотомии сходится медленно. Особенностью является односторонняя сходимость, это уменьшает ценность метода.

Замечание. При проверке условий (4.2-2) в неравенствах могут использоваться разные ε : ε_1 и ε_2 .

4.3. Метод касательных (Ньютона)

Пусть функция $f(x)$ дважды непрерывно дифференцирована. Пусть в окрестности корня уравнения $f(x)=0$, в некоторой точке a_0 выполняется соотношение $f(a_0) \cdot f''(a_0) > 0$. Из точки $A(a_0; f(a_0))$ проведём касательную к графику функции $y=f(x)$. Она пересечёт ось OX в точке a_1 : $a_1 = a_0 - \frac{f(a_0)}{f'(a_0)}$. Из точки $A_1(a_1; f(a_1))$ проведём касательную к графику функции $y=f(x)$. Она пересечёт ось OX в точке $a_2 = a_1 - \frac{f(a_1)}{f'(a_1)}$.



Таким образом, можно получить бесконечную последовательность точек a_n , которые быстро будут приближаться к корню λ . Преимущество метода в его высокой скорости, но есть и недостатки. Это необходимость уметь вычислять производную от функции $f(x)$, а это бывает не всегда возможно. Ещё желательно на каждом i -м шаге проверять выполнение неравенства $f(a_i) \cdot f''(a_i) > 0$.

Пример. На отрезке $[0;8]$ найти положительный корень уравнения $x^2 - 2 = 0$.

Вычислим значения функции $f(x) = x^2 - 2$ в точках 0 и 8: $f(0) = -2$, $f(8) = 62$, $f''(0) = f''(8) = 2$. Так как $f(8) \cdot f''(8) > 0$, за начальную точку a_0 можно взять $a_0 = 8$. Поскольку $f'(x) = 2x$ и $f'(a_0) = 2 \cdot 8 = 16$, вычисляем

$$a_1 = a_0 - \frac{f(a_0)}{f'(a_0)} = 8 - \frac{62}{16} = 4\frac{1}{8} = 4,125. \text{ Вычислив } f(a_1) = f(4,125) = 15,015625$$

$f'(a_1) = 8,25$, можно посчитать и

$$a_2 = a_1 - \frac{f(a_1)}{f'(a_1)} = 4,125 - \frac{15,015625}{8,25} = 2,30492424.$$

Продолжим вычисления, записывая результаты в таблицу:

n	a_{n-1}	$f(a_{n-1})$	$f'(a_{n-1})$	f''	a_n
0	8.00000000	62.00000000	16.00000000	2	4.12500000
1	4.12500000	15.01562500	8.25000000	2	2.30492424
2	2.30492424	3.31267576	4.60984848	2	1.58631586
3	1.58631586	0.51639801	3.17263172	2	1.42354941

4	1.42354941	0.02649292	2.84709882	2	1.41424418
5	1.41424418	0.00008659	2.82848835	2	1.41421356
6	1.41421356	0.00000000	2.82842713	2	1.41421356

Итак, констатируем высокую скорость сходимости метода касательных. К недостаткам метода (кроме необходимости вычислять производную) можно отнести одностороннюю сходимость к корню.

Замечание. В случае трудности вычисления производной $f'(x)$, её вычисляют не на каждой итерации, а используют одно и то же значение производной на нескольких итерациях.

4.4. Метод итераций

Пусть уравнение $f(x)=0$ удалось преобразовать к равносильному уравнению $x=g(x)$, функция $g(x)$ которого обладает свойством: в окрестности корня $x=\lambda$

$$|g'(x)| \leq q < 1, \quad (4.4.1)$$

то итерационный процесс $x_{n+1}=g(x_n)$ будет сходиться к корню λ уравнения $f(x)=0$. Причём, сходимость будет тем быстрее, чем меньше q . Величины $|x_n - \lambda|$ будут убывать не медленнее бесконечно-убывающей прогрессии со знаменателем q . Процесс останавливаем, если для некоторого малого положительного ε

$$\begin{cases} \frac{(x_n - x_{n-1})^2}{|2x_{n-1} - x_n - x_{n-2}|} < \varepsilon \\ |f(x_n)| < \varepsilon \end{cases} \quad (4.4.2)$$

Скорость сходимости метода может быть разной (и хуже, чем у метода половинного деления, и почти такой же высокой, как у метода касательных). Сходимость может быть как односторонней, так и двухсторонней (точки x_n могут ложиться как с одной стороны от корня, так и поочерёдно с разных сторон).

Пример. Вычислить с точностью $\varepsilon=10^{-3}$ корень уравнения $f(x)=0$:

$$x^3 - x^2 + x - 5 = 0. \quad (4.4.3)$$

Преобразуем это уравнение к виду $x=g(x)$:

$$x = \sqrt[3]{x^2 - x + 5}. \quad (4.4.4)$$

Так как $f(0) = -5$, $f(8) = 451$, то $f(0)$ и $f(8)$ разных знаков. Тогда на отрезке $[0;8]$ есть корень. Вычислим $g'(x) = \frac{2 \cdot x - 1}{3 \cdot \sqrt[3]{(x^2 - x + 5)^2}}$ в точке $x_0=8$:

$g'(8) = 0.32266 < 1$, значит имеет смысл вычислять $x_1 = \sqrt[3]{x_0^2 - x_0 + 5} = \sqrt[3]{8^2 - 8 + 5} = 3,9364971831$. Для x_1 вычислим $g'(x_1) = 0.35264 < 1$, вычисляем

$x_2 = \sqrt[3]{x_1^2 - x_1 + 5} = 2.0760625983$. И так далее, контролируя на каждом шаге выполнение (4.4.1) и проверяя (4.4.2).

№	x	$f(x)$	$g(x)$	$g'(x)$
1	8,0000000000	451	3,9364971831	0,32266
2	3,9364971831	44,44048711	2,5488788043	0,35264
3	2,5488788043	7,61160853	2,0760625983	0,31692
4	2,0760625983	1,71393104	1,9340111227	0,28091
5	1,9340111227	0,42758541	1,8951295294	0,26618
6	1,8951295294	0,11000150	1,8848646465	0,26180
7	1,8848646465	0,02853631	1,8821834155	0,26061
8	1,8821834155	0,00741909	1,8814850755	0,26030
9	1,8814850755	0,00192998	1,8813033265	0,26022

4.5. Метод парабол (Мюллера)

Пусть в окрестности корня уравнения $f(x)=0$ известны три точки (возможно вычислены другим методом) x_1, x_2, x_3 . Вычислив значения функции $f(x)$ в этих точках, получим $M_1(x_1; f(x_1))$, $M_2(x_2; f(x_2))$, $M_3(x_3; f(x_3))$. По точкам M_1, M_2, M_3 можно построить параболу $y = ax^2 + bx + c$. Коэффициенты a, b, c можно найти из системы

$$\begin{cases} ax_1^2 + bx_1 + c = y_1 \\ ax_2^2 + bx_2 + c = y_2 \\ ax_3^2 + bx_3 + c = y_3 \end{cases},$$

решив её методом Гаусса, или воспользовавшись следующими формулами:

$$a = \left(\frac{y_1 - y_2}{x_1 - x_2} - \frac{y_1 - y_3}{x_1 - x_3} \right) \cdot \frac{1}{x_2 - x_3}, \quad b = \left(\frac{y_1 - y_2}{x_1 - x_2} \right) - a \cdot (x_1 + x_2), \quad c = y_1 - ax_1^2 + bx_1$$

Из двух корней уравнения

$$ax^2 + bx + c = 0 \quad (4.5.1)$$

выбираем тот, который по ближе к x_3 . Пусть это x_4 . Теперь по точкам $M_2(x_2; f(x_2))$, $M_3(x_3; f(x_3))$, $M_4(x_4; f(x_4))$ строим новую параболу $y = ax^2 + bx + c$ и у квадратного уравнения $ax^2 + bx + c = 0$ выбираем ближайший к x_4 корень (пусть это x_5). Процесс останавливаем, если для заданного малого положительного ε

$$\begin{cases} |x_{n+1} - x_n| < \varepsilon \\ |f(x_{n+1})| < \varepsilon \end{cases} \quad (4.4.2)$$

Скорость сходимости метода хуже, чем у метода касательных, но лучше чем у метода половинного деления. В случае кратных корней

сходимость ухудшается. Преимущество метода в том, что он может найти и комплексный корень.

5. Решение систем нелинейных уравнений.

Рассмотрим несколько способов решения системы из k нелинейных уравнений с k неизвестными:

$$\begin{cases} f_1(x_1, x_2, x_3, \dots, x_k) = 0 \\ f_2(x_1, x_2, x_3, \dots, x_k) = 0 \\ \dots \\ f_k(x_1, x_2, x_3, \dots, x_k) = 0 \end{cases} \quad (5.1)$$

5.1. Метод Ньютона (линеаризации)

Пусть известно приближение $\bar{\chi}^{(n)} = (x_1^{(n)}, x_2^{(n)}, \dots, x_k^{(n)})^T$ к вектору-решению системы (5.1), вычислим его уточнение по формуле

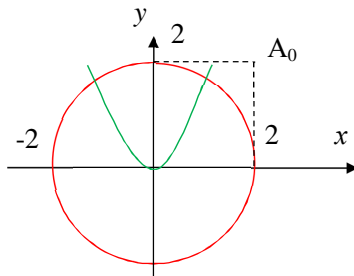
$$\bar{\chi}^{(n+1)} = \bar{\chi}^{(n)} - J_{(n)}^{-1} \cdot \bar{f}^{(n)}(\bar{\chi}^{(n)}) \quad (5.1.1)$$

Тут $J_{(n)}^{-1}$ – обратная матрица для матрицы Якоби $J_{(n)}$ (матрица частных производных, состоящая из элементов $j_{pq}^{(n)} = \frac{\partial f_p}{\partial x_q}$, вычисленных в точке $\bar{\chi}^{(n)}$); $\bar{f}^{(n)}(\bar{\chi}^{(n)})$ – вектор значений функций, стоящих в левой части (5.1).

Чем ближе выберем нулевое (начальное) приближение (например, по графику, или по каким-то другим соображениям), тем быстрее итерационный процесс (5.1-1) сойдётся к решению системы (5-1).

Пример. Решить систему
$$\begin{cases} x^2 + y^2 - 4 = 0 \\ x^2 - y = 0 \end{cases}$$

Решение. Начальное приближение выберем из рисунка:



Начнём вычисление корня с точки $A_0(2;2)$, т.е. $x_0=2$, $y_0=2$. Вычислим частные производные, чтобы сформировать матрицу Якоби:

$$\begin{pmatrix} \frac{\partial f_1}{\partial x} = 2x, & \frac{\partial f_1}{\partial y} = 2y \\ \frac{\partial f_2}{\partial x} = 2x, & \frac{\partial f_2}{\partial y} = -1 \end{pmatrix}, \text{ поэтому } J_{(n)} = \begin{pmatrix} 2 \cdot x_{(n)} & 2 \cdot y_{(n)} \\ 2 \cdot x_{(n)} & -1 \end{pmatrix}.$$

На каждом шаге итерации приходится вычислять обратную матрицу $J_{(n)}^{-1}$, например методом Гаусса. Но в случае, когда уравнений всего два, воспользуемся формулой для вычисления обратной матрицы второго порядка:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \quad (5.1.2)$$

Для нашего случая

$$J_{(n)}^{-1} = \frac{-1}{2x_{(n)} + 4x_{(n)}y_{(n)}} \begin{pmatrix} -1 & -2 \cdot y_{(n)} \\ -2 \cdot x_{(n)} & 2 \cdot x_{(n)} \end{pmatrix} = \frac{1}{2x_{(n)} + 4x_{(n)}y_{(n)}} \begin{pmatrix} 1 & 2 \cdot y_{(n)} \\ 2 \cdot x_{(n)} & -2 \cdot x_{(n)} \end{pmatrix}.$$

Выпишем окончательную формулу для совершения итераций при решении данного примера:

$$\begin{pmatrix} x_{(n+1)} \\ y_{(n+1)} \end{pmatrix} = \begin{pmatrix} x_{(n)} \\ y_{(n)} \end{pmatrix} - \frac{1}{2x_{(n)} + 4 \cdot x_{(n)} \cdot y_{(n)}} \begin{pmatrix} 1 & 2 \cdot y_{(n)} \\ 2 \cdot x_{(n)} & -2 \cdot x_{(n)} \end{pmatrix} \cdot \begin{pmatrix} x_{(n)}^2 + y_{(n)}^2 - 4 \\ x_{(n)}^2 - y_{(n)} \end{pmatrix} \quad (5.1.3)$$

Если $x_0=2$, $y_0=2$, то

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} - \frac{1}{2 \cdot 2 + 4 \cdot 2 \cdot 2} \begin{pmatrix} 1 & 2 \cdot 2 \\ 2 \cdot 2 & -2 \cdot 2 \end{pmatrix} \cdot \begin{pmatrix} 2^2 + 2^2 - 4 \\ 2^2 - 2 \end{pmatrix} = \begin{pmatrix} 1,4 \\ 1,6 \end{pmatrix}.$$

Сделаем вторую итерацию, вычислив x_2 , y_2 :

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1,4 \\ 1,6 \end{pmatrix} - \frac{1}{2 \cdot 1,4 + 4 \cdot 1,4 \cdot 1,6} \begin{pmatrix} 1 & 2 \cdot 1,6 \\ 2 \cdot 1,4 & -2 \cdot 1,4 \end{pmatrix} \cdot \begin{pmatrix} 1,4^2 + 1,6^2 - 4 \\ 1,4^2 - 1,6 \end{pmatrix} = \begin{pmatrix} 1,2578 \\ 1,5619 \end{pmatrix}.$$

т.е. $x_2=1,2578$, $y_2=1,5619$. Теперь точку $A_2(x_2; y_2)$ считаем старой и вычислим координаты новой точки $A_3(x_3; y_3)$:

$$\begin{pmatrix} x_3 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1,2578 \\ 1,5619 \end{pmatrix} - \frac{1}{2 \cdot 1,2578 + 4 \cdot 1,2578 \cdot 1,5619} \begin{pmatrix} 1 & 2 \cdot 1,5619 \\ 2 \cdot 1,2578 & -2 \cdot 1,2578 \end{pmatrix} \cdot \begin{pmatrix} 1,2578^2 + 1,5619^2 - 4 \\ 1,2578^2 - 1,5619 \end{pmatrix} = \begin{pmatrix} 1,249648 \\ 1,561553 \end{pmatrix}$$

Продолжим вычисления и запишем результаты в таблицу:

n	x_n	y_n	$\ \overline{F}\ _1$
0	2	2	4
1	1.4	1.6	0.52
2	1.2578	1.5619	0,022
3	1.24964782178751	1.56155284284615	0,0000067
4	1.24962106797405	1.56155281280883	7×10^{-10}
5	1.24962106768765	1.56155281280883	10^{-16}

Второй корень системы вычисляется аналогично, начиная с точки $A_0(-2;2)$.

Как видно из таблицы, сходимость очень быстрая. Процесс останавливаем, когда $\max(|x_n - x_{n-1}|, |y_n - y_{n-1}|) < \varepsilon$ и $\|\overline{F}\|_1 < \varepsilon$, т.е. когда точки A_n и A_{n-1} почти совпадают и значения всех функций в системе – почти нули.

Для упрощения вычислительного процесса обратную матрицу к матрице Якоби можно вычислять не на каждом шаге итерационного процесса.

5.2. Метод простой итерации

Рассмотрим метод решения системы нелинейных уравнений специального вида. Пусть система (5.1) выглядит (или может быть преобразована к виду) как

$$\begin{cases} x_1 = f_1(x_2, x_3, \dots, x_k) \\ x_2 = f_2(x_1, x_3, \dots, x_k) \\ \dots \\ x_k = f_k(x_1, x_2, x_3, \dots, x_{k-1}) \end{cases} \quad (5.2.1)$$

Подберём начальный вектор $\overline{\chi}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_k^{(0)})^O$ так, чтобы кака-нибудь норма матрицы

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_k} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_k} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_k}{\partial x_1} & \frac{\partial f_k}{\partial x_2} & \dots & \frac{\partial f_k}{\partial x_k} \end{pmatrix} \quad (5.2.2)$$

была бы меньше единицы, тогда итерационный процесс

$$\begin{cases} x_1^{(n+1)} = f_1(x_2^{(n)}, x_3^{(n)}, \dots, x_k^{(n)}) \\ x_2^{(n+1)} = f_2(x_1^{(n)}, x_3^{(n)}, \dots, x_k^{(n)}) \\ \dots \\ x_k^{(n+1)} = f_k(x_1^{(n)}, x_2^{(n)}, x_3^{(n)}, \dots, x_{k-1}^{(n)}) \end{cases} \quad (5.2.3)$$

будет заведомо сходиться к точному решению (медленнее метода Ньютона). Процесс итераций остановим, если норма вектора невязки $(\overline{\chi}^{(n)} - \overline{\chi}^{(n-1)})$ будет меньше некоторого заранее заданного ε ($\varepsilon > 0$).

Пример. Возьмём тот же пример, что и в предыдущем параграфе:

$$\begin{cases} x^2 + y^2 - 4 = 0 \\ x^2 - y = 0 \end{cases}$$

Решение. Преобразуем систему к виду

$$\begin{cases} x = \sqrt{y} \\ y = \sqrt{4 - x^2} \end{cases} \quad (5.2.4)$$

За начальное приближение возьмём ту же, что и в разд. 5.1 точку $(2; 2)$. Тогда результатом первой итерации (после подстановки $x_0=2$ и $y_0=2$ в правую часть системы (5.2.4)) станет точка $x_1=\sqrt{2}$, $y_1=0$. Теперь в правую часть системы (5.2.4) подставим x_1 и y_1 и получим x_2 и y_2 и т.д. Сходимость процесса медленнее метода Ньютона, но программируется проще:

```
n := 0 ; x := 2.0 ; y := 2.0 ; d := 100;
While (n<155) and (d>0.01) do
begin
  n := n+1;
  x1 := sqrt(y) ; y1 := sqrt(4 - x*x) ;
  d := sqrt( sqr(x-x1) + sqr(y-y1) ) ;
  Writeln (ff, n:5, x1:15:10, y1:15:10, d:10:5) ;
  x := x1 ; y := y1 ;
end;
```

Результаты расчёта до достижения точности $\varepsilon < 0,01$ представлены в таблице:

n	x	y	d
1	1.4142135624	0.0000000000	2.08402
2	0.0000000000	1.4142135624	2.00000
3	1.1892071150	2.0000000000	1.32565
4	1.4142135624	1.6080380710	0.45195
5	1.2680844100	1.4142135624	0.24274
6	1.1892071150	1.5465968864	0.15410
7	1.2436224855	1.6080380710	0.08207

8	1.2680844100	1.5663342918	0.04835
9	1.2515327769	1.5465968864	0.02576
10	1.2436224855	1.5600210602	0.01558
11	1.2490080305	1.5663342918	0.00830

5.3. Метод Зейделя

Идея метода в том, что для вычисления компонентов нового вектора $\bar{\chi}^{(n)}$ использовать не только компоненты старого вектора $\bar{\chi}^{(n-1)}$, но и только что вычисленные компоненты нового вектора. Итерационный процесс запишется так:

$$\begin{cases} x_1^{n+1} = f_1(x_2^n, x_3^n, \dots, x_k^n) \\ x_2^{n+1} = f_2(x_1^{n+1}, x_3^n, \dots, x_k^n) \\ x_3^{n+1} = f_3(x_1^{n+1}, x_2^{n+1}, x_4^n, \dots, x_k^n) \\ \dots \\ x_k^{n+1} = f_k(x_1^{n+1}, x_2^{n+1}, x_3^{n+1}, \dots, x_{k-1}^{n+1}) \end{cases} \quad (5.3.1)$$

Условие сходимости ($\|J\| < 1$) и условие прекращения итераций ($\|\bar{\chi}^{(n)} - \bar{\chi}^{(n-1)}\| < \varepsilon$) те же, что и у метода простых итераций.

Пример. Рассмотрим тот же пример, что и в методе простых итераций.

Решение. Итерации будут выполняться по формулам

$$\begin{cases} x_{n+1} = \sqrt{y_n} \\ y_{n+1} = \sqrt{4 - (x_{n+1})^2} \end{cases}$$

```
n := 0 ; x := 2.0 ; y := 2.0 ; d := 100;
While (n<155) and (d>0.01) do
begin
    n := n+1;
    r := sqrt(y) ; d := sqrt(x-x1); x:=r;
    r := sqrt(4 - x*x) ; d := d + sqrt(y-r); y := r ;
    d := sqrt(d) ;
    Writeln (ff, n:5, x:15:10, y:15:10, d:10:5) ;
end;
```

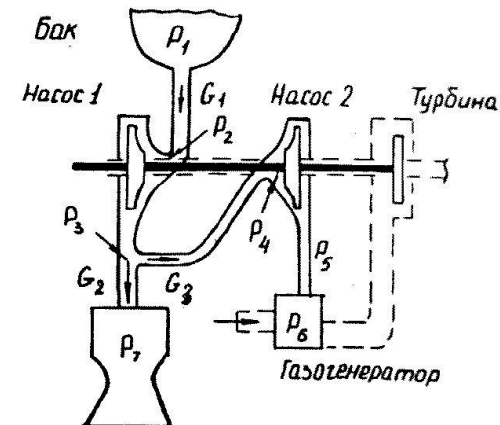
Результаты расчёта до достижения точности $\varepsilon < 0,01$ представлены в таблице:

n	x	y	d
1	1.4142135624	1.4142135624	0.82843

2	1.1892071150	1.6080380710	0.29698
3	1.2680844100	1.5465968864	0.09998
4	1.2436224855	1.5663342918	0.03143
5	1.2515327769	1.5600210602	0.01012
6	1.2490080305	1.5620431940	0.00323

Метод Зейделя в данной задаче сходится быстрее метода простых итераций, но медленнее метода Ньютона (линеаризации).

Пример 2. В качестве второго примера рассмотрим упрощённую схему основной гидравлической подсистемы ракетного двигателя :



Необходимо рассчитать расходы G_i и давление P_i в магистралях при заданных давлении в баке P_1 , коэффициентах гидравлического сопротивления ζ , коэффициентах расхода β . Насосы 1 и 2 находятся на одном валу и вращаются турбиной, магистраль которой в данном случае не рассчитывается. Стационарное движение жидкости, газа в магистралях, насосах, камерах сгорания описывается законами сохранения массы, энергии, их следствиями:

$$G_2 = \beta_1 P_7, \quad (П.1)$$

$$G_3 = \beta_6 P_6, \quad (П.2)$$

$$G_1 = G_2 + G_3, \quad (П.3)$$

$$P_2 = P_1 - \zeta_1 G_1^2, \quad (П.4)$$

$$P_3 = P_2 + a_1 n^2 + b_1 n G_1 + C_1 G_1^2, \quad (П.5)$$

$$P_4 = P_3 - \zeta_3 G_3^2, \quad (П.6)$$

$$P_5 = P_4 + a_2 n^2 + b_2 n G_3 + C_2 G_3^2, \quad (П.7)$$

$$P_6 = P_5 - \zeta_5 G_3^2, \quad (П.8)$$

$$P_7 = P_3 - \zeta_2 G_2^2, \quad (П.9)$$

где n – скорость вращения насоса ($=300$); a, b, c – эмпирические коэффициенты напорной характеристики насоса. Уравнения (П.1) - (П.9)

описывают потери давления в магистральных, напоры насосов 1 и 2, уравнения связи давления и расхода в камере сгорания и в газогенераторе.

Выберем исходные данные: $P_1 = 5$, $\xi_1 = 1,149 \cdot 10^{-5}$, $\xi_2 = 6,85 \cdot 10^{-4}$, $\xi_3 = 0,016$, $\xi_5 = 0,16$, $\beta_6 = 0,083$, $\beta_7 = 2,594$, $a_1 = 1,39 \cdot 10^{-3}$, $b_1 = 6,1 \cdot 10^{-4}$, $c_1 = 3,38 \cdot 10^{-4}$, $a_2 = 9,55 \cdot 10^{-4}$, $b_2 = 0,01$, $c_2 = 0,152$, $n = 300$.

Ниже представлены результаты расчётов по методу Зейделя до относительной погрешности $\varepsilon < 0,01$:

i	P_2	P_3	P_4	P_5	P_6	P_7	G_1	G_2	G_3
1	5.00	130.10	130.10	216.05	216.05	130.10	0.00	0.00	0.00
2	3.55	236.38	231.24	419.86	368.41	158.37	355.41	337.48	17.93
3	2.76	274.48	259.52	579.33	429.73	158.88	441.38	410.81	30.58
4	2.70	277.52	257.17	643.49	439.95	161.17	447.81	412.14	35.67
5	2.63	280.76	259.43	657.60	444.26	161.04	454.59	418.08	36.52
6	2.63	280.77	259.01	662.25	444.71	161.24	454.60	417.73	36.87
7	2.62	281.04	259.24	663.01	445.02	161.21	455.17	418.26	36.91
8	2.62	281.01	259.18	663.33	445.03	161.23	455.11	418.17	36.94
9	2.62	281.04	259.21	663.36	445.06	161.22	455.16	418.23	36.94
10	2.62	281.03	259.20	663.38	445.05	161.23	455.15	418.21	36.94

6. Интерполяция

Пусть в результате не многочисленных, точных и дорогих экспериментов получена табличная функция $\{x_i; y_i, i = 1, \dots, n\}$:

$x =$	x_1	x_2	x_3	\dots	x_n
$y =$	y_1	y_2	y_3	\dots	y_n

Хотелось бы уметь вычислять значения y для любой точки x внутри отрезка $[x_1; x_n]$, не производя дорогостоящих экспериментов.

Одна из самых простых функций – многочлен (полином). Будем искать интерполяционную функцию, которая проходит точно через все узлы в виде многочлена.

Многочлен Лагранжа.

Интерполяционный многочлен Лагранжа имеет вид суммы:

$$L_{n-1}(x) = \sum_{i=1}^n y_i \prod_{\substack{k=1 \\ k \neq i}}^n \frac{(x - x_k)}{(x_i - x_k)} \quad (6.1)$$

Пример. По результатам экспериментов получена таблица

$x =$	0	1	2	3
$y =$	-4	-2	2	14

Получить возможное значение функции y для параметра $x = 1,5$ (внутри таблицы).

Запишем интерполяционный многочлен в форме Лагранжа:

$$L_3(x) = (-4) \frac{1 \cdot (x-1)(x-2)(x-3)}{1 \cdot (0-1)(0-2)(0-3)} + (-2) \frac{(x-0) \cdot 1 \cdot (x-2)(x-3)}{(1-0) \cdot 1 \cdot (1-2)(1-3)} + (2) \frac{(x-0)(x-1) \cdot 1 \cdot (x-3)}{(2-0)(2-1) \cdot 1 \cdot (2-3)} + (14) \frac{(x-0)(x-1)(x-2) \cdot 1}{(3-0)(3-1)(3-2) \cdot 1}.$$

Т.к. нужно получить значение y в одной точке $x=1,5$, не будем раскрывать скобки и приводить подобные члены, а подставим в (6.1) **1,5** вместо x . Получим прогнозируемое значение $y(1,5) = -0,625$.

Многочлен Ньютона.

Пример. По заданной таблице построить интерполяционный многочлен Ньютона и вычислить его значение при $x=0,5$.

x	0	1	2	3
y	-5	-6	3	28

Вычислим и занесём в таблицу разделённые разности 1,2 и 3 порядков.

$$f_1(x_1, x_2) = \frac{y_2 - y_1}{x_2 - x_1} = \frac{(-6) - (-5)}{1 - 0} = -1. \quad f_1(x_2, x_3) = \frac{y_3 - y_2}{x_3 - x_2} = \frac{3 - (-6)}{2 - 1} = 9.$$

$$f_1(x_3, x_4) = \frac{y_4 - y_3}{x_4 - x_3} = \frac{28 - 3}{3 - 2} = 25. \quad f_2(x_1, x_2, x_3) = \frac{f_1(x_2, x_3) - f_1(x_1, x_2)}{x_3 - x_1} = \frac{9 - (-1)}{2 - 0} = 5.$$

$$f_2(x_2, x_3, x_4) = \frac{f_1(x_3, x_4) - f_1(x_2, x_3)}{x_4 - x_2} = \frac{25 - 9}{3 - 1} = 8.$$

$$f_3(x_1, x_2, x_3, x_4) = \frac{f_2(x_2, x_3, x_4) - f_2(x_1, x_2, x_3)}{x_4 - x_1} = \frac{8 - 5}{3 - 0} = 1$$

Выделим верхнюю часть таблицы.

N	x	y	f_1	f_2	f_3
1	0	-5			
			$\frac{(-6) - (-5)}{1 - 0} = \mathbf{-1}$		
2	1	-6		$\frac{9 - (-1)}{2 - 0} = \mathbf{5}$	
			$\frac{3 - (-6)}{2 - 1} = 9$		$\frac{8 - 5}{3 - 0} = \mathbf{1}$
3	2	3		$\frac{25 - 9}{3 - 1} = 8$	
			$\frac{28 - 3}{3 - 2} = 25$		
4	3	28			

Общая формула для n узлов интерполяции:

$$N_{n-1}(x) = f(x_1) + f_1(x_1, x_2) * (x - x_1) + f_2(x_1, x_2, x_3) * (x - x_1) * (x - x_2) + f_3(x_1, x_2, x_3, x_4) * (x - x_1) * (x - x_2) * (x - x_3) + \dots + f_{n-1}(x_1, \dots, x_n) * (x - x_1) * \dots * (x - x_{n-1}). \quad (6.2)$$

Т.е. $N_3(x) = (-5) + (-1) \cdot (x-0) + (5) \cdot (x-0)(x-1) + (1) \cdot (x-0)(x-1)(x-2)$.

Теперь можно вычислить $N_3(0,5) = -6,375$.

Преимущество такого способа заключается в возможности уточнения (6.2) при добавлении нового узла всего лишь прибавлением ещё одного слагаемого: $N_4(x) = N_3(x) + f(x_1, x_2, x_3, x_4, x_5) \cdot (x-0)(x-1)(x-2)(x-3)$.

Канонический многочлен.

Пусть заданы n точек $M_i(x_i; y_i)$ на плоскости XOY . Попытаемся построить многочлен $y=P_{n-1}(x)$ степени $n-1$, который проходил бы точно через все заданные n точек. Так как

$$P_{n-1}(x) = p_1 \cdot x^{n-1} + p_2 \cdot x^{n-2} + p_3 \cdot x^{n-3} + \dots + p_{n-1} \cdot x + p_n, \quad (6.3)$$

то неизвестными являются коэффициенты p_1, p_2, \dots, p_n , которых n штук. Для их нахождения нужно составить n уравнений, для этого в нашем распоряжении координаты n точек $M_i(x_i; y_i)$. Поскольку многочлен $P_{n-1}(x)$ должен проходить через все эти точки, то значение многочлена в каждой точке x_i должно совпадать со значением табличной функции y_i , т.е.

$$P_{n-1}(x_i) = y_i. \quad (6.4)$$

Перепишем уравнение (6.2) для каждой из n точек:

$$\begin{cases} p_1 \cdot x_1^{n-1} + p_2 \cdot x_1^{n-2} + p_3 \cdot x_1^{n-3} + \dots + p_{n-1} \cdot x_1 + p_n = y_1 \\ p_1 \cdot x_2^{n-1} + p_2 \cdot x_2^{n-2} + p_3 \cdot x_2^{n-3} + \dots + p_{n-1} \cdot x_2 + p_n = y_2 \\ \dots\dots\dots \\ p_1 \cdot x_n^{n-1} + p_2 \cdot x_n^{n-2} + p_3 \cdot x_n^{n-3} + \dots + p_{n-1} \cdot x_n + p_n = y_n \end{cases} \quad (6.5)$$

Получилась ничто иное, как линейная система, состоящая из n уравнений с n неизвестными: $p_1, p_2, p_3, \dots, p_n$. Решается она, например, методом Гаусса, и после вычисления неизвестных $p_1, p_2, p_3, \dots, p_n$, можно записать (6.1), подставив туда вместо символов $p_1, p_2, p_3, \dots, p_n$ только что вычисленные их числовые значения. Интерполяционный многочлен $P_{n-1}(x)$ составлен. Его можно использовать для вычисления значений функции (и её производных) между узлами интерполяции. По способу его нахождения его называют *каноническим*. Известны и другие методы вычисления интерполяционного многочлена (многочлены Ньютона, Лагранжа, Чебышева...). Построенные по одинаковой таблице разными способами интерполяционные многочлены совпадают.

Пример. Вычислить коэффициенты многочлена второй степени (параболы), проходящего через три заданные точки:

$x=$	-1	1	5
$y=$	4	-2	10

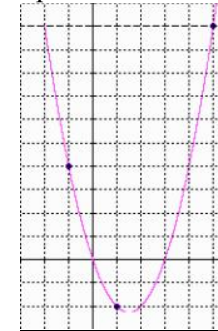
Будем искать функцию $P_2(x) = a \cdot x^2 + b \cdot x + c$. Так как $P_2(x_i) = y_i$, то запишем три уравнения системы (6.3):

$$\begin{cases} a \cdot (-1)^2 + b \cdot (-1) + c = 4 \\ a \cdot (1)^2 + b \cdot (1) + c = -2 \\ a \cdot (5)^2 + b \cdot (5) + c = 10 \end{cases} \quad \text{или} \quad \begin{cases} a - b + c = 4 \\ a + b + c = -2 \\ 25 \cdot a + 5 \cdot b + c = 10 \end{cases}$$

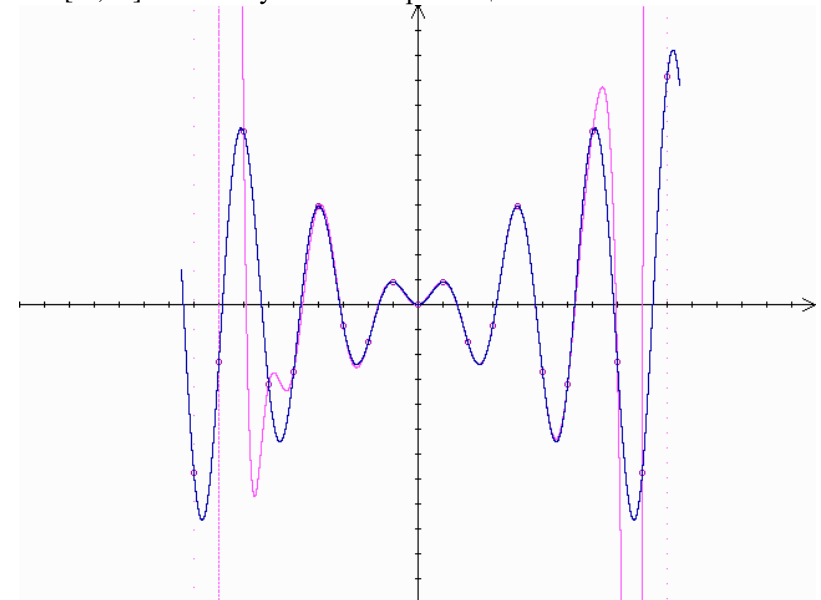
Решим эту систему и получим $a=1, b=-3, c=0$, т.е.

$$P_2(x) = x^2 - 3x.$$

Нарисуем на одних осях координат точки таблицы и график $P_2(x)$:



Замечание 1. Интерполяцию применяют для точек внутри отрезка интерполяции (от x_1 до x_n) для табличной функции, содержащей не более пяти-восьми узлов. При большом количестве узлов интерполяции на краях отрезка (дальше от начала координат) между узлами обычно сильно возрастает ошибка, т.е. колебания многочлена сильно возрастают и ближе к границе аппроксимация (приближение) табличной функции становится неудовлетворительной. В качестве иллюстрации сказанного предлагаем читателю оценить качество аппроксимации функции $y = x \cdot \sin(2 \cdot x)$ на отрезке $[-9; 10]$ с 20-тью узлами интерполяции:



Замечание 2. Не применяют интерполяцию и в случае, когда узлы интерполяции (значения табличной функции) заданы не точно. Чем хуже точность задания табличной функции, тем хуже интерполяционный многочлен будет аппроксимировать таблицу x - y .

Замечание 3. Если табличная функция, которую требуется аппроксимировать, имеет короткие участки монотонности (частые ситуации, когда $y_{i-1} < y_i > y_{i+1}$ или $y_{i-1} > y_i < y_{i+1}$), то на хорошую аппроксимацию с помощью интерполяционного многочлена вряд ли стоит надеяться.

Замечание 4. Улучшить интерполяционный многочлен можно, построив его по специально вычисленным узлам. Узлы интерполяции должны совпадать с корнями так называемых *полиномов Чебышева*, тогда ошибки интерполяции будут минимальны (но они будут).

Замечание 5. Если нужно вычислить значение функции y для одной или нескольких точек x , то наиболее эффективны интерполяционные многочлены Лагранжа и Ньютона. Если нужно вычислять производные от y по x (искать экстремум) или построить график $y(x)$, то канонический интерполяционный многочлен будет удобнее.

7. Аппроксимация методом наименьших квадратов

Рассмотрим табличную функцию, значения которой заданы не точно. Например, таблицу заполняли замером некоторого параметра при проведении каких-то схожих экспериментов. А измерения происходят при внешних воздействиях (изменяющийся ветер, разная температура,...). Проведённые эксперименты не дорогие, не точные и их много. Использование интерполяции невозможно. Но если в силу, например, некоторых теоретических исследований мы подозреваем, что исследуемый процесс должен иметь, ну скажем, квадратичный характер (исследуемая зависимость представляет собой параболу), то можно среди всех существующих парабол подобрать ту, которая лучше всех остальных аппроксимировала бы табличную функцию.

7.1. Квадратичная аппроксимация

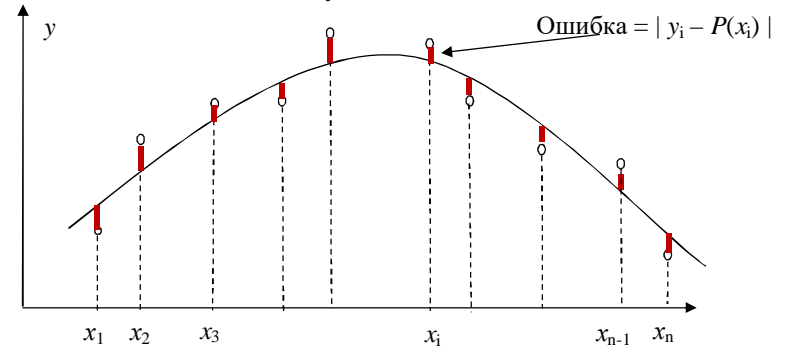
Пусть имеется некоторая табличная функция с n узлами:

x_1	x_2	x_3	\dots	x_{n-1}	x_n
y_1	y_2	y_3	\dots	y_{n-1}	y_n

(7.1-1)

Представляющая собой n точек на плоскости xOy . Нарисуем параболу, которая должна хорошо аппроксимировать таблицу. Но параболу $P(x) = a \cdot x^2 + b \cdot x + c$ вряд ли пройдёт через все узлы. Т.е. для каждого $x = x_i$ значения табличной функции y_i и значения многочлена $P(x_i)$ будут различаться на величину ошибки $d_i = y_i - P(x_i)$. Чтобы при суммировании

отрицательные d_i не сократились с положительными d_i возведём в квадрат все ошибки d_i и сложим их для всех узлов x_i .



Получим выражение

$$\sum_{i=1}^n (y_i - P(x_i))^2 = \sum_{i=1}^n (y_i - a \cdot x_i^2 - b \cdot x_i - c)^2 = \psi(a, b, c) \quad (7.1.2)$$

Для разных парабол (для разных значений a, b, c) значение функционала невязки $\psi(a, b, c)$ будет разным. Попробуем найти такие три числа a, b, c , для которых значение невязки $\psi(a, b, c)$ было бы минимально.

Поскольку ищется минимум функции трёх переменных, все частные производные должны приравняться нулю:

$$\begin{cases} \frac{\partial \psi}{\partial a} = 0 \\ \frac{\partial \psi}{\partial b} = 0 \\ \frac{\partial \psi}{\partial c} = 0 \end{cases} \Leftrightarrow \begin{cases} \sum_{i=1}^n 2(y_i - a \cdot x_i^2 - b x_i - c)(x_i^2) = 0 \\ \sum_{i=1}^n 2(y_i - a \cdot x_i^2 - b x_i - c)(x_i) = 0 \\ \sum_{i=1}^n 2(y_i - a \cdot x_i^2 - b x_i - c) = 0 \end{cases} \quad (7.1.3)$$

После преобразования получаем систему из трёх линейных уравнений с тремя неизвестными a, b, c :

$$\begin{cases} (\sum_{i=1}^n x_i^4) \cdot a + (\sum_{i=1}^n x_i^3) \cdot b + (\sum_{i=1}^n x_i^2) \cdot c = \sum_{i=1}^n (x_i^2 y_i) \\ (\sum_{i=1}^n x_i^3) \cdot a + (\sum_{i=1}^n x_i^2) \cdot b + (\sum_{i=1}^n x_i) \cdot c = \sum_{i=1}^n (x_i y_i) \\ (\sum_{i=1}^n x_i^2) \cdot a + (\sum_{i=1}^n x_i) \cdot b + n \cdot c = \sum_{i=1}^n y_i \end{cases} \quad (7.1.4)$$

Решив эту систему (например, методом Гаусса), найдём числовые значения a, b, c , те самые, для которых квадратическая невязка (7.1.2) будет принимать минимальное значение. Полученный таким образом многочлен $P(x) = a \cdot x^2 + b \cdot x + c$ является для табличной функции (7.1.1) оптимальным в смысле минимизации невязки (7.1.2) среди всех многочленов второй степени.

Пример 1. Вычислить коэффициенты параболы (и построить её), аппроксимирующую таблицу:

$x =$	-2	-1	0	1	2
$y =$	3	4	2	1	1

Вычислим коэффициенты системы (7.1.4) :

$$n=5, \sum_{i=1}^5 x_i = -2-1+0+1+2 = 0; \quad \sum_{i=1}^5 x_i^2 = (-2)^2 + (-1)^2 + (0)^2 + (1)^2 + (2)^2 = 10;$$

$$\sum_{i=1}^5 x_i^3 = (-2)^3 + (-1)^3 + (0)^3 + (1)^3 + (2)^3 = 0;$$

$$\sum_{i=1}^5 x_i^4 = (-2)^4 + (-1)^4 + (0)^4 + (1)^4 + (2)^4 = 34;$$

$$\sum_{i=1}^5 (y_i) = 3 + 4 + 2 + 1 + 1 = 11;$$

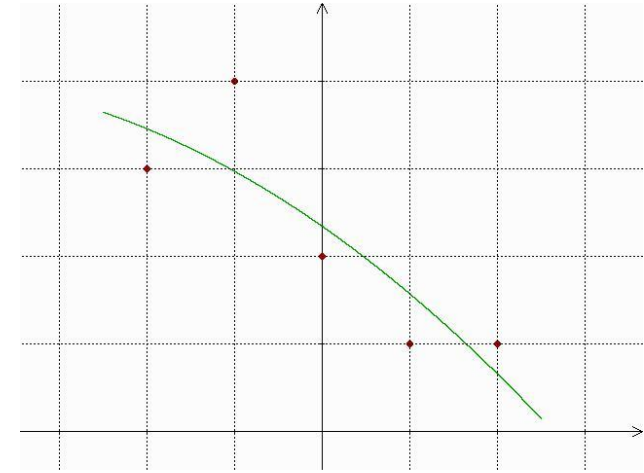
$$\sum_{i=1}^5 (x_i \cdot y_i) = (-2) \cdot 3 + (-1) \cdot 4 + 0 \cdot 2 + 1 \cdot 1 + 2 \cdot 1 = -7;$$

$$\sum_{i=1}^5 (x_i^2 \cdot y_i) = (-2)^2 \cdot 3 + (-1)^2 \cdot 4 + 0^2 \cdot 2 + 1^2 \cdot 1 + 2^2 \cdot 1 = 21.$$

Теперь система (7.1-4) примет вид

$$\begin{cases} 34 \cdot a + 0 \cdot b + 10 \cdot c = 21 \\ 0 \cdot a + 10 \cdot b + 0 \cdot c = -7 \\ 10 \cdot a + 0 \cdot b + 5 \cdot c = 11 \end{cases}.$$

Решив эту систему, получим $a = -0,0714285$, $b = -0,7$, $c = 2,3428571$, т.е. $P(x) = a \cdot x^2 + b \cdot x + c = -0,0714285 \cdot x^2 - 0,7 \cdot x + 2,342857$. Построим этот многочлен и точки таблицы:



Квадратическая невязка $\psi(a, b, c) = (P(-2) - y_1)^2 + (P(-1) - y_2)^2 + \dots = (-0,0714285 \cdot (-2)^2 - 0,7 \cdot (-2) + 2,342857 - 3)^2 + (-0,0714285 \cdot (-1)^2 - 0,7 \cdot (-1) + 2,342857 - 4)^2 + \dots + (-0,0714285 \cdot (2)^2 - 0,7 \cdot (2) + 2,342857 - 1)^2 \approx 1,83$. Для любой другой параболы с коэффициентами a^* , b^* , c^* квадратическая невязка $\psi(a^*, b^*, c^*)$ будет больше.

Пример 2. Вычислить коэффициенты параболы (и построить её), аппроксимирующую таблицу:

X	-4	-3	-2	-1	0	1	2	3	4	5
Y	2,4	3,7	4	3,9	4,8	4,7	4	3,1	2,4	1,7

Вычислим коэффициенты системы (7.1.4) :

$$\sum_{i=1}^{10} x_i = -4-3-2-1+0+1+2+3+4+5=5;$$

$$\sum_{i=1}^{10} x_i^2 = (-4)^2 + (-3)^2 + (-2)^2 + (-1)^2 + (0)^2 + (1)^2 + (2)^2 + (3)^2 + (4)^2 + (5)^2 = 85;$$

$$\sum_{i=1}^{10} x_i^3 = (-4)^3 + (-3)^3 + (-2)^3 + (-1)^3 + (0)^3 + (1)^3 + (2)^3 + (3)^3 + (4)^3 + (5)^3 = 125;$$

$$\sum_{i=1}^{10} x_i^4 = (-4)^4 + (-3)^4 + (-2)^4 + (-1)^4 + (0)^4 + (1)^4 + (2)^4 + (3)^4 + (4)^4 + (5)^4 = 1333;$$

$$\sum_{i=1}^{10} (y_i) = 2,4 + 3,7 + 4,0 + 3,9 + 4,8 + 4,7 + 4,0 + 3,1 + 2,4 + 1,7 = 34,7;$$

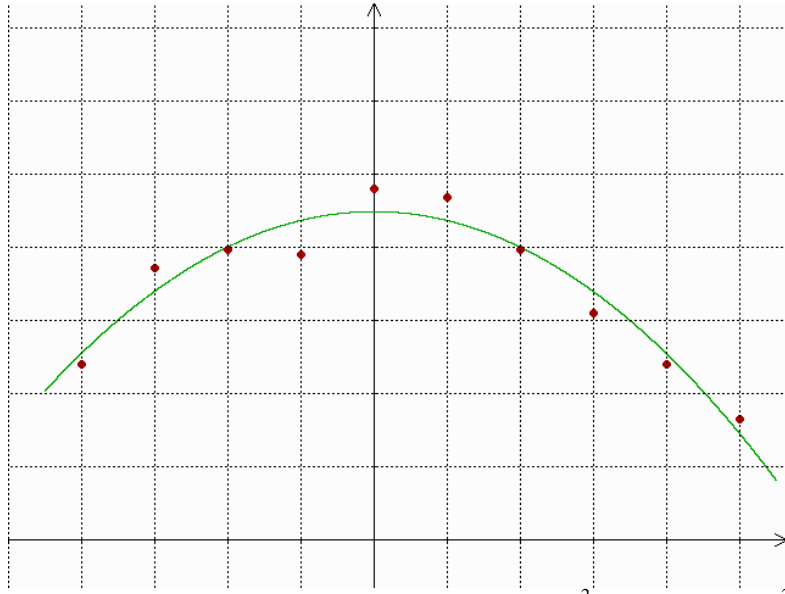
$$\sum_{i=1}^{10} (x_i \cdot y_i) = (-4) \cdot 2,4 + (-3) \cdot 3,7 + \dots + 4 \cdot 2,4 + 5 \cdot 1,7 = 7,5;$$

$$\sum_{i=1}^{10} (x_i^2 \cdot y_i) = (-4)^2 \cdot 2,4 + (-3)^2 \cdot 3,7 + \dots + 4^2 \cdot 2,4 + 5^2 \cdot 1,7 = 221,1.$$

После подстановки коэффициентов система примет вид

$$\begin{cases} 1333 \cdot a + 125 \cdot b + 85 \cdot c = 221,1 \\ 125 \cdot a + 85 \cdot b + 5 \cdot c = 7,5 \\ 85 \cdot a + 5 \cdot b + 10 \cdot c = 34,7 \end{cases}$$

Решив эту систему, получим $a = -0,12$, $b = 0,0018$, $c = 4,4993$, т.е. $P(x) = a \cdot x^2 + b \cdot x + c = -0,12121212 \cdot x^2 + 0,00181818 \cdot x + 4,49939393$. Построим этот многочлен и точки таблицы:



Квадратическая невязка $\psi(a, b, c) = (P(-4) - y_1)^2 + (P(-3) - y_2)^2 + \dots = (-0,1212 \cdot (-4)^2 - 0,0018 \cdot (-4) + 4,499 - 2,4)^2 + (-0,1212 \cdot (-3)^2 - 0,0018 \cdot (-3) + 4,499 - 3,7)^2 + \dots + (-0,1212 \cdot (5)^2 - 0,0018 \cdot (5) + 4,499 - 1,7)^2 \approx 0,707$. Для любой другой параболы с коэффициентами a^* , b^* , c^* квадратическая невязка $\psi(a^*, b^*, c^*)$ будет больше.

7.2. Линейная аппроксимация

Аппроксимировать табличную функцию (7.1.1) линейной функцией $P(x) = b \cdot x + c$ проще. В функции $P(x)$ отсутствует коэффициент при x^2 , т.е. $a = 0$, и невязка $\psi(b, c)$ зависит только от b и c :

$$\sum_{i=1}^n (y_i - P(x_i))^2 = \sum_{i=1}^n (y_i - b \cdot x_i - c)^2 = \psi(b, c). \quad (7.2.1)$$

В системе (7.1.3) и (7.1.4) будут отсутствовать первый столбец и первая строка, и она примет вид

$$\begin{cases} (\sum_{i=1}^n x_i^2) \cdot b + (\sum_{i=1}^n x_i) \cdot c = \sum_{i=1}^n (x_i \cdot y_i) \\ (\sum_{i=1}^n x_i) \cdot b + n \cdot c = \sum_{i=1}^n y_i \end{cases} \quad (7.2.2)$$

Решив эту систему, найдём числовые значения b и c , те самые, для которых квадратическая невязка (7.2.1) будет принимать минимальное значение. Полученный таким образом многочлен $P(x) = b \cdot x + c$ является для табличной функции (7.1.1) оптимальным в смысле минимизации невязки (7.2.2) среди всех линейных многочленов.

Пример. Вычислить коэффициенты прямой (и построить её), аппроксимирующую таблицу:

x	1	2	3	4	5	6	7	8	9	10
y	7	7	6	7	6	5	4	4	4	3

Вычислим коэффициенты системы (7.1-4):

$$\sum_{i=1}^{10} x_i = 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 = 55;$$

$$\sum_{i=1}^{10} x_i^2 = (1)^2 + (2)^2 + (3)^2 + (4)^2 + (5)^2 + (6)^2 + (7)^2 + (8)^2 + (9)^2 + (10)^2 = 385;$$

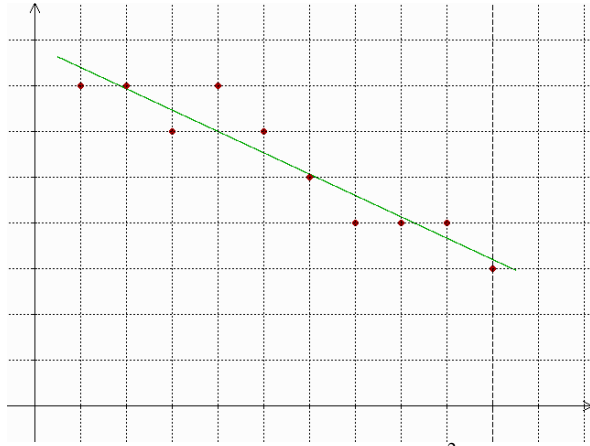
$$\sum_{i=1}^{10} y_i = 7 + 7 + 6 + 7 + 6 + 5 + 4 + 4 + 4 + 3 = 53;$$

$$\sum_{i=1}^{10} (x_i \cdot y_i) = 1 \cdot 7 + 2 \cdot 7 + 3 \cdot 6 + 4 \cdot 7 + 5 \cdot 6 + 6 \cdot 5 + 7 \cdot 4 + 8 \cdot 4 + 9 \cdot 4 + 10 \cdot 3 = 253.$$

После подстановки коэффициентов система примет вид

$$\begin{cases} 385 \cdot b + 55 \cdot c = 253 \\ 55 \cdot b + 10 \cdot c = 53 \end{cases}.$$

Решив эту систему, получим $b = -0,4(6)$, $c = 7,8(6)$, т.е. $P(x) = b \cdot x + c = -0,466666666 \cdot x + 7,866666666$. Построим этот многочлен и точки таблицы:



Квадратическая невязка $\psi(b, c) = (P(1)-y_1)^2 + (P(2)-y_2)^2 + \dots = (-0,467 \cdot (1) + 7,867 - 7)^2 + (-0,467 \cdot (2) + 7,867 - 7)^2 + \dots + (-0,467 \cdot (10) + 7,867 - 3)^2 \approx 2,13$. Для любой другой прямой с коэффициентами b^*, c^* невязка $\psi(b^*, c^*)$ будет больше.

Замечание. Чтобы выбрать степень аппроксимирующего многочлена $P(x)$ можно сначала построить линейный $P_1(x)=ax+b$ и вычислить для него $\psi_1(a, b)$. Если для заданного $\varepsilon > 0$ окажется что $\psi_1 < \varepsilon$, то аппроксимирующем многочленом можно считать $P_1(x)$, если нет, то следует проверить $\psi_2(a, b, c)$ для квадратичного $P_2(x)=ax^2+bx+c$ или ψ_3 для кубического $P_3(x)$.

8. Интерполяция сплайнами

8.1. c-spline

Рассмотрим ситуацию, когда табличная функция задана точно и значений в этой таблице много. В этом случае полезнее всего использовать аппроксимацию сплайнами. Строится непрерывная функция $S(x)$, проходящая (как и интерполяционный многочлен) точно через все узлы, которая не только гладкая сама, но существует и является гладкой первая производная от неё. Её называют кубическим сплайном дефекта 1 (c-spline).

Пусть имеется таблица из n точек $(x_i; y_i)$:

x_1	x_2	x_3	\dots	x_{n-1}	x_n
y_1	y_2	y_3	\dots	y_{n-1}	y_n

(8.1)

На каждом отрезке $[x_{i-1}; x_i]$ будем искать (вычислять его коэффициенты) свой кубический многочлен:

$$S_i(x) = a_i \cdot x^3 + b_i \cdot x^2 + c_i \cdot x + d_i, \quad (8.2)$$

$$\text{Причём } S_i(x_i) = y_i, \quad S_i(x_{i-1}) = y_{i-1} \quad \text{для } i=2, 3, 4, \dots, n. \quad (8.3)$$

Для внутренних узлов (для $i=2, 3, \dots, n-1$)

$$S''_i(x_i-0) = S''_{i+1}(x_i+0), \quad S'_i(x_i-0) = S'_{i+1}(x_i+0). \quad (8.4)$$

Для двух граничных узлов ($i=1$ и $i=n$)

$$S''_2(x_1) = S''_n(x_n) = 0. \quad (8.5)$$

Пусть известны значения вторых производных сплайна во внутренних точках ($S''_i(x_i) = m_i$). Т.к. сплайн $S_i(x)$ на отрезке $[x_{i-1}; x_i]$ длиной $h_i = x_i - x_{i-1}$ является кубическим многочленом, то $S''_i(x)$ - линейный многочлен

$$S''_i(x) = m_i \frac{x - x_{i-1}}{x_i - x_{i-1}} + m_{i-1} \frac{x - x_i}{x_{i-1} - x_i}$$

Дважды проинтегрируем его и, учитывая (8.3), получим

$$S_i(x) = m_i \frac{(x - x_{i-1})^3}{6h_i} + m_{i-1} \frac{(x_i - x)^3}{6h_i} + \left(y_i - m_i \frac{h_i^2}{6} \right) \frac{x - x_{i-1}}{h_i} + \left(y_{i-1} - m_{i-1} \frac{h_i^2}{6} \right) \frac{x_i - x}{h_i}. \quad (8.6)$$

Значения вторых производных сплайна во внутренних точках $m_i = S''_i(x_i-0) = S''_{i+1}(x_i+0)$ найдём, используя условие $S'_i(x_i-0) = S'_{i+1}(x_i+0)$ для составления и решения СЛАУ с трёхдиагональной матрицей

$$\begin{cases} \frac{h_i}{6} m_{i-1} + \frac{h_i + h_{i+1}}{3} m_i + \frac{h_{i+1}}{6} m_{i+1} = \frac{y_{i+1} - y_i}{h_{i+1}} - \frac{y_i - y_{i-1}}{h_i}, & m_1 = m_n = 0 \end{cases} \quad (8.7)$$

$$i = 2, 3, 4, \dots, n-1.$$

Обозначив правые части в (8.7) за Y_i :

$$Y_i = \frac{y_{i+1} - y_i}{h_{i+1}} - \frac{y_i - y_{i-1}}{h_i},$$

воспользуемся формулами метода прогонки. Вычислим прогоночные коэффициенты P_i и Q_i :

$P_2 = Q_2 = 0$, для $i=2, 3, 4, \dots, n-1$:

$$P_{i+1} = -\frac{h_{i+1}}{2 \cdot (h_i + h_{i+1}) + P_i \cdot h_i}; \quad Q_{i+1} = \frac{6 \cdot Y_i - h_i \cdot Q_i}{2 \cdot (h_i + h_{i+1}) + P_i \cdot h_i} \quad (8.8)$$

Затем вычисляются все m_i . $m_n = 0$. Для $i=n, n-1, \dots, 2$:

$$m_{i-1} = P_i \cdot m_i + Q_i. \quad (8.9)$$

Итак, на каждом отрезке $[x_{i-1}; x_i]$ табличной функции x - y мы имеем свою кубическую параболу (8.6). Если нужно вычислить $y(x_0)$, то нужно найти тот отрезок $[x_{i-1}; x_i]$, который содержит x_0 и вычислить $S_i(x_0)$. Если нужна производная $y'(x_0)$, то вычисляют $S'_i(x_0)$.

Пример 1. Для заданной таблицы с помощью кубических сплайнов дефекта 1 вычислить $y(2)$.

$i=$	1	2	3	4	5
$x=$	0	1	3	6	10
$y=$	2	4	6	7	3
$h=$	-	1	2	3	4

Решение.

Вычислим длины отрезков:

$$h_2 = x_2 - x_1 = 1 - 0 = 1; \quad h_3 = x_3 - x_2 = 3 - 1 = 2; \quad h_4 = x_4 - x_3 = 6 - 3 = 3; \quad h_5 = x_5 - x_4 = 10 - 6 = 4;$$

Тогда системе (8.7) примет вид

$$\begin{cases} \frac{h_2}{6} m_1 + \frac{h_2 + h_3}{3} m_2 + \frac{h_3}{6} m_3 &= \frac{y_3 - y_2}{h_3} - \frac{y_2 - y_1}{h_2} \\ \frac{h_3}{6} m_2 + \frac{h_3 + h_4}{3} m_3 + \frac{h_4}{6} m_4 &= \frac{y_4 - y_3}{h_4} - \frac{y_3 - y_2}{h_3} \\ \frac{h_4}{6} m_3 + \frac{h_4 + h_5}{3} m_4 + \frac{h_5}{6} m_5 &= \frac{y_5 - y_4}{h_5} - \frac{y_4 - y_3}{h_4} \end{cases}$$

Вычислим правые части Y_i ($i=2,3,4$) уравнений (8.7) - для внутренних точек отрезка:

$$Y_2 = (y_3 - y_2)/h_3 - (y_2 - y_1)/h_2 = (6-4)/2 - (4-2)/1 = -1;$$

$$Y_3 = (y_4 - y_3)/h_4 - (y_3 - y_2)/h_3 = (7-6)/3 - (6-4)/2 = -2/3;$$

$$Y_4 = (y_5 - y_4)/h_5 - (y_4 - y_3)/h_4 = (3-7)/4 - (7-6)/3 = -4/3;$$

Выпишем систему (8,7) с учётом вычисленных правых частей Y_i и того, что $m_1 = m_5 = 0$:

$$\begin{cases} 1 \cdot m_2 + \frac{1}{3} \cdot m_3 &= -1 \\ \frac{1}{3} \cdot m_2 + \frac{5}{3} \cdot m_3 + \frac{1}{2} \cdot m_4 &= -\frac{2}{3} \\ \frac{1}{2} \cdot m_3 + \frac{7}{3} \cdot m_4 &= -\frac{4}{3} \end{cases}$$

Теперь по формулам (8.8) вычислим прогоночные коэффициенты P_i и

$$Q_i: \quad P_2 = 0, \quad Q_2 = 0, \quad P_3 = -h_3 / (2(h_2 + h_3) + P_2 \cdot h_2) = -2 / (2(1+2)+0) = -1/3,$$

$$Q_3 = (6 \cdot Y_2 - h_2 \cdot Q_2) / (2(h_2 + h_3) + P_2 \cdot h_2) = 6(-1) / (2(1+2)+0) = -1,$$

$$P_4 = -h_4 / (2(h_3 + h_4) + P_3 \cdot h_3) = -3 / (2(2+3) + (-1/3) \cdot 2) = -9/28,$$

$$Q_4 = (6 \cdot Y_3 - h_3 \cdot Q_3) / (2(h_3 + h_4) + P_3 \cdot h_3) = (6(-2/3) - 2 \cdot (-1)) / (2(2+3) + (-1/3) \cdot 2) = -3/14,$$

$$P_5 = -h_5 / (2(h_4 + h_5) + P_4 \cdot h_4) = -4 / (2(3+4) + (-9/28) \cdot 3) = -112/365,$$

$$Q_5 = (6 \cdot Y_4 - h_4 \cdot Q_4) / (2(h_4 + h_5) + P_4 \cdot h_4) = (6(-4/3) - 3 \cdot (-3/14)) / (2(3+4) + (-9/28) \cdot 3) = -206/365,$$

теперь по формулам (8.9), помня, что $m_1 = m_5 = 0$, вычисляем

$$m_4 = Q_5 = -206/365 = -0.564384,$$

$$m_3 = P_4 \cdot m_4 + Q_4 = (-206/365) \cdot (-9/28) - 3/14 = -12/365 = -0.032877,$$

$$m_2 = P_3 \cdot m_3 + Q_3 = (-1/3) \cdot (-12/365) - 1 = -361/365 = -0.989041.$$

Искомое значение y ищется в точке $x^* = 2$, которое лежит во втором отрезке ($i=3$). По формуле (8.6) вычислим

$$S_3(2) = m_3 \frac{(2-x_2)^3}{6h_3} + m_2 \frac{(x_3-2)^3}{6h_3} + \left(y_3 - m_3 \frac{h_3^2}{6} \right) \frac{2-x_2}{h_3} + \left(y_2 - m_2 \frac{h_3^2}{6} \right) \frac{x_3-2}{h_3} = 5.0109589$$

Если есть необходимость вычисления нескольких значений y на этом отрезке, то, зная $m_2, m_3, y_2, y_3, x_2, x_3$, можно восстановить кубический многочлен из (8.6):

$$S_3(x) = -0.0295282 \cdot x^3 + 0.2493151 \cdot x^2 - 0.0502283 \cdot x + 4.7041096.$$

Замечание. Если к таблице добавить (или удалить) узлы, то изменится решение системы (8.6), а значит, и значение $S_i(x^*)$.

Пример 2. Рассмотрим табличную функцию из девяти точек: x меняется от -4 до 4 с шагом 1, а y вычислим с помощью образующей функции $F(x) = x \cdot \sin(x)$. Таблица будет выглядеть так:

$i=$	1	2	3	4	5	6	7	8	9
$x=$	-4	-3	-2	-1	0	1	2	3	4
$y=$	-3.02721	0.42336	1.81859	0.84147	0	0.84147	1.81859	0.42336	-3.02721

Первым делом вычислим все прогоночные коэффициенты по формулам (8.8):

$i=$	2	3	4	5	6	7	8	9
$P=$	0	-0.25	-0.2(6)	-0.2678	-0.2679	-0.2679	-0.2679	-0.2679
$Q=$	0	-3.083	-2.9736	1.01452	2.4338	-0.43403	-3.69773	-2.3135

Теперь по формулам (8.9) вычислим все m :

$i=$	9	8	7	6	5	4	3	2	1
$m=$	0	-2.3135	-3.0778	0.39066	2.32908	0.39066	-3.0778	-2.3135	0

Осталось на каждом отрезке $[x_{i-1}; x_i]$, воспользовавшись формулами (8.6), вычислить коэффициенты (и привести подобные члены) кубических многочленов $S_i(x)$:

для $-4 \leq x \leq -3$

$$S_1(x) = -0.385591 \cdot x^3 - 4.627097 \cdot x^2 - 14.672227 \cdot x - 12.360416;$$

для $-3 \leq x \leq -2$

$$S_2(x) = -0.127378 \cdot x^3 - 2.303176 \cdot x^2 - 7.700465 \cdot x - 5.388654;$$

для $-2 \leq x \leq -1$

$$S_3(x) = 0.578080 \cdot x^3 + 1.929571 \cdot x^2 + 0.765031 \cdot x + 0.255010;$$

для $-1 \leq x \leq 0$

$$S_4(x) = 0.323070 \cdot x^3 + 1.164541 \cdot x^2;$$

для $0 \leq x \leq 1$

$$S_5(x) = -0.323070 \cdot x^3 + 1.164541 \cdot x^2;$$

для $1 \leq x \leq 2$

$$S_6(x) = -0.578080 \cdot x^3 + 1.929571 \cdot x^2 - 0.765031 \cdot x + 0.255010;$$

для $2 \leq x \leq 3$

$$S_7(x) = 0.127378 \cdot x^3 - 2.303176 \cdot x^2 + 7.700465 \cdot x - 5.388654;$$

для $3 \leq x \leq 4$

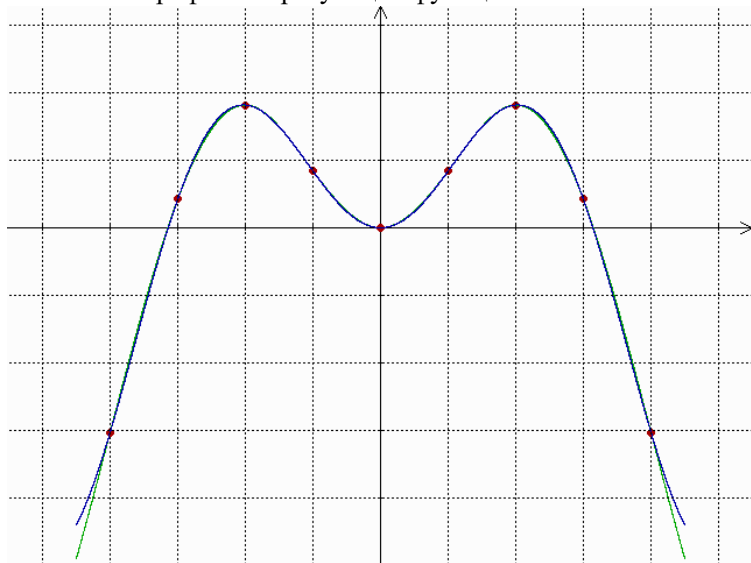
$$S_8(x) = 0.385591 \cdot x^3 - 4.627097 \cdot x^2 + 14.672227 \cdot x - 12.360416.$$

Расчёт показывает, что максимальное значение отклонения сплайна от образующей функции менее 0,076:

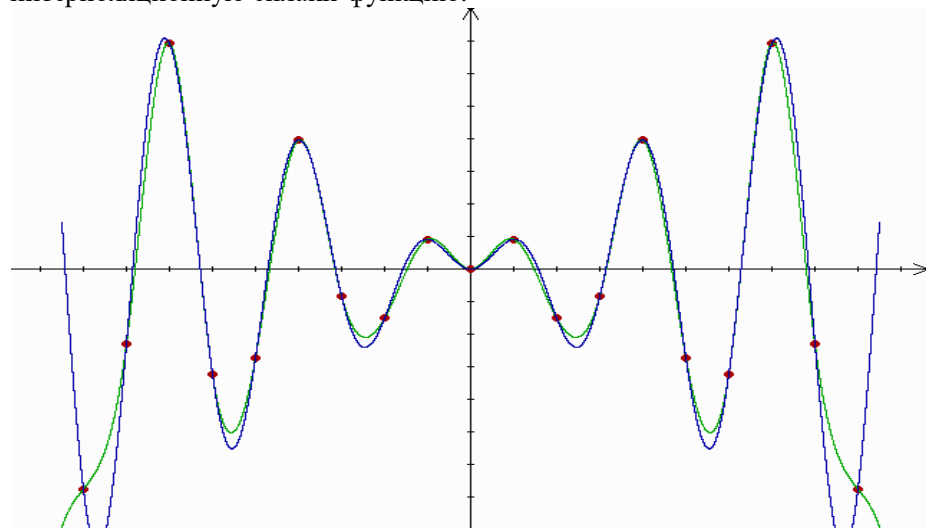
$$\max |F(x) - S(x)| < 0,076, \quad \text{для } x \in [-4; 4].$$

Для уменьшения этого отклонения необходимо уменьшать шаг изменения x в таблице, увеличивая количество узлов интерполяции. На

рисунке показаны графики образующей функции и сплайна:



Пример 3. В качестве примера возьмём уже использованную при интерполяции функцию $y = x \cdot \sin(2 \cdot x)$. С помощью этой образующей функции будем вычислять значения табличных y для табличных x . В качестве значений табличных x возьмём целые числа от -9 до 10 включительно. По этим двадцати точкам на плоскости XOY будем строить интерполяционную сплайн-функцию:



Из этого графика видно, что аппроксимация неудовлетворительна

лишь на двух крайних отрезках. Внутри отрезка $[x_2; x_{n-1}]$ аппроксимация точнее. Аппроксимация интерполяционным многочленом (см. пример разд. 6) в середине отрезка выглядит гораздо лучше. Но интерполяционный многочлен мы существенно улучшить не можем (при увеличении количества точек интерполяции увеличивается степень многочлена, а значит и разболтка). А аппроксимацию сплайнами можно улучшать, увеличивая количество узлов (уменьшая шаг сетки) на сколько это требуется.

8.2. A-spline

Рассмотрим такую же ситуацию, когда табличная функция задана точно и значений в этой таблице много. Строится непрерывная функция $S(x)$, проходящая точно через все узлы, которая гладкая сама, но первая производная от неё может не быть гладкой. Её называют кубическим сплайном дефекта 2.

Пусть имеется таблица из n точек $(x_i; y_i)$:

x_1	x_2	x_3	\dots	x_{n-1}	x_n
y_1	y_2	y_3	\dots	y_{n-1}	y_n

(8.10)

На каждом отрезке $[x_{i-1}; x_i]$ будем искать (вычислять его коэффициенты) свой кубический многочлен:

$$S_i(x) = a_i \cdot x^3 + b_i \cdot x^2 + c_i \cdot x + d_i, \quad (8.11)$$

$$\text{Причём } S_i(x_i) = y_i, \quad S_i(x_{i-1}) = y_{i-1} \quad \text{для } i=2,3,4,\dots,n. \quad (8.12)$$

$$\text{Для внутренних узлов (для } i=2,3,\dots,n-1) \quad S'_i(x_i-0) = S'_{i+1}(x_i+0). \quad (8.13)$$

Вычислим разделённые разности первого порядка:

$$d_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \quad \text{для } i = 1, \dots, n-1$$

9. Численное дифференцирование

Пусть задана табличная функция $\{x_i; y_i, i=1, \dots, n\}$:

$x=$	x_1	x_2	x_3	\dots	x_n
$y=$	y_1	y_2	y_3	\dots	y_n

Т.е. заданы n точек $M_i(x_i; y_i)$ на плоскости XOY , характеризующие значения некоторой неизвестной функции $f(x)$. Требуется найти $f'(x_0)$, $f''(x_0), \dots$ для $x_0 \in [x_1; x_n]$. Для этого построим интерполяционный многочлен $y=P(x)$ и от него вычислим первую производную, вторую производную и т.д. в точке $x=x_0$. Вычисление канонического интерполяционного многочлена легко программируется, ещё проще программируется вычисление производной от многочлена с известными коэффициентами. Часто строят многочлен не по всем узлам, а только по тем, что ближе к x_0 .

Встречаются ситуации, когда необходимо уметь вычислять производные по нескольким соседним точкам. В этих случаях пригодятся следующие формулы, полученные с помощью разложения в ряд Тейлора или дифференцирования интерполяционного многочлена Лагранжа:

$$y'(x) = \frac{y(x+h) - y(x)}{h} + O(h), \quad y'(x) = \frac{y(x) - y(x-h)}{h} + O(h),$$

$$y'(x) = \frac{y(x+h) - y(x-h)}{2h} + O(h^2),$$

$$y'(x) = \frac{-3y(x) + 4y(x+h) - y(x+2h)}{2h} + O(h^2),$$

$$y'(x) = \frac{y(x-2h) - 4y(x-h) + 3y(x)}{2h} + O(h^2),$$

$$y'(x) = \frac{-2y(x-h) - 3y(x) + 6y(x+h) - y(x+2h)}{6h} + O(h^3)$$

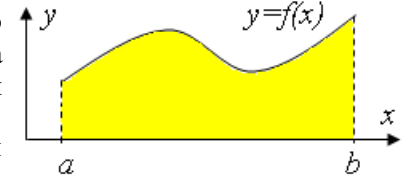
$$y''(x) = \frac{y(x+h) - 2y(x) + y(x-h)}{h^2} + O(h^2),$$

$$y''(x) = \frac{-y(x-3h) + 4y(x-2h) - 5y(x-h) + 2y(x)}{h^2} + O(h^2),$$

$$y''(x) = \frac{-2y(x) + 5y(x+h) - 4y(x+2h) + y(x+3h)}{h^2} + O(h^2).$$

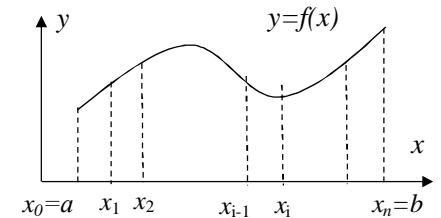
10. Численное интегрирование

Пусть функция $f(x)$ кусочно непрерывна на отрезке $[a; b]$ и ограничена на нём. Тогда существует интеграл $\int_a^b f(x) dx$, который численно равен площади S криволинейной трапеции (площади фигуры, ограниченной линией $y=f(x)$, осью OX , прямыми $x=a$ и $x=b$).

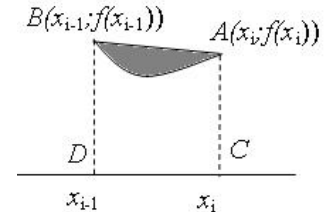


10.1. Формула трапеций

Разобьём отрезок $[a; b]$ на n одинаковых по длине отрезков. Тогда площадь всей криволинейной трапеции будет равна сумме площадей S_i маленьких криволинейных трапеций.



Рассмотрим отдельно маленькую криволинейную трапецию.



Соединим прямой две точки $B(x_{i-1}; f(x_{i-1}))$ и $A(x_i; f(x_i))$, получим прямоугольную трапецию $ABDC$, площадь которой равна T_i . Вычислим T_i как произведение высоты h и полусуммы оснований $f(x_{i-1})$, $f(x_i)$:

$$h = (x_i - x_{i-1}) = \frac{b-a}{n}, \quad T_i = \frac{f(x_{i-1}) + f(x_i)}{2} \cdot (x_i - x_{i-1}) = \frac{h}{2} (f(x_{i-1}) + f(x_i)).$$

Ошибка замены S_i на T_i составляет площадь заштрихованного сегмента. Чем меньше шаг h (чем больше n), тем меньше эта ошибка.

Просуммируем все T_i . Правое основание $f(x_i)$ в T_i является левым основанием для соседней трапеции T_{i+1} , а левое основание $f(x_{i-1})$ является правым для другой соседней трапеции T_{i-1} . Так почти все $f(x_i)$ будут присутствовать сумме дважды. Исключение составляют $f(a)$ и $f(b)$, они будут входить в сумму единожды:

$$T_n = \frac{h}{2} \left(f(a) + 2 \sum_{i=1}^{n-1} f(x_i) + f(b) \right) + O(h^2). \quad (10.1.1)$$

Выражение (10.1-1) называют *формулой трапеций* для вычисления значения определённого интеграла. Слагаемое $O(h^2)$ означает, что порядок

точности формулы равен двум, ошибка (отличие T_n от точного значения интеграла) уменьшается с уменьшением шага h разбиения отрезка пропорционально квадрату шага. Есть более точные представления этой ошибки, связанные с максимальным значением модуля второй производной от подынтегральной функции. Для многих функций вычисление их второй производной и поиск её максимума, если и возможен, то проблематичен.

На практике иногда поступают так. Для некоторого n вычисляют T_n , затем вычисляют T_{2n} и сравнивают T_n и T_{2n} . Если они почти одинаковы

$$|T_n - T_{2n}| < \varepsilon, \quad (10.1.2)$$

то вычисления прекращаются и T_{2n} объявляется значением интеграла, вычисленным с точностью ε ($\varepsilon > 0$). А если они ещё сильно отличаются друг от друга, то n удваивается и опять производится сравнение (10.1.2). В конце концов, для какого-то большого n условие (10.1.2) выполнится.

Для оценки погрешности уточнения результатов вычислений также применяют метод Рунге–Ромберга (см. главу 11).

Пример. Вычислить $I = \int_0^{\pi/2} \sin x dx$. (10.1.3)

Решение. Выпишем таблицу значений подынтегральной функции для $n=4$ (с шагом $\pi/8$):

X	0	$\pi/8$	$\pi/4$	$3\pi/8$	$\pi/2$
Y	0	0,382683	0,707107	0,92388	1

$$T_4 = \frac{\pi}{16} (0 + 2(0,382683 + 0,707107 + 0,92388) + 1) \approx 0,987116.$$

Выпишем таблицу значений подынтегральной функции для $n=8$ ($h=\pi/16$):

X	0	$\pi/16$	$\pi/8$	$3\pi/16$	$\pi/4$	$5\pi/16$	$3\pi/8$	$7\pi/16$	$\pi/2$
Y	0	0,1951	0,3827	0,5556	0,7071	0,8315	0,92388	0,9808	1

$$T_8 = \frac{\pi}{32} (0 + 2(0,1951 + 0,3827 + 0,5556 + 0,7071 + 0,8315 + 0,92388 + 0,9808) + 1) \approx 0,9968.$$

Так как $|T_4 - T_8| < 0,01$, то $T_8 = 0,9968$ можно считать приближённым с точностью 0,01 значением интеграла (10.1-3). После округления получим $I \approx 1,00$ ($\pm 0,01$).

Для $n=16$ ($h = \pi/32$) $T_{16} \approx 0,9998$.

10.2. Формула Симпсона

Рассмотрим не пары соседних узлов, а тройки узлов (по два отрезка). По трём точкам строится парабола. Она интегрируется и суммируются все такие короткие интегралы. Получится *формула Симпсона*:

$$S = \frac{h}{3} \left(f(a) + 4 \cdot \sum_{i=1,3,5..}^{n-1} f(x_i) + 2 \cdot \sum_{i=2,4,6..}^{n-2} f(x_i) + f(b) \right) + O(h^4). \quad (10.2.1)$$

Формула имеет четвёртый порядок точности, что гораздо лучше формулы трапеций. В первой сумме (которая учетверяется) присутствуют только значения подынтегральной функции в узлах с нечётными номерами, а во второй сумме – с чётными. При программировании следует учесть, что в первой сумме на одно слагаемое больше.

Так как метод рассматривает пары отрезков (тройки узлов), то n в формуле Симпсона должно быть чётным.

Пример. Вычислить

$$I = \int_0^{\pi/2} \sin x dx.$$

Решение. Выпишем таблицу значений подынтегральной функции для $n=4$ (с шагом $\pi/8$):

X	0	$\pi/8$	$\pi/4$	$3\pi/8$	$\pi/2$
Y	0	0,382683	0,707107	0,92388	1

$$S_4 = \frac{\pi}{24} (0 + 4(0,382683 + 0,92388) + 2 \cdot 0,707107 + 1) \approx 1,00013.$$

Выпишем таблицу значений подынтегральной функции для $n=8$ ($h=\pi/16$):

X	0	$\pi/16$	$\pi/8$	$3\pi/16$	$\pi/4$	$5\pi/16$	$3\pi/8$	$7\pi/16$	$\pi/2$
Y	0	0,1951	0,3827	0,5556	0,7071	0,8315	0,92388	0,9808	1

$$S_8 = \frac{\pi}{48} (0 + 4(0,1951 + 0,5556 + 0,8315 + 0,9808) + 2(0,3827 + 0,7071 + 0,92388) + 1) \approx 1,000008.$$

Так как $|S_4 - S_8| < 0,000122$, то $S_8 = 1,000008$ можно считать приближённым с точностью 0,00012 значением интеграла (10.2.2). После округления получим $I \approx 1,0000$ ($\pm 0,0001$). На самом деле (так как мы знаем точное значение этого интеграла, оно равно 1) точность значения 1,000008 равна 0,000008.

11. Формулы Рунге и Рунге–Ромберга

Рассмотрим ситуацию, когда у Вас есть несколько результатов расчётов по одной и той же формуле, но с разными шагами. И возникла необходимость иметь более точный результат, чем самый точный из имеющихся уже расчётов. Предположим, что новый более точный расчёт не возможен или слишком дорог. Как уточнить результаты имеющихся расчётов?

Пусть z_1 и z_2 - расчёты, сделанные по одной и той же формуле соответственно с шагами h_1 и h_2 . Пусть известен p (порядок точности

формулы). Обозначим $R=h_2/h_1$. Тогда уточнённое значение найдём по формуле $z_{pp} = z_1 + \frac{z_1 - z_2}{R^p - 1} + O(h^{p+1})$. (11.1)

Это формула Рунге–Ромберга. Порядок точности её $p+1$.

Если уточняемых q расчётов более двух, то составим два детерминанта:

$$D_1 = \begin{vmatrix} z_1 & h_1^p & h_1^{p+1} & \dots & h_1^{p+q-2} \\ z_2 & h_2^p & h_2^{p+1} & \dots & h_2^{p+q-2} \\ \dots & \dots & \dots & \dots & \dots \\ z_q & h_q^p & h_q^{p+1} & \dots & h_q^{p+q-2} \end{vmatrix}, D_2 = \begin{vmatrix} 1 & h_1^p & h_1^{p+1} & \dots & h_1^{p+q-2} \\ 1 & h_2^p & h_2^{p+1} & \dots & h_2^{p+q-2} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & h_q^p & h_q^{p+1} & \dots & h_q^{p+q-2} \end{vmatrix}. \quad (11.2)$$

Результатом первого расчёта с шагом h_1 явилось вычисление z_1 . Результатом второго расчёта с шагом h_2 явилось вычисление z_2 . В последнем расчёте с шагом h_q вычислили z_q . Детерминанты отличаются друг от друга только первым столбцом. Разделив их друг на друга, получим формулу Рунге:

$$z_p = D_1 / D_2 + O(h^{p+1}). \quad (11.3)$$

Эта формула тоже увеличивает на единицу порядок точности уточнённого значения.

Пример 1. Пусть имеем три расчёта по формуле трапеций с шагами $h_1=\pi/8$, $h_2=\pi/16$ и $h_3=\pi/32$ с результатами $z_1=0,9871158$, $z_2=0,996785$, $z_3=0,9991967$. Порядок точности формулы трапеций $p=2$. Для уточнения воспользуемся формулой (11.3):

$$z_p = \frac{\begin{vmatrix} 0,98711580 & 0,15421257 & 0,06055913 \\ 0,99678517 & 0,03855314 & 0,00756989 \\ 0,99919668 & 0,00963829 & 0,00094624 \end{vmatrix}}{\begin{vmatrix} 1 & 0,15421257 & 0,06055913 \\ 1 & 0,03855314 & 0,00756989 \\ 1 & 0,00963829 & 0,00094624 \end{vmatrix}}$$

$$z_p = \frac{-0,000766087717948747}{-0,000766088173447956} = 0,9999994054.$$

Погрешность этого результата равна 0,00000006, что гораздо лучше самого точного из трёх расчётов.

Пример 2. Пусть имеем два расчёта по формуле Симпсона с шагами $h_1=\pi/8$ и $h_2=\pi/16$ с результатами $z_1=1,00013458$, $z_2=1,000008296$. Порядок точности формулы Симпсона $p=4$. Для уточнения воспользуемся формулой (10-1):

$$z_{pp} = 1,0001345850 + \frac{1,0001345850 - 1,0000082955}{0,5^4 - 1} = 0,999999876.$$

Как видно из примеров, при высокой стоимости экспериментов (вычислений) можно успешно прогнозировать более точный результат при известных результатах менее дорогих экспериментов (вычислений).

Замечание. В случае уточнения двух расчётов результаты уточнения по обеим формулам совпадают ($z_p = z_{pp}$).

12. Решение задачи Коши для ОДУ

Рассмотрим движение ракеты вблизи поверхности (без учёта сферичности) Земли. Запишем второй закон Ньютона:

$$\frac{d}{dt}(m\bar{V}) = \bar{F}, \quad (12.1)$$

где m – масса ракеты, \bar{V} – вектор скорости, \bar{F} – результирующая действующих на ракету сил (\bar{P} – тяга двигателя, $\bar{F}_{\text{сопр}}$ – сила аэродинамического сопротивления, $\bar{F}_{\text{гр}}$ – сила тяжести).

$$\bar{F} = \bar{P} + \bar{F}_{\text{сопр}} + \bar{F}_{\text{гр}}; \quad P = \dot{m} P_{y0}; \quad \bar{F}_{\text{сопр}} = c \frac{\rho V^2}{2} s; \quad \bar{F}_{\text{гр}} = m \bar{g},$$

где \dot{m} – расход топлива, P_{y0} – удельная тяга, c – коэффициент аэродинамического сопротивления, s – мидель ракеты, ρ – плотность воздуха, g – ускорение свободного падения.

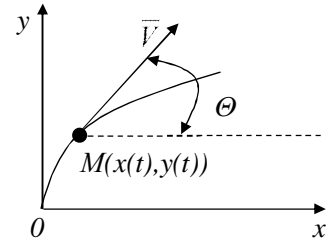
Пусть $O(0;0)$ – точка старта, а $x(t)$ и $y(t)$ – текущие координаты ракеты в момент времени t ; Θ – угол к горизонту вектора скорости $\bar{V}(\dot{x}(t); \dot{y}(t))$:

$$V(t) = \sqrt{(\dot{x}(t))^2 + (\dot{y}(t))^2};$$

$$\theta(t) = \arctg(\dot{y}(t)/\dot{x}(t)).$$

Проецируя (12.1) на оси OX и OY , получаем:

$$\begin{cases} m\ddot{x} + m\dot{x} = (P - c \frac{\rho V^2}{2} s) \cos \theta \\ m\ddot{y} + m\dot{y} = (P - c \frac{\rho V^2}{2} s) \sin \theta - mg \end{cases} \quad \text{или} \quad \begin{cases} \ddot{x} = \frac{1}{m} (P - c \frac{\rho V^2}{2} s) \cos \theta - \frac{\dot{m}}{m} \dot{x} \\ \ddot{y} = \frac{1}{m} (P - c \frac{\rho V^2}{2} s) \sin \theta - \frac{\dot{m}}{m} \dot{y} - g \end{cases}$$



Случай 1. Активный (до полной отработки топлива) участок. Пренебрегаем сопротивлением воздуха и получаем систему

$$\begin{cases} \frac{dx}{dt} = V \cos \theta, \\ \frac{dy}{dt} = V \sin \theta, \\ \frac{dV}{dt} = \frac{\dot{m}}{m} P_{y0} - g \sin \theta, \\ \frac{d\theta}{dt} = -\frac{g}{V} \cos \theta \end{cases}$$

с начальными условиями

$$\begin{cases} x(0) = 0, \\ y(0) = 0, \\ V(0) = 0, \\ \theta(0) = \theta_0 \end{cases} \quad (12.2)$$

Случай 2. После отработки топлива (отделения головной части) в момент времени t_1 оканчивается решение предыдущей системы с результатом x_1, y_1, V_1, θ_1 . Теперь двигатель не работает ($P=0, \dot{m}=0$). Система станет такой:

$$\begin{cases} \frac{dx}{dt} = V \cos \theta, \\ \frac{dy}{dt} = V \sin \theta, \\ \frac{dV}{dt} = -c \frac{\rho V^2}{2} \frac{s}{m} \cos \theta, \\ \frac{d\theta}{dt} = -\frac{g}{V} \cos \theta \end{cases}$$

с начальными условиями

$$\begin{cases} x(t_1) = x_1, \\ y(t_1) = y_1, \\ V(t_1) = V_1, \\ \theta(t_1) = \theta_1 \end{cases} \quad (12.3)$$

Условием окончания интегрирования системы может служить выполнение условия $y(t)=0$.

12.1. Решение задачи Коши для ОДУ первого порядка

Пусть на отрезке $x \in [a; b]$ дано обыкновенное дифференциальное уравнение (ОДУ) $y' = f(x; y)$. И известно значение y в начальной точке отрезка: $y(a) = y_0$. Такую задачу называют *задачей Коши* для ОДУ.

Весь отрезок $[a; b]$ разобьём на n отрезков с шагом $h = (b-a)/n$. Важно тут же отметить, что шаг разбиения отрезка (шаг сетки) может меняться, тогда и количество узлов сетки (строк табличной функции) заранее не известно.

На каждом i -м шаге алгоритма будем получать новый узел x_i сетки и вычислять для него y_i . После достижения правой границы отрезка $[a; b]$ вычисляется последний y_n и задача объявляется решённой. Ответом служит табличная функция x - y .

12.1.1. Метод Эйлера для ОДУ первого порядка

Если известно y_i – значение табличной функции при $x = x_i$, то можно вычислить новый узел $x = x_{i+1}$ и соответствующее ему y_{i+1} по формуле

$$y_{i+1} = y_i + h \cdot f(x_i; y_i) + O(h^2), \quad x_{i+1} = x_i + h. \quad (12.1.1.1)$$

В этой формуле шаг постоянен, но на практике шаг можно менять в зависимости от скорости изменения y . *Метод Эйлера* (12.1.1.1) имеет первый порядок точности.

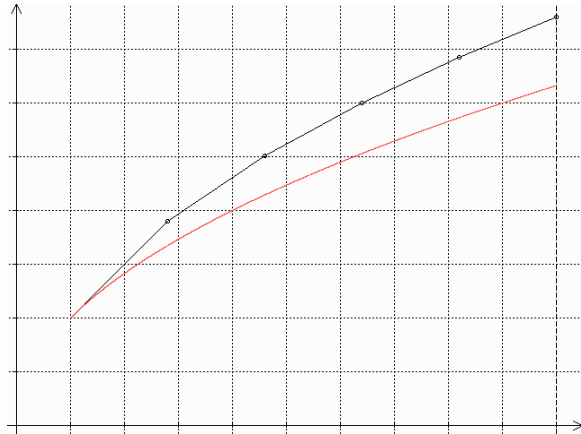
Пример. Для $x \in [1; 10]$ и известном $y(1) = 2$ найти решение дифференциального уравнения $y' = y / (2 \cdot x)$, разделив отрезок $[1; 10]$ на 5 частей одинаковой длины.

Решение. По условию шаг разбиения отрезка $[1; 10]$ постоянен и равен $h = 9/5 = 1,8$. Так как $x_0 = 1$ и $y_0 = 2$, то вычислим $x_1 = x_0 + h = 2,8$, $y_1 = y_0 + h \cdot f(x_0; y_0) = 2 + 1,8 \cdot 2 / (2 \cdot 1) = 3,8$. Теперь вычислим $x_2 = x_1 + h = 4,6$, $y_2 = y_1 + h \cdot f(x_1; y_1) = 3,8 + 1,8 \cdot 3,8 / (2 \cdot 2,8) = 5,02142857$. И так далее до последнего x_5 и y_5 . Результаты представлены в таблице:

i	x	y
0	1,0	2,0
1	2,8	3,8
2	4,6	5.0214285714
3	6,4	6.0038819876
4	8,2	6.8481778921
5	10,0	7.5998071729

Точным решением данного обыкновенного уравнения является функция $y = 2\sqrt{x}$. Для $x = 10$ значение точного решения $y \approx 6,32455532$. Таким образом, видно, что вычисленное y_5 сильно отличается от точного решения. Улучшить точность можно, уменьшив шаг интегрирования h .

Построим график точного решения и узлы, получившиеся при численном решении по методу Эйлера:



12.1.2. Метод Рунге–Кутты для ОДУ первого порядка

Если известно y_i — значение табличной функции при $x=x_i$, то можно вычислить для нового узла $x=x_{i+1}=x_i+h$ дополнительно четыре параметра:

$$\begin{aligned} K_1 &= f(x_i; y_i), \\ K_2 &= f(x_i + h/2; y_i + h/2 \cdot K_1), \\ K_3 &= f(x_i + h/2; y_i + h/2 \cdot K_2), \\ K_4 &= f(x_i + h; y_i + h \cdot K_3). \end{aligned} \quad (12.1.2.1)$$

После этого вычисляется y_{i+1} в новой точке x_{i+1} по формуле

$$y_{i+1} = y_i + h/6 \cdot (K_1 + 2 \cdot K_2 + 2 \cdot K_3 + K_4) + O(h^4). \quad (12.1.2.2)$$

Это формулы метода Рунге–Кутты четвёртого порядка. Как и метод Эйлера, метод Рунге–Кутты программируется одним циклом и допускает переменный шаг интегрирования. Если y_i и y_{i+1} мало отличаются друг от друга, шаг можно увеличить, а если функция y делает вдруг большой скачок, можно вернуться на шаг и, уменьшив шаг, сделать расчёт заново.

Пример. Для $x \in [1; 10]$ и известного $y(1)=2$ найти решение дифференциального уравнения $y' = y/(2 \cdot x)$, разделив отрезок $[1; 10]$ на 5 частей одинаковой длины $h=9/5$.

Решение. По условию шаг разбиения отрезка $[1; 10]$ постоянен и равен $h=9/5=1,8$. Так как $x_0=1$ и $y_0=2$, то вычислим $h/2=0,9$, $x_1=x_0+h=2,8$,

$$\begin{aligned} K_1 &= f(x_0; y_0) = 2/(2 \cdot 1) = 1, \\ K_2 &= f(x_0 + h/2; y_0 + h/2 \cdot K_1) = (2 + 0,9 \cdot 1)/(2 \cdot (1 + 0,9)) = 0,76316, \\ K_3 &= f(x_0 + h/2; y_0 + h/2 \cdot K_2) = (2 + 0,9 \cdot 0,76316)/(2 \cdot (1 + 0,9)) = 0,70706, \\ K_4 &= f(x_0 + h; y_0 + h \cdot K_3) = (2 + 1,8 \cdot 0,70706)/(2 \cdot (1 + 1,8)) = 0,58441, \end{aligned}$$

После этого вычисляется y_1 в новой точке x_1 по формуле

$$y_1 = y_0 + h/6 \cdot (K_1 + 2 \cdot K_2 + 2 \cdot K_3 + K_4) = 3,3574569648.$$

Теперь, используя только что вычисленные значения x_1 и y_1 , вычислим сначала четыре коэффициента K , а потом и y_2 :

$$\begin{aligned} K_1 &= f(x_1; y_1) = 3,3574569648 / (2 \cdot 2,8) = 0,59955, \\ K_2 &= f(x_1 + h/2; y_1 + h/2 \cdot K_1) = (3,3574569648 + 0,9 \cdot 0,59955) / (2 \cdot (2,8 + 0,9)) = 0,52663, \\ K_3 &= f(x_1 + h/2; y_1 + h/2 \cdot K_2) = (3,3574569648 + 0,9 \cdot 0,52663) / (2 \cdot (2,8 + 0,9)) = 0,51776, \\ K_4 &= f(x_1 + h; y_1 + h \cdot K_3) = (3,3574569648 + 1,8 \cdot 0,51776) / (2 \cdot (2,8 + 1,8)) = 0,46624 \end{aligned}$$

После этого для $x_2 = x_1 + h = 4,6$ вычисляется y_2 :

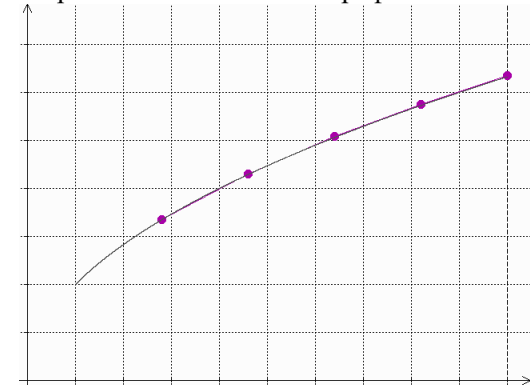
$$y_2 = y_1 + h/6 \cdot (K_1 + 2 \cdot K_2 + 2 \cdot K_3 + K_4) = 4,3038260183.$$

Продолжим вычисления и запишем результаты в таблицу:

i	x_i	y_i	K_1	K_2	K_3	K_4
0	1	2	1	0,76316	0,70706	0,58441
1	2,8	3,3574569648	0,59955	0,52663	0,51776	0,46624
2	4,6	4,3038260183	0,46781	0,42953	0,42640	0,39620
3	6,4	5,0765873529	0,39661	0,37216	0,37065	0,35023
4	8,2	5,7463263456	0,35039	0,33306	0,33220	0,31721
5	10	6,3457631160				

Для $x=10$ точное решение даёт $y \approx 6,32455532$, т.е. погрешность вычисленного $y_5 \approx 6,346$ не превышает 0,03.

Нарисуем точное решение и нанесём на график вычисленные точки:



Из рисунка видно, что точное решение почти сливается с точками таблицы x - y .

12.2. Решение задачи Коши для ОДУ второго порядка

Рассмотрим задачу Коши для дифференциального уравнения второго порядка, разрешённого относительно второй производной:

$$y'' = f(x, y, y'), x \in [a; b], y(a) = c, y'(a) = d. \quad (12.2.1)$$

С помощью замены $z = y'$ уравнение (11.2-1) превратится в систему

$$\begin{cases} y' = z \\ z' = f(x, y, z) \end{cases}, x \in [a; b], y(a) = c, z(a) = d. \quad (12.2.2)$$

Аналогично дифференциальное уравнение n -го порядка сводится к системе n -го порядка (n уравнений, каждое уравнение первого порядка).

Численное решение (12.2.2) заключается в построении (вычислении) таблицы x - y - z , в которой x меняется от a до b с некоторым шагом (одинаковым или разным). Рассмотрим те же два метода решения этой задачи: метод Эйлера (первого порядка) и метод Рунге–Кутты (четвёртого порядка).

12.2.1. Метод Эйлера для ОДУ второго порядка

Если известно y_i и z_i = значения табличной функции x - y - z при $x = x_i$, то можно вычислить новый узел $x = x_{i+1}$ и ему соответствующие y_{i+1} и z_{i+1} по формуле Эйлера для системы (11.2.2):

$$\begin{aligned} x_{i+1} &= x_i + h \\ y_{i+1} &= y_i + h \cdot f_1(x_i, y_i, z_i) = y_i + h \cdot z_i, \\ z_{i+1} &= z_i + h \cdot f_2(x_i, y_i, z_i) = z_i + h \cdot f(x_i, y_i, z_i). \end{aligned} \quad (12.2.1.1)$$

Для последнего $x_n = b$ вычисляется только y_n и таблица x - y (а если необходимо и z) готова.

В формулах (12.2.1.1) шаг постоянен, но на практике шаг можно менять.

Пример. Для $x \in [1; 10]$ и известном $y(1) = 2$, $y'(1) = 1$ найти решение дифференциального уравнения $y'' = -y/(2 \cdot x)$, разделив отрезок $[1; 10]$ на 10 частей одинаковой длины.

Решение. Для известных x_i, y_i, z_i вычисление $x_{i+1}, y_{i+1}, z_{i+1}$ производится по формулам: $x_{i+1} = x_i + h$; $y_{i+1} = y_i + h \cdot z_i$; $z_{i+1} = z_i + h \cdot (-z_i / (2 \cdot x_i))$.

По условию шаг разбиения отрезка $[1; 10]$ постоянен и равен $h = (10 - 1) / 10 = 0,9$. Так как $x_0 = 1$, $y_0 = 2$, $z_0 = 1$, то вычислим y_1 и z_1 :

$$\begin{aligned} x_1 &= x_0 + h = 1 + 0,9 = 1,9; \\ y_1 &= y_0 + h \cdot z_0 = 2 + 0,9 \cdot 1 = 2,9; \\ z_1 &= z_0 + h \cdot (-z_0 / (2 \cdot x_0)) = 1 - 0,9 \cdot (1 / (2 \cdot 1)) = 0,55. \end{aligned}$$

Теперь, зная x_1, y_1, z_1 , вычислим x_2, y_2, z_2 :

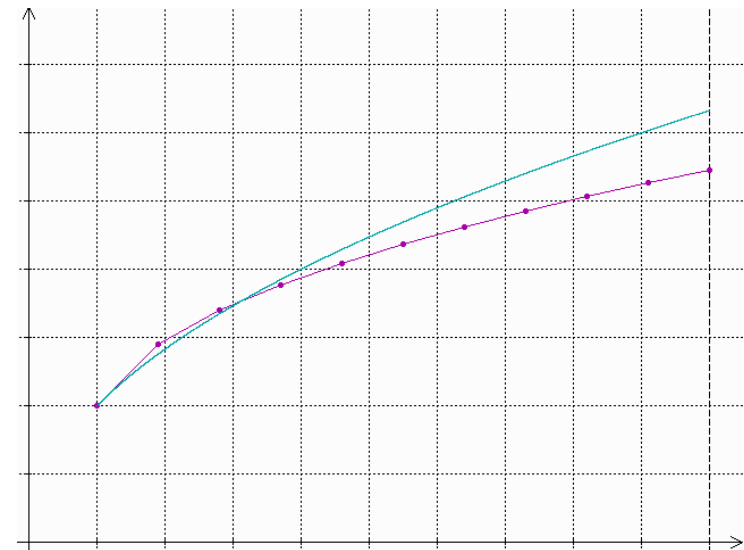
$$\begin{aligned} x_2 &= x_1 + h = 1,9 + 0,9 = 2,8; \\ y_2 &= y_1 + h \cdot z_1 = 2,9 + 0,9 \cdot 0,55 = 3,395; \\ z_2 &= z_1 + h \cdot (-z_1 / (2 \cdot x_1)) = 0,55 - 0,9 \cdot (0,55 / (2 \cdot 1,9)) = 0,3866606. \end{aligned}$$

Проведём расчёты до $x_{10} = 10$ и запишем их в таблицу:

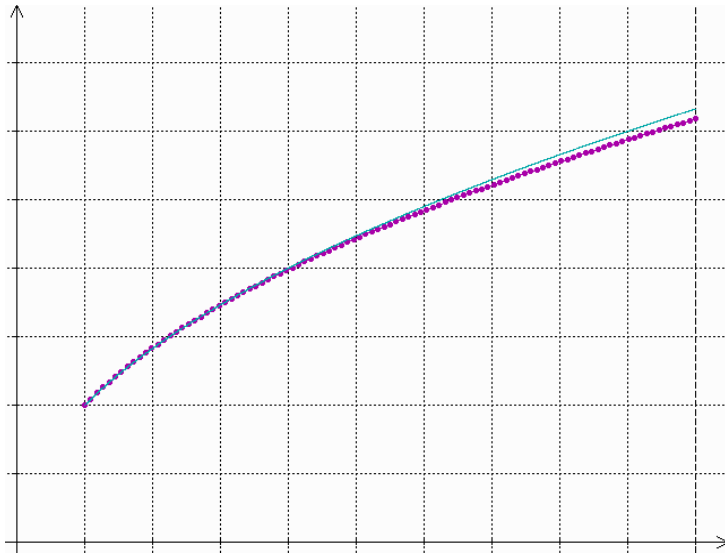
i	x	y	z
0	1.	2.	1.
1	1.9	2.9	0.55

2	2.80000	3.3950000000	0.4197368421
3	3.70000	3.7727631579	0.3522791353
4	4.60000	4.0898143797	0.3094343756
5	5.50000	4.3683053178	0.2791636215
6	6.40000	4.6195525771	0.2563229616
7	7.30000	4.8502432425	0.2383002533
8	8.20000	5.0647134705	0.2236105117
9	9.10000	5.2659629310	0.2113392031
10	10.00000	5.4561682138	0.2008883634

Сравнить результаты расчёта с точным решением (линия без узелков) можно по графику:



С увеличением числа узлов (уменьшением шага), погрешность метода уменьшается линейно относительно шага. Для $n = 100$ график выглядит так:



Пример 2. Численное моделирование орбиты спутника (задача Кеплера). Рассмотрим движение спутника вокруг планеты. Начало координат поместим в центр планеты. Спутник представим точкой $M(x;y)$. Движение точки M на плоскости XOY согласно закону Ньютона описывается системой дифференциальных уравнений :

$$\begin{cases} \frac{d^2 x}{dt^2} = -\frac{GM}{r^3} x \\ \frac{d^2 y}{dt^2} = -\frac{GM}{r^3} y \end{cases} \quad (12.2.1.2)$$

где $r^2 = x^2 + y^2$, а нормировочный коэффициент $GM = 4\pi^2$ (а.е.)³ / год².

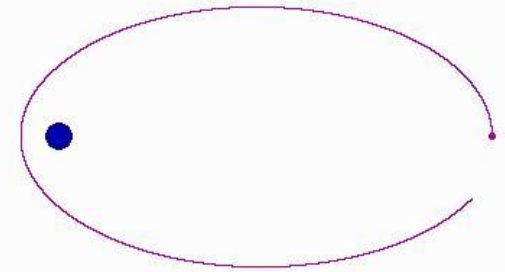
Преобразуем систему (12.2.1-2) в систему

$$\begin{cases} \frac{dx}{dt} = q \\ \frac{dq}{dt} = -\frac{GM}{r^3} x \\ \frac{dy}{dt} = z \\ \frac{dz}{dt} = -\frac{GM}{r^3} y \end{cases} \quad (12.2.1.3)$$

В качестве начальных данных (для $t=0$) возьмём, например, $x_0=10$, $y_0=0$, $q_0=0$, $z_0=0,8$. Формулы расчёта с шагом по t , равным h_t до момента времени T_{end} , будут выглядеть так:

$$\begin{aligned} x_{n+1} &= x_n + h_t \cdot q_n, & y_{n+1} &= y_n + h_t \cdot z_n, \\ q_{n+1} &= q_n + (-GM \cdot x_n / (x_n^2 + y_n^2)^{3/2}), & z_{n+1} &= z_n + (-GM \cdot y_n / (x_n^2 + y_n^2)^{3/2}). \end{aligned}$$

Программируются вычисления в одном цикле, а на экране траектория движения спутника выглядит:



12.2.2. Метод Рунге–Кутты для ОДУ второго порядка

Если известно y_i и z_i = значения табличной функции $x-y-z$ при $x=x_i$, то для системы (11.2-2) можно вычислить новый узел $x=x_{i+1} = x_i + h$ и ему соответствующие y_{i+1} и z_{i+1} по формуле Рунге-Кутты (четвёртого/порядка), вычислив сначала восемь промежуточных коэффициентов (четыре пары):

$$\begin{aligned} K_{1y} &= z_i, & K_{1z} &= f(x_i; y_i; z_i), \\ K_{2y} &= z_i + h/2 \cdot K_{1z}, & K_{2z} &= f(x_i + h/2; y_i + h/2 \cdot K_{1y}; z_i + h/2 \cdot K_{1z}), \\ K_{3y} &= z_i + h/2 \cdot K_{2z}, & K_{3z} &= f(x_i + h/2; y_i + h/2 \cdot K_{2y}; z_i + h/2 \cdot K_{2z}), \\ K_{4y} &= z_i + h \cdot K_{3z}, & K_{4z} &= f(x_i + h; y_i + h \cdot K_{3y}; z_i + h \cdot K_{3z}). \end{aligned}$$

После этого вычисляется y_{i+1} и z_{i+1} в новой точке $x_{i+1} = x_i + h$ по формулам :

$$\begin{aligned} y_{i+1} &= y_i + h/6 \cdot (K_{1y} + 2 \cdot K_{2y} + 2 \cdot K_{3y} + K_{4y}), \\ z_{i+1} &= z_i + h/6 \cdot (K_{1z} + 2 \cdot K_{2z} + 2 \cdot K_{3z} + K_{4z}). \end{aligned}$$

Для последнего $x_n = b$ вычисляется только y_n и таблица $x-y$ (а если необходимо и z) готова.

Пример. Для $x \in [1;10]$ при известном $y(1)=2$ и $y'(1)=1$ найти решение дифференциального уравнения $y'' = -y/(2 \cdot x)$, разделив отрезок $[1;10]$ на 5 частей одинаковой длины $h=9/5$.

Решение. Для известных x_i, y_i, z_i вычисление $x_{i+1}, y_{i+1}, z_{i+1}$ производится по формулам : $x_{i+1} = x_i + h$;

$$\begin{aligned} K_{1y} &= z_i, & K_{1z} &= -z_i / (2 \cdot x_i), \\ K_{2y} &= z_i + h/2 \cdot K_{1z}, & K_{2z} &= -(z_i + h/2 \cdot K_{1z}) / (2 \cdot (x_i + h/2)), \\ K_{3y} &= z_i + h/2 \cdot K_{2z}, & K_{3z} &= -(z_i + h/2 \cdot K_{2z}) / (2 \cdot (x_i + h/2)), \\ K_{4y} &= z_i + h \cdot K_{3z}, & K_{4z} &= -(z_i + h \cdot K_{3z}) / (2 \cdot (x_i + h)). \end{aligned}$$

По условию шаг разбиения отрезка $[1;10]$ постоянен и равен $h=(10-1)/5=1,8$. Так как $x_0=1$, $y_0=2$, $z_0=1$ то вычислим x_1 , восемь коэффициентов K , y_1 и z_1 : $x_1=x_0+h=1+1,8=2,8$;

$$\begin{aligned} K_{1y} &= z_0 = 1, & K_{1z} &= -z_0/(2 \cdot x_0) = -1/(2 \cdot 1) = -0,5, \\ K_{2y} &= z_0 + h/2 \cdot K_{1z} = 0,55, & K_{2z} &= -(z_0 + h/2 \cdot K_{1z}) / (2 \cdot (x_0 + h/2)) = -0,144737, \\ K_{3y} &= z_0 + h/2 \cdot K_{2z} = 0,869737, & K_{3z} &= -(z_0 + h/2 \cdot K_{2z}) / (2 \cdot (x_0 + h/2)) = -0,228878, \\ K_{4y} &= z_0 + h \cdot K_{3z} = 0,588019, & K_{4z} &= -(z_0 + h \cdot K_{3z}) / (2 \cdot (x_0 + h)) = -0,105003, \end{aligned}$$

После этого вычисляется y_1 и z_1 в новой точке x_1 по формулам

$$y_1 = y_0 + h/6 \cdot (K_{1y} + 2 \cdot K_{2y} + 2 \cdot K_{3y} + K_{4y}) = 3,328248,$$

$$z_1 = z_0 + h/6 \cdot (K_{1z} + 2 \cdot K_{2z} + 2 \cdot K_{3z} + K_{4z}) = 0,594330.$$

Проведём расчёты до конца отрезка и запишем их в таблицу:

i	1	2	3	4	5
K_{1y}	1.000000	0.594330	0.463629	0.393054	0.347243
K_{1z}	-0.500000	-0.106130	-0.050394	-0.030707	-0.021173
K_{2y}	0.550000	0.498813	0.418274	0.365418	0.328187
K_{2z}	-0.144737	-0.067407	-0.038025	-0.025029	-0.018032
K_{3y}	0.869737	0.533664	0.429407	0.370529	0.331014
K_{3z}	-0.228878	-0.072117	-0.039037	-0.025379	-0.018188
K_{4y}	0.588019	0.464520	0.393363	0.347373	0.314506
K_{4z}	-0.105003	-0.050491	-0.030731	-0.021181	-0.015725
X	2.800000	4.600000	6.400000	8.200000	10.000000
Y	3.328248	4.265389	5.031095	5.694791	6.288836
Z	0.594330	0.463629	0.393054	0.347243	0.314442

Точное значение $y(10)=6,324555$, а вычисленное $y_5=6.288836$, т.е. относительная погрешность менее 0,01 (менее 1%).

Метод Рунге-Кутты настолько точен, что даже для такого большого шага расчётные точки на графике визуально точно ложатся на график точного решения:



Рассмотренные методы решения задачи Коши допускают применение приёма, называемого *автоматическим выбором шага*.

Воспользуемся правилом Рунге-Ромберга для оценки погрешности метода вычисления y в точке $x=x_i$. Пусть для $x=x_i$ вычислены два разных y (y_{ih} и $y_{ih/2}$) с разными шагами (h и $h/2$). Если p — порядок метода и ε — допустимая погрешность, то на каждом шаге (в каждой точке $x=x_i$) будем вычислять:

$$\varepsilon_i^* = \frac{|y_{ih} - y_{ih/2}|}{2^p - 1}.$$

Если ε_i^* не выходит за пределы интервала $(\varepsilon; \varepsilon/2^{p+1})$, то шаг интегрирования не меняют. Если $\varepsilon_i^* < \varepsilon$, то шаг увеличивают, а если $\varepsilon_i^* > \frac{\varepsilon}{2^{p+1}}$, то уменьшают в два раза.

Известно много других методов решения задачи Коши. При решении конкретной практической задачи можно выбрать метод, который лучше других подходит для решения данной конкретной задачи.

13. Решение краевой задачи для ОДУ

Рассмотрим численное решение дифференциального уравнения второго порядка на отрезке $x \in [a; b]$:

$$K(x) \cdot y'' + L(x) \cdot y' + M(x) \cdot y = F(x). \quad (13.1)$$

С заданными краевыми (граничными) условиями третьего рода на границах отрезка (при $x=a$ и $x=b$):

$$\begin{aligned} R \cdot y'(a) + S \cdot y(a) &= T, \\ V \cdot y'(b) + W \cdot y(b) &= Z. \end{aligned} \quad (13.2)$$

Краевыми условиями первого рода называют условия на функцию (например, $S \cdot y(a) = T$). Краевыми условиями второго рода называют условия на производную функции (например, $R \cdot y'(a) = T$). Смешанными краевыми условиями называют условия разного рода на левой и правой границах отрезка $[a; b]$.

Коэффициенты K, L, M уравнения (13.1) могут быть постоянными. Это не сильно упростит решение, поэтому рассмотрим более общий случай, когда они зависят от x , как и искомая функция $y(x)$.

13.1. Метод конечных разностей

Разделим отрезок $[a; b]$ n точками на $n-1$ отрезков одинаковой длины $h=(b-a)/(n-1)$ точками $x_i=a+i \cdot h$ ($x_1=a$, $x_2=a+h, \dots, x_n=b$). Значение искомой функции в точке $x=x_i$ обозначим y_i . Целью вычислений является нахождение таблицы $x-y$. В этой таблице неизвестными являются n значений y_i , значит для их нахождения необходимо найти n уравнений. Два

уравнения дадут краевые условия (13.2) и $n-2$ уравнений получим, записав дифференциальное уравнение (13.1) в $n-2$ внутренних точках (x_2, x_3, \dots, x_{n-1}) отрезка $[a;b]$. Воспользуемся соотношениями:

$$y'(a) = \frac{y_2 - y_1}{h} + O(h), \quad y'(b) = \frac{y_n - y_{n-1}}{h} + O(h) \quad (13.3)$$

для записи краевых условий (13.2) в конечно-разностном виде

$$R \cdot \frac{y_2 - y_1}{h} + S \cdot y_1 = T \quad \text{и} \quad V \cdot \frac{y_n - y_{n-1}}{h} + W \cdot y_n = Z \quad \text{или окончательно}$$

$$-\left(\frac{R}{h} - S\right)y_1 + \frac{R}{h}y_2 = T \quad \text{и} \quad \left(\frac{V}{h}\right)y_{n-1} - \left(\frac{V}{h} + W\right)y_n = -Z. \quad (13.4)$$

Это первое и последнее (n -е) уравнения будущей системы (СЛАУ) с трёхдиагональной матрицей.

Остальные $n-2$ уравнения получим из дифференциального уравнения, записывая его для каждой внутренней точки $x_i, i=2,3,\dots,n-1$, заменяя y, y', y'' в точке x_i по формулам

$$y(x_i) = y_i, \quad y'(x_i) = \frac{y_{i+1} - y_{i-1}}{2h} + O(h^2), \quad y''(x_i) = \frac{y_{i-1} - 2y_i + y_{i+1}}{h^2} + O(h^2). \quad (13.5)$$

Итак, для каждой внутренней точки $x_i, i=2,3,\dots,n-1$ получим конечно-разностное соотношение

$$\left(\frac{K(x_i)}{h^2} - \frac{L(x_i)}{2h}\right)y_{i-1} - \left(\frac{2K(x_i)}{h^2} - M(x_i)\right)y_i + \left(\frac{K(x_i)}{h^2} + \frac{L(x_i)}{2h}\right)y_{i+1} = F(x_i) \quad (13.6)$$

Это соотношение является уравнением (с номером i) искомой СЛАУ с трёхдиагональной матрицей.

Если соберём вместе в систему уравнения (13-4) и (13-6), получим

$$\begin{pmatrix} b_1 & c_1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & d_1 \\ a_2 & b_2 & c_2 & 0 & 0 & \dots & 0 & 0 & 0 & d_2 \\ 0 & a_3 & b_3 & c_3 & 0 & \dots & 0 & 0 & 0 & d_3 \\ 0 & 0 & a_4 & b_4 & c_4 & \dots & 0 & 0 & 0 & d_4 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & b_{n-2} & c_{n-2} & 0 & d_{n-2} \\ 0 & 0 & 0 & 0 & 0 & \dots & a_{n-1} & b_{n-1} & c_{n-1} & d_{n-1} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & a_n & b_n & d_n \end{pmatrix} \quad (13.7)$$

В этой расширенной матрице

$$\begin{aligned} b_1 &= -R/h + S, & c_1 &= R/h, & d_1 &= T, \\ a_2 &= K(x_2)/h^2 - L(x_2)/2h, & b_2 &= -2K(x_2)/h^2 + M(x_2), \\ c_2 &= K(x_2)/h^2 + L(x_2)/2h, & d_2 &= F(x_2), & a_3 &= K(x_3)/h^2 - L(x_3)/2h, \\ b_3 &= -2K(x_3)/h^2 + M(x_3), & c_3 &= K(x_3)/h^2 + L(x_3)/2h, & d_3 &= F(x_3), \dots, \\ a_n &= V/h, & b_n &= -V/h - W, & d_n &= -Z. \end{aligned} \quad (13.8)$$

Решим эту систему методом прогонки (см. раздел 1.3) и получим в качестве ответа значения y_i ($i=1,\dots,n$) табличной функции x - y .

Имеющуюся таблицу можно аппроксимировать теми методами, которые были рассмотрены ранее.

Пример 1. Найти решение уравнения $2x^2y'' + xy' + y = 2\sqrt{x}$ на отрезке $[1;10]$ при заданных соотношениях $y(1)=2, y(10)=2\sqrt{10}$.

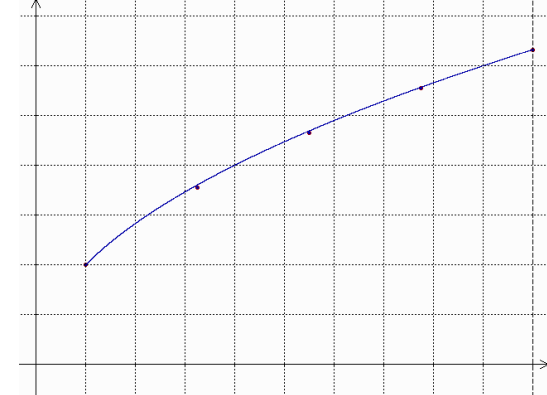
Решение. Разделим отрезок на четыре отрезка равной длины $h=9/4$ точками $x_1=1, x_2=1+2,25=3,25, x_3=3,25+2,25=5,5, x_4=5,5+2,25=7,75$ и $x_5=10$. Краевые условия первого рода дают первое и пятое уравнения системы: $y_1=2, y_5=2\sqrt{10}$. $K(x)=2x^2, L(x)=x, M(x)=1, R=0, S=1, T=2, V=0, W=1, Z=2\sqrt{10}$. Используя формулы (13-8), получаем систему:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 2 \\ 3,45062 & -7,34568 & 4,89506 & 0 & 0 & 3,60555 \\ 0 & 10,72840 & -22,90123 & 13,17284 & 0 & 4,69042 \\ 0 & 0 & 22,00617 & -46,45679 & 25,45062 & 5,56776 \\ 0 & 0 & 0 & 0 & 1 & 6,32456 \end{pmatrix}$$

Используем формулы метода прогонки (1.3.3) и найдём прогоночные коэффициенты p и q : $p_1=0, q_1=0, p_2=0, q_2=2, p_3=0,6663865546, q_3=0,4486560449, p_4=0,8362658257, q_4=0,0078049594, p_5=0,9072081354, q_5=0,1923450997, p_6=0, q_6=6.3245553203$. Теперь с обратным ходом метода прогонки (1.3.4) найдём все y и запишем их в таблицу:

X	1	3,25	5,5	7,75	10
Y	2	3.5441453737	4.6451857520	5.5453429398	6.3245553203

Построим полученные точки и график точного решения $y=2\sqrt{x}$:



Для $x=5,5$ точное решение $y_T \approx 4,6904$. Относительная погрешность вычисленного методом конечных разностей y менее 1%. Как видно, уже для $N=5$ таблица метода конечных разностей даёт хорошее представление о решении исходной краевой задачи. При увеличении N (уменьшении шага h

разбиения отрезка) вычисленные y_i быстро стремятся к точному решению. Порядок точности метода – два (h^2).

Изложение материала этого параграфа подразумевает необходимость программирования решения поставленной задачи. Теперь рассмотрим простейший пример, который без труда решается за 15-25 минут, например, на лекции.

Пример 2. Найти решение уравнения $2 \cdot y'' + 2 \cdot y' - 4 \cdot y = 1 - 2 \cdot x$ для $x \in [0; 4]$ при краевых условиях $y(0)=1$ и $2 \cdot y'(4) + 4 \cdot y(4) = 9$.

Решение. Разделим отрезок $[0; 4]$ на четыре отрезка единичной длины ($h=1$). Тогда $x_1=0, x_2=1, x_3=2, x_4=3, x_5=4$. Будем искать $y_i=y(x_i)$. Первым уравнением будущей системы будет первое краевое условие: $y_1=1$. Чтобы получить ещё три уравнения будущей системы, воспользуемся соотношениями (13.1.3).

$$\text{Для } x_2=1: 2 \cdot \frac{y_1 - 2y_2 + y_3}{1^2} + 2 \cdot \frac{y_3 - y_1}{2 \cdot 1} - 4 \cdot y_2 = 1 - 2 \cdot 1 \quad \text{или} \quad y_1 - 8 \cdot y_2 + 3 \cdot y_3 = -1.$$

$$\text{Для } x_3=2: 2 \cdot \frac{y_2 - 2y_3 + y_4}{1^2} + 2 \cdot \frac{y_4 - y_2}{2 \cdot 1} - 4 \cdot y_3 = 1 - 2 \cdot 2 \quad \text{или} \quad y_2 - 8 \cdot y_3 + 3 \cdot y_4 = -3.$$

$$\text{Для } x_3=3: 2 \cdot \frac{y_3 - 2y_4 + y_5}{1^2} + 2 \cdot \frac{y_5 - y_3}{2 \cdot 1} - 4 \cdot y_4 = 1 - 2 \cdot 3 \quad \text{или} \quad y_3 - 8 \cdot y_4 + 3 \cdot y_5 = -5.$$

Для записи последнего уравнения системы используем (13.3) для краевого условия в точке $x_5=4$:

$$2 \cdot \frac{y_5 - y_4}{1} + 4 \cdot y_5 = 9 \quad \text{или} \quad -2 \cdot y_4 + 6 \cdot y_5 = 9.$$

Соберём все уравнения в систему:

$$\begin{cases} y_1 & & & & & = 1 \\ y_1 - 8 \cdot y_2 & + 3 \cdot y_3 & & & & = -1 \\ & y_2 & - 8 \cdot y_3 & + 3 \cdot y_4 & & = -3 \\ & & y_3 & - 8 \cdot y_4 & + 3 \cdot y_5 & = -5 \\ & & & -2 \cdot y_4 & + 6 \cdot y_5 & = 9 \end{cases}$$

Эту систему решим методом прогонки (разд. 1.3), а для этого уточним, как обозначаются и чему равны коэффициенты.

В первом уравнении $a_1=0, b_1=1, c_1=0, d_1=1$.

Из второго уравнения: $a_2=1, b_2=-8, c_2=3, d_2=-1$.

Из третьего уравнения: $a_3=1, b_3=-8, c_3=3, d_3=-3$.

Из четвёртого уравнения: $a_4=1, b_4=-8, c_4=3, d_4=-5$.

Из пятого уравнения: $a_5=-2, b_5=6, c_5=0, d_5=9$.

По формулам (1.3.3) вычислим прогоночные коэффициенты: ($p_1=0, q_1=0$),

$$p_2 = \frac{c_1}{-b_1 - a_1 \cdot p_1} = \frac{0}{-1 - 0 \cdot 0} = 0; \quad q_2 = \frac{a_1 \cdot q_1 - d_1}{-b_1 - a_1 \cdot p_1} = \frac{0 \cdot 0 - 1}{-1 - 0 \cdot 0} = 1;$$

$$p_3 = \frac{c_2}{-b_2 - a_2 \cdot p_2} = \frac{3}{8 - 1 \cdot 0} = 0,375; \quad q_3 = \frac{a_2 \cdot q_2 - d_2}{-b_2 - a_2 \cdot p_2} = \frac{1 \cdot 1 - (-1)}{8 - 1 \cdot 0} = 0,25;$$

$$p_4 = \frac{c_3}{-b_3 - a_3 \cdot p_3} = \frac{3}{8 - 1 \cdot 0,375} = 0,39344; \quad q_4 = \frac{a_3 \cdot q_3 - d_3}{-b_3 - a_3 \cdot p_3} = \frac{1 \cdot 0,25 - (-3)}{8 - 1 \cdot 0,375} = 0,4262295;$$

$$p_5 = \frac{c_4}{-b_4 - a_4 \cdot p_4} = \frac{3}{8 - 1 \cdot 0,39344} = 0,394396;$$

$$q_5 = \frac{a_4 \cdot q_4 - d_4}{-b_4 - a_4 \cdot p_4} = \frac{1 \cdot 0,4262295 - (-5)}{8 - 1 \cdot 0,39344} = 0,71336;$$

$$p_6 = \frac{c_5}{-b_5 - a_5 \cdot p_5} = \frac{0}{-6 - (-2) \cdot 0,3943} = 0;$$

$$q_6 = \frac{a_5 \cdot q_5 - d_5}{-b_5 - a_5 \cdot p_5} = \frac{(-2) \cdot 0,71336 - (9)}{-6 - (-2) \cdot 0,3943} = 2,0008.$$

Теперь по формулам (1.3.4) вычислим y :

$$y_5 = q_6 = 2,0008;$$

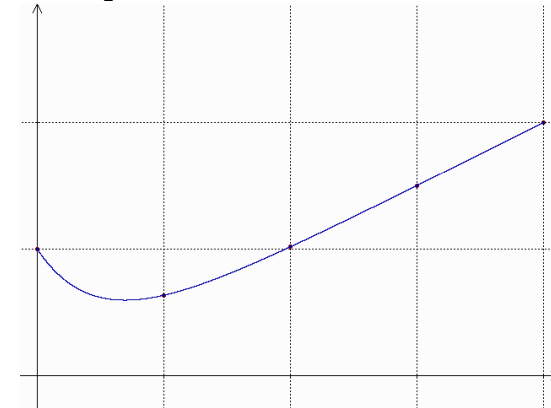
$$y_4 = q_5 + p_5 \cdot y_5 = 0,71336 + 0,394396 \cdot 2,0008 = 1,50248;$$

$$y_3 = q_4 + p_4 \cdot y_4 = 0,426229 + 0,39344 \cdot 1,50248 = 1,017369;$$

$$y_2 = q_3 + p_3 \cdot y_3 = 0,25 + 0,375 \cdot 1,017369 = 0,6315136;$$

$$y_1 = q_2 + p_2 \cdot y_2 = 1.$$

Осталось нарисовать пять точек (x_i, y_i) : $i=1, 2, \dots, 5$ и график точного решения $y(x) = e^{-2x} + x/2$:



13.2. Метод стрельбы

Рассмотрим метод решения краевой задачи для нелинейного дифференциального уравнения второго порядка

$$y'' = f(x, y, y'), \quad x \in [a; b] \quad (13.2.1)$$

с граничными условиями:

$$y(a) = c, \quad y(b) = d. \quad (13.2.2)$$

Метод решения заключается в многократном решении задачи Коши

$$y'' = f(x, y, y'), \quad x \in [a; b], \quad y(a) = c, \quad y'(a) = g. \quad (13.2.3)$$

В первом расчёте задачи Коши выбирается любое значение g и вычисляется $y(b)$. Если полученное $y(b)$ меньше заданного d , то в следующем решении задачи Коши (13.2.3) уменьшим g , если больше - увеличим. Так много раз решая задачу Коши, будем добиваться приближения $y(b)$ к d . Получается, что разность $\psi = y(b) - d$ зависит от выбираемого g , т.е. $\psi = \psi(g)$. Подбирая g , нужно добиться, чтобы ψ стало как можно ближе к нулю. Т.е. следует решить уравнение

$$\psi(g) = 0. \quad (13.2.4)$$

Методы решения таких уравнений мы уже рассматривали ранее, например, метод половинного деления или метод хорд.

Пример. Найти решение уравнения $y'' = (y')^2 \sqrt{x+1} / y^2$ на отрезке $x \in [0; 3]$ при заданных соотношениях $y(0)=1$, $y(3)=2$.

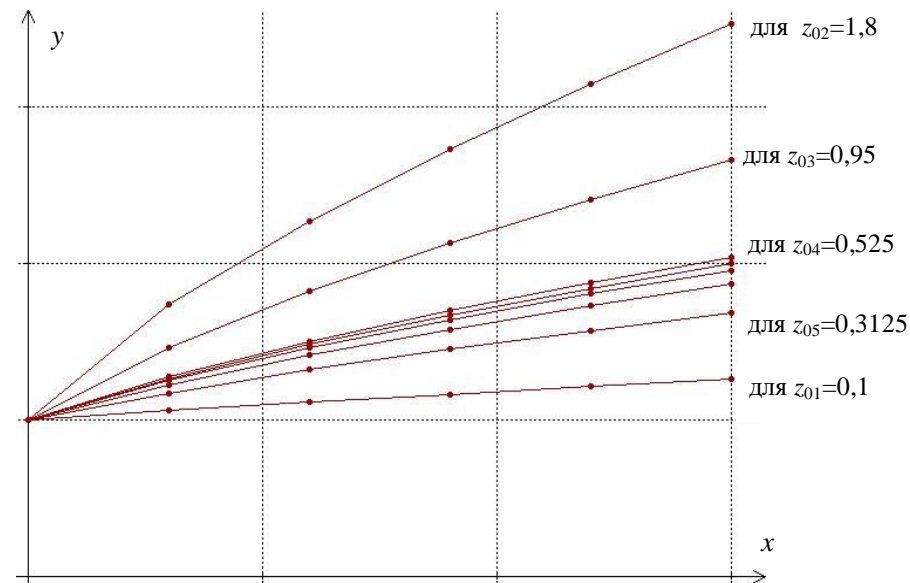
Решение. Разделим отрезок на пять отрезков равной длины $h=3/5$. В задаче даны краевые условия первого рода. Для решения задачи Коши к начальному условию $y(0)=1$ добавим условие $y'(0)=z_0$. Решим дважды задачу Коши методом Рунге–Кутты для $y'(0)=z_{01}=0,1$ и для $y'(0)=z_{02}=1,8$. В результате двух таких расчётов получим два разных $y(3)=y_{end_1}=1,25775$ и $y(3)=y_{end_2}=3,53024$ соответственно.

На самом деле $y(3)$ должно равняться 2. Так как $1,25775 < 2 < 3,53024$ ($y_{end_1} < 2 < y_{end_2}$), то разделим отрезок $[z_{01}; z_{02}]$ пополам точкой $z_{03}=0,95$. Третий раз решим задачу Коши с начальными данными $y(0)=1$, $y'(0)=z_{03}=0,95$. Получим $y(3)=2,66193$. Теперь сужаем отрезки $[z_{01}; z_{02}]$ и $[y_{end_1}; y_{end_2}]$: так как $2 < 2,66193 < 3,53024 = y_{end_2}$, то $z_{02}=z_{03}=0,95$ и $y_{end_2}=2,66193$.

Разделим новый отрезок $[z_{01}; z_{02}]$ пополам точкой $z_{04}=0,525$ и, четвёртый раз решив задачу Коши, получим $y(3)=2,03938$. Это значение ещё ближе к тому, которое должно быть (к $y(3)=2$). Опять сужаем отрезок $[z_{01}; z_{02}]$ и получим новый отрезок $[y_{end_1}; y_{end_2}]$: т.к. $2 < 2,03938 < 2,66193 = y_{end_2}$, то $z_{02}=z_{04}=0,525$ и $y_{end_2}=2,03938$.

Так, перемещая то левые, то правые границы отрезка $[z_{01}; z_{02}]$, будем получать новый всё меньшей длины отрезок $[y_{end_1}; y_{end_2}]$, содержащий значение $y(3)$.

Когда длина отрезка $[y_{end_1}; y_{end_2}]$ станет приемлемо малой, вычисления можно прекратить и считать результаты расчёта последней задачи Коши ответом. На рисунке представлены результаты расчётов нескольких задач Коши для этого примера:



14. Решение уравнения теплопроводности.

Метод сеток

Рассмотрим решение дифференциального уравнения в частных производных параболического типа с начально-краевыми условиями:

$$\frac{\partial y}{\partial t} = \alpha_1 \cdot \frac{\partial^2 y}{\partial x^2} + \alpha_2 \frac{\partial y}{\partial x} + \alpha_3 y + f(t, x), \quad a \leq x \leq b, \quad 0 \leq t \leq T, \quad \alpha_i > 0 \quad (14.1)$$

$$\text{для } x=a \quad \varphi_1 \frac{\partial y}{\partial x} + \varphi_2 y = f_1(t), \quad (14.2)$$

$$\text{для } x=b \quad \varphi_4 \frac{\partial y}{\partial x} + \varphi_5 y = f_2(t), \quad (14.3)$$

$$\text{для } t=0 \quad y(0, x) = f_3(x). \quad (14.4)$$

Накроем область сеткой с шагом по x равным h и с шагом по t равным τ . Тогда $x_0=a$, $x_1=a+h$, ..., $x_N=b$, $N=(b-a)/h$, $t_0=0$, $t_1=\tau$, $t_2=2\tau$, $t_M=T$, $M=T/\tau$.

Геометрически область представляет собой «стакан», с трёх сторон которого заданы начальные условия (14.4), слева и справа заданы краевые условия (14.2) и (14.3), а на верхней кромке (при $t=T$) значения функции $y(t, x)$ не известны. Их вычисление и является целью рассматриваемых алгоритмов. Коэффициенты α_i уравнения и φ_i в краевых условиях представляют собой константы (могут быть и равны 0).

Тогда $y(t_k, x_i) = y_i^k$. Назовём её сеточной функцией. Рассмотрим несколько способов решения этой задачи.

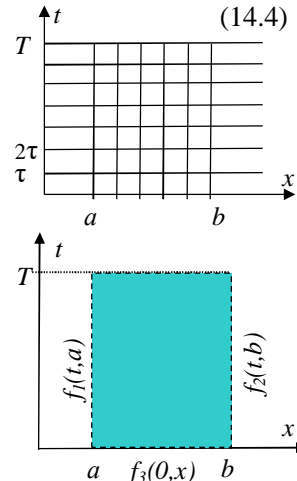
14.1. Явная схема

Для замены производных конечно-разностными соотношениями воспользуемся формулами главы 9.

$$\frac{\partial y}{\partial t} = \frac{y_i^{k+1} - y_i^k}{\tau} + O(\tau^1), \quad \frac{\partial^2 y}{\partial x^2} = \frac{y_{i+1}^k - 2y_i^k + y_{i-1}^k}{h^2} + O(h^2), \quad \frac{\partial y}{\partial x} = \frac{y_{i+1}^k - y_{i-1}^k}{2h} + O(h^2) \quad (14.5)$$

Подставим формулы (14.5) в (14.1) и выразим y_i^{k+1} на новом $k+1$ слое:

$$y_i^{k+1} = \left(\frac{\alpha_1 \tau}{h^2} - \frac{\alpha_2 \tau}{2h} \right) y_i^k + \left(\tau \alpha_3 + 1 - \frac{2\alpha_1 \tau}{h^2} \right) y_i^k + \left(\frac{\alpha_1 \tau}{h^2} + \frac{\alpha_2 \tau}{2h} \right) y_{i+1}^k + \tau \cdot f(x_i; t_k) + O(\tau + h^2). \quad (14.6)$$



Весь нулевой слой (при $t=0$) вычисляем, используя (14.4): $y_i^0 = f_3(x_i)$, $i=0, 1, 2, \dots, N$. Затем по формуле (14.6) вычисляем y_i^1 для внутренних точек $i=1, 2, 3, \dots, N-1$. Для вычисления y_0^1 и y_N^1 в граничных точках $x=a$ и $x=b$ воспользуемся формулами главы 9 первого порядка точности:

$$y^-(a) = \frac{y(a+h) - y(a)}{h} + O(h); \quad y^-(b) = \frac{y(b) - y(b-h)}{h} + O(h).$$

Получим $y_0^1 = \frac{hf_1(t_1) - \varphi_1 y_1^1}{h\varphi_2 - \varphi_1} + O(h); \quad y_N^1 = \frac{hf_2(t_1) + \varphi_4 y_{N-1}^1}{h\varphi_5 + \varphi_4} + O(h) \quad (14.7)$

Или второго порядка точности:

$$y^-(a) = \frac{-3y(a) + 4y(a+h) - y(a+2h)}{2h} + O(h^2) = \frac{1}{2h}(-3y_0^1 + 4y_1^1 - y_2^1)$$

$$y^-(b) = \frac{y(b-2h) - 4y(b-h) + 3y(b)}{2h} + O(h^2) = \frac{1}{2h}(y_{N-2}^1 - 4y_{N-1}^1 + 3y_N^1) \quad (14.8)$$

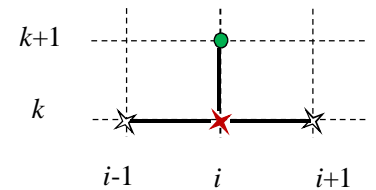
Использование формул (14.7) существенно ухудшит точность, поэтому использовать их не будем.

Подставляя (14.8) в (14.2) и (14.3), получим y_0^1 и y_N^1 :

$$y_0^1 = \frac{2hf_1(t_1) - 4\varphi_1 y_1^1 + \varphi_1 y_2^1}{2h\varphi_2 - 3\varphi_1} + O(h^2), \quad (14.9)$$

$$y_N^1 = \frac{2hf_2(t_1) + 4\varphi_4 y_{N-1}^1 - \varphi_4 y_{N-2}^1}{2h\varphi_5 + 3\varphi_4} + O(h^2) \quad (14.10)$$

В результате весь первый слой (для $t=1$) будет вычислен. Аналогично при полностью известном первом слое вычисляем второй (при $t_2=2\tau$) слой y_i^2 , $i=1, \dots, N-1$: сначала по (14.6) во внутренних точках, затем по (14.9) и (14.10) – в крайних. И так вплоть до $t_M=T$. Приведём графическое изображение явной конечно-разностной схемы:



Программируется явная схема двумя вложенными циклами: по t и по x . Используются два массива размерности $N+1$ каждый – для хранения значений функции y на старом слое k и вычисляемых на новом $k+1$ слое.

Преимущество алгоритма в простоте. Недостаток его в плохой (условной) сходимости. Чтобы получить достоверный результат на последнем слое $t=T$, τ (шаг по t) должен быть очень мал ($\tau < h^2/(2\alpha_1)$). Поэтому для получения результатов вычисления с хорошей точностью, надо уменьшать

не только шаг по x , но и многократно шаг по t . Это может существенно увеличить время расчёта.

Пример. Решить для $t \in [0; 2]$, $x \in [0; \pi]$ уравнение $y'_t = y''_{xx}$ при $y(t, 0) = 0$, $y(t, \pi) = 0$, $y(0, x) = \sin x$

Решение. Наложим на область сетку с шагом по x : $h_x = \pi/10$ и по t : $\tau = 0,1$. На нулевом слое при $t=0$ $y_{0j} = \sin x_j = \sin(j \cdot \pi/10)$. На новом (первом) слое известны (вычисляемы по формулам из условия) самое левое y_{10} и самое правое y_{110} , которые вычисляются:

$y_{10} = y(t, a) = f_1(t) = 0$, $y_{110} = y(t, b) = f_2(t) = 0$. Запишем формулу (14.1.2) для вычисления y в узле $(i+1, j)$:

$$y_{i+1j} \approx y_{ij} + \frac{\tau}{h^2} \cdot (y_{ij-1} - 2y_{ij} + y_{ij+1}),$$

И с её помощью вычислим $y_{11}, y_{12}, y_{13}, \dots, y_{19}$ на первом слое ($t=0,1$):
 $y_{11} = y_{01} + 0,1 / (\pi/10)^2 \cdot (y_{00} - 2y_{01} + y_{02}) = \sin(\pi/10) + 1,0132118 \cdot (\sin(0) - 2 \cdot \sin(\pi/10) + \sin(2 \cdot \pi/10)) \approx 0,2783686166$;

$y_{12} = y_{02} + 0,1 / (\pi/10)^2 \cdot (y_{01} - 2y_{02} + y_{03}) = \sin(2 \cdot \pi/10) + 1,0132118 \times (\sin(\pi/10) - 2 \cdot \sin(2 \cdot \pi/10) + \sin(3 \cdot \pi/10)) \approx 0,5294885734$;

.....

$y_{19} = y_{09} + 0,1 / (\pi/10)^2 \cdot (y_{08} - 2y_{09} + y_{010}) = \sin(9 \cdot \pi/10) + 1,0132118 \times (\sin(8 \cdot \pi/10) - 2 \cdot \sin(9 \cdot \pi/10) + \sin(\pi)) \approx 0,2783686166$.

Закончен просчёт значений функции y на первом неизвестном слое. Затем вычисляются y на втором слое ($t=0,2$):

$y_{21} = y_{11} + 0,1 / (\pi/10)^2 \cdot (y_{10} - 2y_{11} + y_{12}) = 0,2783686166 + 1,0132118 \cdot (0 - 2 \cdot 0,2783686166 + 0,5294885734) \approx 0,2507599520$;

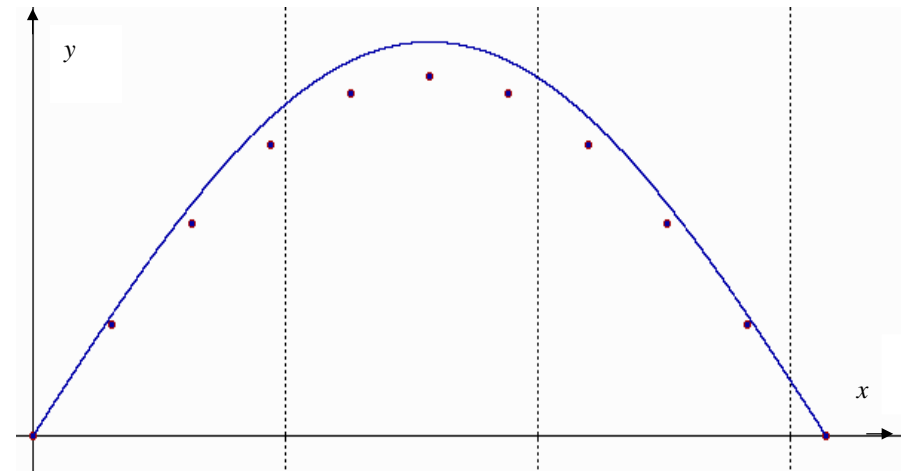
$y_{22} = y_{12} + 0,1 / (\pi/10)^2 \cdot (y_{11} - 2y_{12} + y_{13}) = \dots$

... ..

$y_{29} = y_{19} + 0,1 / (\pi/10)^2 \cdot (y_{18} - 2y_{19} + y_{110}) = \dots$

Так слой за слоем вычисляем y и через 20 таких шагов-слоёв вычислим значения y на последнем двадцатом слое при $t=2$. Эти одиннадцать чисел ($y_{200}, y_{201}, y_{202}, \dots, y_{2010}$), соответствующие одиннадцати значениям x ($x_0=0, x_1=\pi/10, x_2=\pi/5, \dots, x_{10}=\pi$) на координатной плоскости XOY , дают одиннадцать точек.

Построим на одной координатной плоскости одиннадцать найденных точек и точное решение этой задачи ($y=e^{-t} \cdot \sin x$):



Относительная погрешность вычисленных в узлах y оказалась одинаковой и равной $\approx 8,5\%$.

Если решать эту задачу, с шагом h по x , равным $\pi/100$, то для достижения устойчивости вычислительного процесса шаг τ по t придётся делать не более 0,00045, т.е. рассчитывать 4445 слоёв. Относительная погрешность вычисленных на последнем слое y равна 0,0003.

Вычисления по явной схеме программируются двумя вложенными циклами и двумя массивами (а можно обойтись и одним массивом) для хранения значений y на старом слое y_0 и на новом слое y , например, так:

```
t := 0 ; for j:= 0 to N do y0[j] := f3(0+j*h);
for i := 1 to Nt do
begin t := t + tau ;
for j:= 1 to N-1 do
y[j] := y0[j] + alfa*tau/h/h * ( y0[j+1]-
2*y0[j]+y0[j-1] ) +
tau * f( t-tau , x[j] ) ;
y[0] := f1( t ) ; y[N] := f2( t ) ;
for j:= 0 to N do y0[j] := y[j] ;
end ;
```

14.2. Неявная схема

В (14.1) производные будем вычислять в узле $(x_i; t_{k+1})$. Заменим производные на конечно-разностные соотношения

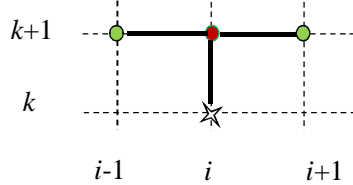
$$\frac{\partial y}{\partial t} = \frac{y_i^{k+1} - y_i^k}{\tau} + O(\tau^1); \quad \frac{\partial^2 y}{\partial x^2} = \frac{y_{i+1}^{k+1} - 2y_i^{k+1} + y_{i-1}^{k+1}}{h^2} + O(h^2);$$

$$\frac{\partial y}{\partial x} = \frac{y_{i+1}^{k+1} - y_{i-1}^{k+1}}{2h} + O(h^2). \quad (14.11)$$

Подставим эти соотношения в (14.1) и получим для каждой внутренней точки $(x_i; t_{k+1})$:

$$(2\alpha_1 - h \cdot \alpha_2) \frac{\tau}{2h^2} y_{i-1}^{k+1} + \left(\alpha_3 \tau - \frac{2\alpha_1 \tau}{h^2} - 1 \right) y_i^{k+1} + (2\alpha_1 + h \cdot \alpha_2) \frac{\tau}{2h^2} y_{i+1}^{k+1} = -y_i^k - \tau \cdot f(t_{k+1}; x_i) + O(\tau + h^2), \quad i=1, 2, 3, \dots, N-1. \quad (14.12)$$

Это уравнение содержит три неизвестных для $i=1, 3, \dots, n-1$. Решение (14.1) по такой схеме носит название «неявная схема» и имеет графическое изображение:



Осталось дополнить систему (14.12) первым (для $i=0$, т.е. $x=a$) и последним (для $i=N$, т.е. $x=b$) уравнениями. Можно использовать формулы (14.7), но это ухудшит точность решения:

$$\text{Для } x=a \text{ первое уравнение: } (h\varphi_2 - \varphi_1) y_0^1 + \varphi_1 y_1^1 = f_1(t_1) \cdot h.$$

$$\text{Для } x=b \text{ последнее уравнение: } -\varphi_4 y_{N-1}^1 + (h\varphi_5 + \varphi_4) y_N^1 = f_2(t_1) \cdot h. \quad (14.13)$$

Используя (14.8), (14.9), (14.10), мы получим второй порядок точности по h , но эта трёхточечная схема сделает матрицу системы не трёхдиагональной:

$$\begin{aligned} \text{Для } x=a: & \quad (2h\varphi_2 - 3\varphi_1) y_0 + 4\varphi_1 y_1 - \varphi_1 y_2 = 2hf_1(t). \\ \text{Для } x=b: & \quad \varphi_4 y_{N-2} - 4\varphi_4 y_{N-1} + (2h\varphi_5 + 3\varphi_4) y_{N-2} = 2hf_2(t). \end{aligned} \quad (14.14)$$

Исправить её не трёхдиагональность можно арифметическими операциями со строками: первой со второй и последней с предпоследней. После такого исправления матрица системы станет трёхдиагональной, но вероятно потеряет свойство диагонального преобладания, что для матриц больших порядков может привести к потерям точности при применении метода прогонки.

Для сохранения трёхдиагональности матрицы системы и второго порядка точности вычислений относительно h , разложим y_1^{k+1} в ряд Тейлора в окрестности точки $(t_{k+1}; x_0)$:

$$y_1^{k+1} = y_0^{k+1} + \frac{\partial y}{\partial x} \cdot h + \frac{\partial^2 y}{\partial x^2} \cdot \frac{h^2}{2} + O(h^2). \quad (14.15)$$

Сюда вместо y''_{xx} подставим его выражение из (14.1) и из полученного соотношения выразим y''_{xx} в точке $(t_{k+1}; x_0)$:

$$\begin{aligned} \frac{\partial y}{\partial x} \Big|_{x=0}^{k+1} &= \frac{2\alpha_1}{h(2\alpha_1 - \alpha_2 h)} (y_1^{k+1} - y_0^{k+1}) - \frac{h}{2\alpha_1 - \alpha_2 h} \cdot \frac{\partial y}{\partial t} \Big|_0^{k+1} + \frac{\alpha_3 h}{2\alpha_1 - \alpha_2 h} y_0^{k+1} + \\ &+ \frac{h}{2\alpha_1 - \alpha_2 h} f(t_{k+1}, x_0) + O(h^2). \end{aligned}$$

Учтём, что $\frac{\partial y}{\partial t} \Big|_0^{k+1} = \frac{(y_0^{k+1} - y_0^k)}{\tau} + O(\tau)$ и получим первое уравнение (для $x=x_0$) будущей трёхдиагональной системы в случае, когда $\varphi_1 \neq 0$:

$$\begin{aligned} \left(\frac{2\alpha_1}{h} + \frac{h}{\tau} - \alpha_3 h - \frac{\varphi_2}{\varphi_1} (2\alpha_1 - \alpha_2 h) \right) y_0^{k+1} - \frac{2\alpha_1}{h} y_1^{k+1} = \\ = \frac{h}{\tau} y_0^k + h \cdot f(t_{k+1}, x_0) - \frac{2\alpha_1 - \alpha_2 h}{\varphi_1} f_1(t_{k+1}). \end{aligned} \quad (14.16)$$

Если $\varphi_1=0$, то первое уравнение будет выглядеть: $y_0^{k+1} = f_1(t_{k+1})$.

Аналогично последнее уравнение (для $x=x_N$), будет для $\varphi_4=0$ выглядеть: $y_N^{k+1} = f_2(t_{k+1})$, а если $\varphi_4 \neq 0$ так:

$$\begin{aligned} -\frac{2\alpha_1}{h} y_{N-1}^{k+1} + \left(\frac{2\alpha_1}{h} + \frac{h}{\tau} - \alpha_3 h + \frac{\varphi_5}{\varphi_4} (2\alpha_1 + \alpha_2 h) \right) y_N^{k+1} = \\ = \frac{h}{\tau} y_N^k + h \cdot f(t_{k+1}, x_N) + \frac{2\alpha_1 + \alpha_2 h}{\varphi_4} f_2(t_{k+1}). \end{aligned} \quad (14.17)$$

Остальные $N-1$ уравнений для внутренних точек (для $i=1, 2, \dots, N-1$) записываются по формуле (14.12).

При программировании неявной схемы надо учесть, что на каждом новом слое приходится решать систему линейных алгебраических уравнений с трёхдиагональной матрицей. Недостаток неявной схемы в необходимости решения трёхдиагональной СЛАУ (например, методом прогонки). Это несколько усложняет программирование, увеличивает количество арифметических операций (а значит ухудшение точности). Но решение получается устойчивым по сравнению с явной схемой. Поэтому вычисления возможно проводить с большим шагом по t , а это, даже с учётом метода прогонки, существенно уменьшает общее время вычисления до $t=T$, а значит может уменьшаться общее количество арифметических операций.

Пример. В качестве примера логично рассмотреть решение той же задачи, что мы брали для явной схемы: решить для $t \in [0; 2]$, $x \in [0; \pi]$ уравнение $y'_t = y''_{xx}$ при $y(t, 0)=0$, $y(t, \pi)=0$, $y(0, x) = \sin x$.

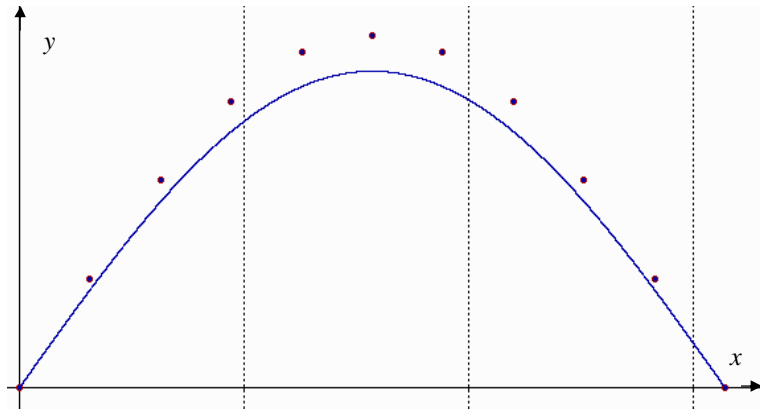
Решение. Наложим на область сетку с шагом по x : $h_x = \pi/10$ и по t : $\tau=0,1$. На нулевом слое при $t=0$ $y_{0j} = \sin x_j = \sin(j \cdot \pi/10)$. На новом (первом) слое известны (вычисляемы по формулам из условия) самое

левое y_{10} и самое правое y_{110} , которые вычисляются и равны 0. Чтобы найти остальные y_1, y_2, \dots, y_9 , надо составить и решить систему уравнений

$$\begin{cases} -2,98696 y_1 + y_2 = -0,30499 \\ y_1 - 2,98696 y_2 + y_3 = -0,58012 \\ \dots\dots\dots \\ y_8 - 2,98696 y_9 = -0,30499 \end{cases}$$

Граничные y_0, y_{10} равны 0. Эта система решается методом прогонки. Вычислив сначала прогоночные коэффициенты, затем вычисляем неизвестные y_1, y_2, \dots, y_9 . Это и есть значения неизвестной функции на первом слое (при $t=\tau=0,1$). Теперь, считая первый слой старым, вычислим (решив при этом свою систему) значения функции на новом втором слое (при $t=2\cdot\tau=0,2$). И так далее до последнего слоя при $t=T=2$.

Построим на одной координатной плоскости одиннадцать найденных точек $(x_i; y_i)$ и точное решение этой задачи ($y=e^{-t} \cdot \sin x$):



Максимальная относительная погрешность вычисленных в узлах y оказалась равной $\approx 11,5\%$.

Если решать эту задачу с шагом h по x равным $\pi/100$, то процесс вычислений будет устойчив для любого шага τ по t . С уменьшением τ будет уменьшаться и относительная погрешность ε вычисляемых y_i .

В следующей таблице представлены относительные погрешности вычисленных на последнем слое y_i для двух разных разбиений h по x и пяти разбиений τ по t (N_t = количество вычисляемых слоёв):

		$h=\pi/10$	$h=\pi/100$
τ	N_t	ε	ε
0,1	20	0,11	0,096
0,01	200	0,026	0,010
0,001	2000	0,017	0,0011

0,0001	20000	0,016	0,00026
0,00001	200000	0,016	0,00017

Из таблицы видно (интуитивно понятно и без таблицы), что получать более точные результаты надо уменьшая оба шага (и h , и τ). Эффективным явилось соотношение между шагами: $\tau^* \approx h^2/2$. Для явной схемы это соотношение являлось обязательным. В неявной схеме шаг τ может быть и больше чем τ^* . Но при $\tau < \tau^*$ улучшения точности расчётов не произошло.

14.3. Явно-неявная схема

Рассмотрим решение уравнения (14.1) используя обе схемы с весами.

Уравнение $\frac{\partial y}{\partial t} = \alpha_1 \cdot \frac{\partial^2 y}{\partial x^2} + \alpha_2 \frac{\partial y}{\partial x} + \alpha_3 y + f(t, x)$ запишется в виде:

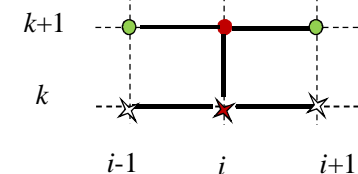
$$\frac{y_i^{k+1} - y_i^k}{\tau} = \theta \left(\alpha_1 \frac{y_{i+1}^{k+1} - 2y_i^{k+1} + y_{i-1}^{k+1}}{h^2} + \alpha_2 \frac{y_{i+1}^{k+1} - y_{i-1}^{k+1}}{2h} + \alpha_3 y_i^{k+1} + f(t_{k+1}, x_i) \right) + (1-\theta) \left(\alpha_1 \frac{y_{i+1}^k - 2y_i^k + y_{i-1}^k}{h^2} + \alpha_2 \frac{y_{i+1}^k - y_{i-1}^k}{2h} + \alpha_3 y_i^k + f(t_k, x_i) \right). \quad (14.14)$$

Здесь $0 \leq \theta \leq 1$, $i=1, 2, \dots, N-1$ (для внутренних точек отрезка $[a; b]$).

Неизвестными являются значения y_{i-1}^{k+1} , y_i^{k+1} , y_{i+1}^{k+1} на новом $k+1$ слое.

$$\left(2\alpha_1 - h \cdot \alpha_2 \right) \frac{\tau \theta}{2h^2} y_{i-1}^{k+1} + \left(\alpha_3 \theta \tau - \frac{2\alpha_1 \theta \tau}{h^2} - 1 \right) y_i^{k+1} + \left(2\alpha_1 + h \cdot \alpha_2 \right) \frac{\tau \theta}{2h^2} y_{i+1}^{k+1} = -\theta \tau \cdot f(t_{k+1}, x_i) - y_i^k - (1-\theta) \left(\alpha_1 \tau \frac{y_{i+1}^k - 2y_i^k + y_{i-1}^k}{h^2} + \alpha_2 \tau \frac{y_{i+1}^k - y_{i-1}^k}{2h} + \alpha_3 \tau \cdot y_i^k + \tau \cdot f(t_k, x_i) \right) \quad (14.15)$$

Т.к. $i=1, 2, \dots, N-1$, получаем $N-1$ уравнение с $N+1$ неизвестными $y_0^{k+1}, y_1^{k+1}, \dots, y_N^{k+1}$ на новом $k+1$ слое. На старом k слое все $y_0^k, y_1^k, \dots, y_N^k$ должны быть известны. Графически такая схема изображается так



Для получения ещё двух уравнений, описывающих краевые условия, воспользуемся формулами предыдущего параграфа:

- двухточечные формулы (14.13) первого порядка точности относительно h .

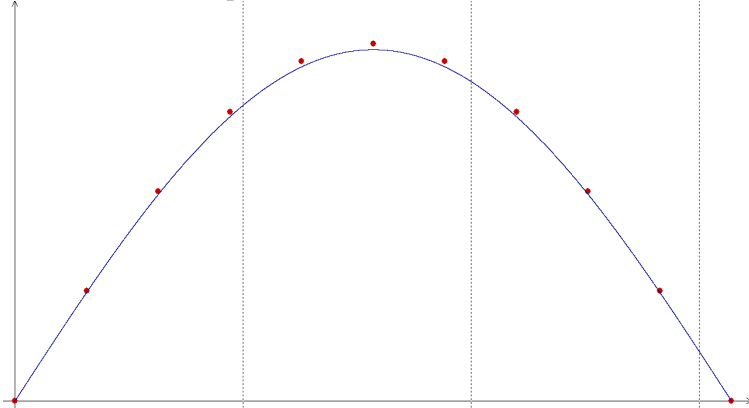
- трёхточечные формулы (14.14) второго порядка точности относительно h .
- двухточечные формулы (14.16) и (14.17) второго порядка точности относительно h .

При программировании явно-неявной схемы, как и для неявной на каждом новом слое приходится решать систему линейных алгебраических уравнений с трёхдиагональной матрицей (методом прогонки). Решение получается устойчивым по сравнению с явной схемой если $0,5 \leq \Theta \leq 1$. Для $\Theta=0,5$ схему (14.14) называют схемой Кранка-Никлсона. Как будет видно из примера, она точнее, чем явная и неявная схемы.

Пример. В качестве примера логично рассмотреть решение той же задачи, что мы брали для явной схемы: решить для $t \in [0;2]$, $x \in [0;\pi]$ уравнение $y'_t = y''_{xx}$ при $y(t,0)=0$, $y(t,\pi)=0$, $y(0,x) = \sin x$.

Решение. Схема Кранка-Никлсона ($\Theta=0,5$). Наложим на область сетку с шагом по x : $h_x = \pi/10$ и по t : $\tau=0,1$. На нулевом слое при $t=0$ $y_{0j} = \sin x_j = \sin(j \cdot \pi/10)$. На новом (первом) слое известны (вычисляются по формулам из условия) самое левое y_{10} и самое правое y_{110} , которые вычисляются и равны 0. Чтобы найти остальные y_1, y_2, \dots, y_9 , надо составить и решить систему уравнений, которая решается методом прогонки. Вычислив сначала прогоночные коэффициенты, затем вычисляем неизвестные y_1, y_2, \dots, y_9 (граничные y_0, y_{10} равны 0). Это и есть значения неизвестной функции на первом слое (при $t=\tau=0,1$). Решение аналогично неявной схеме.

Построим на одной координатной плоскости одиннадцать найденных точек $(x_i; y_i)$ и точное решение этой задачи ($y=e^{-t} \cdot \sin x$):



Решение ожидается точнее обеих предыдущих схем.

15. Решение уравнения гиперболического типа.

Рассмотрим решение дифференциального уравнения в частных производных гиперболического типа с начально-краевыми условиями:

$$\frac{\partial^2 u}{\partial t^2} + \alpha_1 \frac{\partial u}{\partial t} = \alpha_2 \frac{\partial^2 u}{\partial x^2} + \alpha_3 \frac{\partial u}{\partial x} + \alpha_4 u + f(t, x), \quad (15.1)$$

$$a \leq x \leq b, \quad 0 \leq t \leq T, \quad \alpha_2 > 0$$

$$\text{для } x=a \quad \varphi_1 \frac{\partial u}{\partial x} + \varphi_2 u = f_1(t), \quad (15.2)$$

$$\text{для } x=b \quad \varphi_4 \frac{\partial u}{\partial x} + \varphi_5 u = f_2(t), \quad (15.3)$$

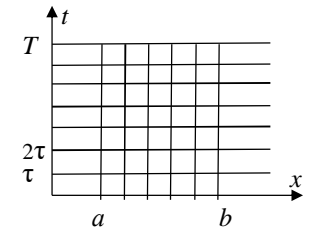
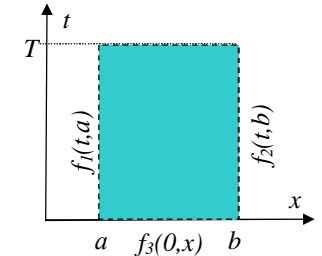
$$\text{для } t=0 \quad u(0, x) = f_3(x). \quad (15.4)$$

$$\text{для } t=0 \quad \left. \frac{\partial u}{\partial t} \right|_x = f_4(x) \quad (15.5)$$

Накроем область сеткой с шагом по x равным h и с шагом по t равным τ . Тогда $x_0=a$, $x_1=a+h$, ..., $x_N=b$, $h=(b-a)/N$, $t_0=0$, $t_1=\tau$, $t_2=2\tau$, ..., $t_M=T=M\tau$, $M=T/\tau$.

Геометрически область представляет собой «стакан», с трёх сторон которого заданы условия: начальные условия (15.4) и (15.5), краевые условия слева и справа заданы (15.2) и (15.3), а на верхней кромке (при $t=T$) значения функции $u(t, x)$ не известны. Их вычисление и является целью рассматриваемых алгоритмов. Коэффициенты α_i уравнения и φ_i в краевых условиях представляют собой константы (могут быть и равны 0).

Тогда $u(t_k, x_i) = u_i^k$. Назовём её сеточной функцией. Рассмотрим несколько способов решения этой задачи.



15.1. Явная схема

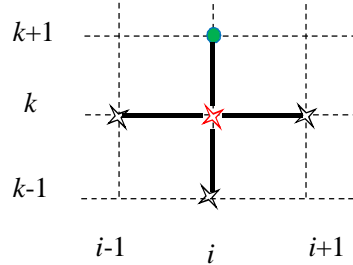
Для замены производных конечно-разностными соотношениями воспользуемся формулами главы 9 для $(t_k; x_i)$ внутри области:

$$\begin{aligned}\frac{\partial^2 u}{\partial t^2} &= \frac{u_i^{k+1} - 2u_i^k + u_i^{k-1}}{\tau^2} + O(\tau^2), & \frac{\partial u}{\partial t} &= \frac{u_i^{k+1} - u_i^{k-1}}{2\tau} + O(\tau^2), \\ \frac{\partial^2 u}{\partial x^2} &= \frac{u_{i+1}^k - 2u_i^k + u_{i-1}^k}{h^2} + O(h^2), & \frac{\partial u}{\partial x} &= \frac{u_{i+1}^k - u_{i-1}^k}{2h} + O(h^2)\end{aligned}\quad (15.6)$$

Для точки $(t_k; x_i)$: $a < x_i < b$, $t_k > 0$ получим уравнение

$$\begin{aligned}\frac{u_i^{k+1} - 2u_i^k + u_i^{k-1}}{\tau^2} + \alpha_1 \frac{u_i^{k+1} - u_i^{k-1}}{2\tau} = \alpha_2 \frac{u_{i+1}^k - 2u_i^k + u_{i-1}^k}{h^2} + \alpha_3 \frac{u_{i+1}^k - u_{i-1}^k}{2h} + \alpha_4 u_i^k + \\ + f(t_k, x_i) + O(\tau^2 + h^2)\end{aligned}$$

Такую конечно-разностную схему называют «крест» и изображают на сетке $x_0=a$, $x_1=a+h, \dots, x_N=b$, $h=(b-a)/N$, $\tau=T/M$, $t_0=0$, $t_1=\tau$, $t_2=2\tau, \dots, t_M=T$. Все производные уравнения (15.1) вычисляются в точке (t_k, x_i) . После подстановки соотношений (15.6) в (15.1) можно выразить u_i^{k+1} нового $k+1$ слоя в зависимости от значений $u_{i-1}^k, u_i^k, u_{i+1}^k, u_i^{k-1}$ на уже вычисленных слоях k и $k-1$.



$$\begin{aligned}u_i^{k+1} = \left(\frac{1}{2 + \alpha_1 \tau} \right) \left(4u_i^k + (\alpha_1 \tau - 2)u_i^{k-1} + \frac{2\alpha_2 \tau^2}{h^2} (u_{i-1}^k - 2u_i^k + u_{i+1}^k) + \frac{\alpha_3 \tau^2}{h} (u_{i+1}^k - u_{i-1}^k) + \right. \\ \left. + 2\tau^2 (\alpha_4 u_i^k + f(t_k, x_i)) \right) + O(\tau^2 + h^2)\end{aligned}\quad (15.7)$$

После вычисления u_i^{k+1} для $i=1, 2, \dots, N-1$ можно будет вычислить u_0^{k+1} и u_N^{k+1} . Аппроксимация граничных условий может быть первого порядка точности по h (но тогда уменьшается точность решения всей задачи):

$$\left. \frac{\partial u}{\partial x} \right|_{x=a}^{k+1} = \frac{u_1^{k+1} - u_0^{k+1}}{h} + O(h) \quad \text{и} \quad \left. \frac{\partial u}{\partial x} \right|_{x=b}^{k+1} = \frac{u_N^{k+1} - u_{N-1}^{k+1}}{h} + O(h)$$

подставим в (15.2) и (15.3) и тогда крайние u_0^{k+1} и u_N^{k+1} вычислим по формулам

$$u_0^{k+1} = \frac{\varphi_1}{\varphi_1 - \varphi_2 h} u_1^{k+1} + \frac{h}{\varphi_2 h - \varphi_1} f_1(t_{k+1}) + O(h), \quad (15.8)$$

$$u_N^{k+1} = \frac{\varphi_4}{\varphi_4 + \varphi_5 h} u_{N-1}^{k+1} + \frac{h}{\varphi_5 h + \varphi_4} f_2(t_{k+1}) + O(h). \quad (15.9)$$

Для более точных вычислений (порядка h^2) можно использовать формулы

$$\left. \frac{\partial u}{\partial x} \right|_{x=a}^{k+1} = \frac{-3u_0^{k+1} + 4u_1^{k+1} - u_2^{k+1}}{2h} + O(h^2) \quad \text{и} \quad \left. \frac{\partial u}{\partial x} \right|_{x=b}^{k+1} = \frac{u_{N-2}^{k+1} - 4u_{N-1}^{k+1} + 3u_N^{k+1}}{2h} + O(h^2),$$

тогда, подставив их в (15.2) и (15.3), вычислим крайние u_0^{k+1} и u_N^{k+1} по формулам

$$u_0^{k+1} = \frac{\varphi_1}{2\varphi_2 h - 3\varphi_1} (u_2^{k+1} - 4u_1^{k+1}) + \frac{2h}{2\varphi_2 h - 3\varphi_1} f_1(t_{k+1}) + O(h^2), \quad (15.10)$$

$$u_N^{k+1} = \frac{\varphi_4}{2\varphi_5 h + 3\varphi_4} (-u_{N-2}^{k+1} + 4u_{N-1}^{k+1}) + \frac{2h}{2\varphi_5 h + 3\varphi_4} f_2(t_{k+1}) + O(h^2). \quad (15.11)$$

Чтобы начать пользоваться формулами (15.7) надо знать $u_i^0 (i=1, \dots, N-1)$ на нулевом ($t=0$) слое и $u_i^1 (i=0, \dots, N)$ на первом слое ($t=t_1=\tau$).

На нулевом слое $u_i^0 = f_3(x_i) (i=0, \dots, N)$.

Для вычисления u_i^1 можно воспользоваться формулами *первого* порядка точности относительно τ :

$$\frac{\partial u}{\partial t} = \frac{u_i^1 - u_i^0}{\tau} + O(\tau), \quad \text{тогда} \quad u_i^1 = u_i^0 + \tau \cdot f_4(x_i) + O(\tau). \quad (15.12)$$

Для получения формул *второго* порядка точности относительно τ для вычисления u_i^1 , разложим его в ряд Тейлора в окрестности узла $(t_0; x_i)$:

$$u_i^1 = u_i^0 + \left. \frac{\partial u}{\partial t} \right|_i^0 \cdot \tau + \left. \frac{\partial^2 u}{\partial t^2} \right|_i^0 \cdot \frac{\tau^2}{2} + O(\tau^2). \quad \text{Подставим в него из (15.1) выражение для}$$

$$\left. \frac{\partial^2 u}{\partial t^2} \right|_i^0 = -\alpha_1 \left. \frac{\partial u}{\partial t} \right|_i^0 + \alpha_2 \left. \frac{\partial^2 u}{\partial x^2} \right|_i^0 + \alpha_3 \left. \frac{\partial u}{\partial x} \right|_i^0 + \alpha_4 u(0; x_i) + f(0; x_i) =$$

$$-\alpha_1 f_4(x_i) + \alpha_2 \frac{\partial^2 f_3}{\partial x^2}(x_i) +$$

$$+ \alpha_3 \frac{\partial f_3}{\partial x}(x_i) + \alpha_4 f_3(x_i) + f(0; x_i). \quad \text{Тогда получим формулу второго порядка}$$

точности относительно τ :

$$u_i^1 = f_3(x_i) + f_4(x_i) \cdot \tau + \left(-\alpha_1 f_4(x_i) + \alpha_2 \frac{\partial^2 f_3}{\partial x^2}(x_i) + \alpha_3 \frac{\partial f_3}{\partial x}(x_i) + \alpha_4 f_3(x_i) + f(0; x_i) \right) \frac{\tau^2}{2} + O(\tau^2) \quad (15.13)$$

Итак, алгоритм решения таков:

- 1) Вычисляем на нулевом слое $u_i^0 = f_3(x_i) (i=0, \dots, N)$ по (15.4).
- 2) По формулам (15.12) или (15.13) на первом слое вычисляем $u_i^1 (i=0, \dots, N)$.
- 3) По формулам (15.7) вычисляем $u_i^k (k=2, i=1, \dots, N-1)$.
- 4) По формулам (15.8) и (15.9) или (15.10) и (15.11) вычисляем u_0^k и $u_N^k (k=2)$.

Теперь $k=k+1$. Возвращаемся к пунктам 3) и 4) для вычисления u_i^k (для $k=3$) по уже известным слоям $k=1$ и 2 ($k-1$ и $k-2$). И так с помощью пунктов 3) и 4) вычислим все u_i^k на всей сетке до $t=T$.

15.2. Неявная схема

Рассмотрим вначале более простой и менее точный способ. Для аппроксимации производных будем использовать соотношения:

$$\frac{\partial^2 u}{\partial \tau^2} = \frac{u_i^k - 2u_i^{k-1} + u_i^{k-2}}{\tau^2} + O(\tau),$$

$\frac{\partial u}{\partial \tau} = \frac{u_i^k - u_i^{k-2}}{2\tau} + O(\tau)$. Порядок точности по τ первый, т.к. эти формулы центрально-симметричны для точки $(t_{k-1}; x_i)$, а мы будем применять в точке $(t_k; x_i)$.

Подставим $\frac{\partial u}{\partial x} = \frac{u_{i+1}^k - u_{i-1}^k}{2h} + O(h^2)$ и

$\frac{\partial^2 u}{\partial x^2} = \frac{u_{i+1}^k - 2u_i^k + u_{i-1}^k}{h^2} + O(h^2)$ в (15.1) и получим для точки $(t_k; x_i)$: $a < x_i < b$ уравнение:

$$\frac{u_i^k - 2u_i^{k-1} + u_i^{k-2}}{\tau^2} + \alpha_1 \frac{u_i^k - u_i^{k-2}}{2\tau} = \alpha_2 \frac{u_{i+1}^k - 2u_i^k + u_{i-1}^k}{h^2} + \alpha_3 \frac{u_{i+1}^k - u_{i-1}^k}{2h} + \alpha_4 u_i^k + O(\tau + h^2).$$

Умножив на $2h^2$, получим:

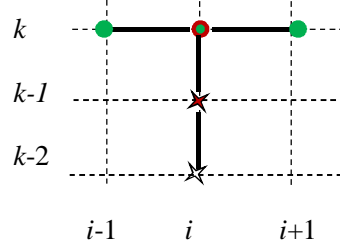
$$(2\alpha_2 - h\alpha_3)u_{i-1}^k + \left(2h^2\left(-\frac{\alpha_1}{2\tau} - \frac{1}{\tau^2} + \alpha_4\right) - 4\alpha_2\right)u_i^k + (h\alpha_3 + 2\alpha_2)u_{i+1}^k = -\frac{4h^2}{\tau^2}u_i^{k-1} + \left(\frac{2h^2}{\tau^2} - \frac{\alpha_1 h^2}{\tau}\right)u_i^{k-2} - 2h^2 f(t_k; x_i) + O(\tau + h^2), \quad (i=1, 2, \dots, N-1). \quad (15.14)$$

Для u_i^1 можно было бы воспользоваться и формулой $\frac{\partial u}{\partial t} = \frac{u_i^k - u_i^{k-1}}{\tau} + O(\tau)$, тогда вместо (15.14) получится:

$$(2\alpha_2 - h\alpha_3)u_{i-1}^k + \left(2h^2\left(-\frac{\alpha_1}{\tau} - \frac{1}{\tau^2} + \alpha_4\right) - 4\alpha_2\right)u_i^k + (h\alpha_3 + 2\alpha_2)u_{i+1}^k = -\frac{2\alpha_1 h^2}{\tau}u_i^{k-1} + \frac{2h^2}{\tau^2}(-2u_i^{k-1} + u_i^{k-2}) - 2h^2 f(t_k; x_i) + O(\tau + h^2), \quad (i=1, 2, \dots, N-1). \quad (15.15)$$

В системах (15.14) и (15.15) нет первого и последнего уравнений. Получим их.

Для вычисления u_0^k и u_N^k в крайних точках ($i=0$ и $i=N$) можно воспользоваться формулами первого порядка точности по h



$$\left. \frac{\partial u}{\partial x} \right|_{x=a}^k = \frac{u_1^k - u_0^k}{h} + O(h) \quad \text{и} \quad \left. \frac{\partial u}{\partial x} \right|_{x=b}^k = \frac{u_N^k - u_{N-1}^k}{h} + O(h), \quad \text{получим первое и}$$

последнее уравнения системы:

$$(\varphi_2 h - \varphi_1)u_0^k - \varphi_1 u_1^k = h \cdot f_1(t_k), \quad -\varphi_4 u_{N-1}^k + (\varphi_5 h + \varphi_4)u_N^k = h \cdot f_2(t_k)$$

Но эти два уравнения понизят порядок точности по h всей задачи до *первого*. Чтобы сохранить *второй* порядок точности по h воспользуемся формулами второго порядка точности по h :

$$\left. \frac{\partial u}{\partial x} \right|_{x=a}^k = \frac{-3u_0^k + 4u_1^k - u_2^k}{2h} + O(h^2) \quad \text{и} \quad \left. \frac{\partial u}{\partial x} \right|_{x=b}^k = \frac{u_{N-2}^k - 4u_{N-1}^k + 3u_N^k}{2h} + O(h^2) \quad (15.16)$$

И получим первое и последнее уравнения системы:

$$(2\varphi_2 h - 3\varphi_1)u_0^k + 4\varphi_1 u_1^k - \varphi_1 u_2^k = 2h \cdot f_1(t_k), \quad (15.17)$$

$$\varphi_4 u_{N-2}^k - 4\varphi_4 u_{N-1}^k + (2\varphi_5 h + 3\varphi_4)u_N^k = 2h \cdot f_2(t_k). \quad (15.18)$$

Но применение этих уравнений сделает матрицу системы не трёхдиагональной. Это можно исправить с помощью арифметических операций первой со второй и последней с предпоследней строками. Хотя после этого возможно будет потеряно свойство диагонального преобладания матрицы системы.

Алгоритм:

- 1) На нулевом слое $u_i^0 = f_3(x_i)$, ($i=0, \dots, N$).
- 2) Для вычисления u_i^1 можно воспользоваться формулами первого порядка точности относительно τ т.к. точность аппроксимации дифференциального уравнения по τ всё равно первый: $\frac{\partial u}{\partial t} = \frac{u_i^1 - u_i^0}{\tau} + O(\tau)$, тогда $u_i^1 = u_i^0 + \tau \cdot f_4(x_i)$, ($i=0, \dots, N$).
- 3) Все следующие слои вычисляются решением системы (15.17), (15.14), (15.18), или (15.17), (15.15), (15.18), вычисляя u_i^2 ($i=0, 1, \dots, N$), затем u_i^3 ($i=0, 1, \dots, N$). И так до u_i^M для $t_k = T$ ($k=M$).

Теперь рассмотрим *более точный и менее простой способ* с аппроксимацией второго порядка точности относительно τ и h .

Согласно пунктам 1)-4) предыдущего параграфа вычисляем u_i^0 , u_i^1 на слоях 0, 1, используя формулы численного дифференцирования *второго* порядка. На слое 2 вычисляем u_i^2 по явной схеме со вторым порядком точности по τ и h . Начиная со слоя 3, будем использовать неявную схему. Для этого конечно-разностные выражения для производных применяем в точке $(t_k; x_i)$:

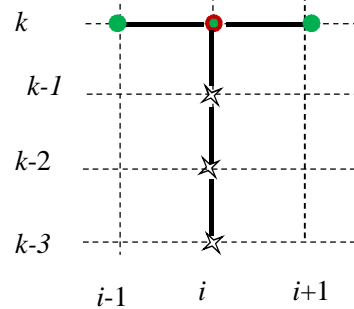
$$\frac{\partial^2 u}{\partial t^2} = \frac{2u_i^k - 5u_i^{k-1} + 4u_i^{k-2} - u_i^{k-3}}{\tau^2} + O(\tau^2);$$

$$\frac{\partial u}{\partial t} = \frac{3u_i^k - 4u_i^{k-1} + u_i^{k-2}}{2\tau} + O(\tau^2);$$

$$\frac{\partial^2 u}{\partial x^2} = \frac{u_{i+1}^k - 2u_i^k + u_{i-1}^k}{h^2} + O(h^2);$$

$$\frac{\partial u}{\partial x} = \frac{u_{i+1}^k - u_{i-1}^k}{2h} + O(h^2). \quad (15.18)$$

Графически неявную схему второго порядка точности относительно h и τ можно изобразить в виде «зонта»:



Итак, начиная со слоя $k=3$, используем неявную схему. В (15.1) подставим (15.18) и получим:

$$\begin{aligned} & (2\alpha_2 - \alpha_3 h)u_{i-1}^k + \left(2h^2\alpha_4 - \frac{4h^2}{\tau^2} - \frac{3h^2\alpha_1}{\tau} - 4\alpha_2 \right)u_i^k + \\ & + (2\alpha_2 + \alpha_3 h)u_{i+1}^k = + \frac{2h^2}{\tau^2}(-5u_i^{k-1} + 4u_i^{k-2} - u_i^{k-3}) + \\ & + \frac{h^2\alpha_1}{\tau}(-4u_i^{k-1} + u_i^{k-2}) - 2h^2 \cdot f((t_k; x_i) + O(\tau^2 + h^2)), \quad i=1, 2, \dots, N-1 \end{aligned} \quad (15.19)$$

Для вычисления u_0^k и u_N^k в крайних точках ($i=0$ и $i=N$) можно воспользоваться формулами второго порядка точности (15.15) И получим первое и последнее уравнения системы (15.16) и (15.17).

При этом матрица системы перестанет быть трёхдиагональной. Это можно исправить с помощью арифметических операций со строками: первой со второй и последней с предпоследней. Но после этого возможно будет потеряно свойство диагонального преобладания матрицы системы.

Есть способ добавления в систему (15.14) двух уравнений (первого и последнего), не нарушающих трёхдиагональность матрицы системы.

Разложим u_i^k в ряд Тейлора в окрестности точки $(t_k; x_0)$.

$$u_i^k = u_0^k + \frac{\partial u}{\partial x} \Big|_0^k \cdot h + \frac{\partial^2 u}{\partial x^2} \Big|_0^k \cdot \frac{h^2}{2} + O(h^2). \text{ Подставим в него из (15.1) выражение для}$$

$$\frac{\partial^2 u}{\partial x^2} \Big|_0^k = \frac{1}{\alpha_2} \left(\frac{\partial^2 u}{\partial t^2} \Big|_0^k + \alpha_1 \frac{\partial u}{\partial t} \Big|_0^k - \alpha_3 \frac{\partial u}{\partial x} \Big|_0^k - f(t_k, x_0) \right) + O(h^2). \text{ Используя (15.18) :}$$

$$\frac{\partial^2 u}{\partial t^2} \Big|_0^k = \frac{2u_0^k - 5u_0^{k-1} + 4u_0^{k-2} - u_0^{k-3}}{\tau^2} + O(\tau^2); \quad \frac{\partial u}{\partial t} \Big|_0^k = \frac{3u_0^k - 4u_0^{k-1} + u_0^{k-2}}{2\tau} + O(\tau^2),$$

получим

$$\begin{aligned} u_1^k = u_0^k + h \frac{\partial u}{\partial x} \Big|_0^k + \frac{h^2}{2\alpha_2} \left(\frac{1}{\tau^2} (2u_0^k - 5u_0^{k-1} + 4u_0^{k-2} - u_0^{k-3}) + \frac{\alpha_1}{2\tau} (3u_0^k - 4u_0^{k-1} + u_0^{k-2}) - \right. \\ \left. - \alpha_3 \frac{\partial u}{\partial x} \Big|_0^k - \alpha_4 u_0^k - f(t_k; x_0) \right) + O(h^2 + \tau^2). \end{aligned} \quad (15.19)$$

Выразим отсюда $\frac{\partial u}{\partial x} \Big|_0^k$ и подставим в краевое условие (15.2). Получим

$$\left\{ \varphi_2 + \frac{\varphi_1 2\alpha_2}{h(2\alpha_2 - h\alpha_3)} \left[-1 - \frac{h^2}{2\alpha_2} \left(\frac{2}{\tau^2} + \frac{3\alpha_1}{2\tau} - \alpha_4 \right) \right] \right\} u_0^k + \frac{\varphi_1 2\alpha_2}{h(2\alpha_2 - h\alpha_3)} u_1^k = f_1(t_k) + O(\tau^2 + h^2)$$

Это первое уравнение системы. Оно не нарушает трёхдиагональность матрицы системы.

Аналогично получается последнее уравнение системы для краевого условия (15.3).

Теперь можно применять трёхдиагональную прогонку.

Замечание. Применение более точной но трудоёмкой схемы может не оправдаться из-за роста погрешности при существенном увеличении количества арифметических операций.

16. Решение уравнения эллиптического типа.

Рассмотрим дифференциальное уравнение

$$\frac{\partial^2 u}{\partial x^2} + \alpha_1 \cdot \frac{\partial^2 u}{\partial y^2} + \alpha_2 \frac{\partial u}{\partial x} + \alpha_3 \frac{\partial u}{\partial y} + \alpha_4 u = f(x, y), \quad \alpha_1 > 0 \quad (16.1)$$

$a_x \leq x \leq b_x$, $a_y \leq y \leq b_y$, т.е. будем интегрировать уравнение в прямоугольной области. Краевые условия в общем виде выглядят так:

$$\text{для } x=a_x \quad \varphi_1 \frac{\partial u}{\partial x} + \varphi_2 u = f_1(y), \quad (16.2)$$

$$\text{для } x=b_x \quad \varphi_5 \frac{\partial u}{\partial x} + \varphi_6 u = f_3(y), \quad (16.3)$$

$$\text{для } y=a_y \quad \varphi_3 \frac{\partial u}{\partial y} + \varphi_4 u = f_2(x), \quad (16.4)$$

$$\text{для } y=b_y \quad \varphi_7 \frac{\partial u}{\partial y} + \varphi_8 u = f_4(x), \quad (16.5)$$

Наложим сетку на область интегрирования: $x_0=a_x$, $x_l=x_0+h_x$, $x_2=x_l+h_x$, ..., $x_N=b_x$; $y_0=a_y$, $y_l=y_0+h_y$, $y_2=y_l+h_y$, ..., $y_M=b_y$. Используем центрально-симметричную схему «крест». Тогда в каждой внутренней точке $(x_i; y_j)$ уравнение (16.1) примет конечно-разностный вид

$$\begin{aligned} & \frac{1}{h_x^2} (u_{i+1}^j - 2u_i^j + u_{i-1}^j) + \frac{\alpha_1}{h_y^2} (u_i^{j+1} - 2u_i^j + u_i^{j-1}) + \frac{\alpha_2}{2h_x} (u_{i+1}^j - u_{i-1}^j) + \\ & + \frac{\alpha_3}{2h_y} (u_i^{j+1} - u_i^{j-1}) + \alpha_4 u_i^j = f(x_i, y_j) + O(h_x^2 + h_y^2), \quad \alpha_1 > 0 \end{aligned} \quad (16.6)$$

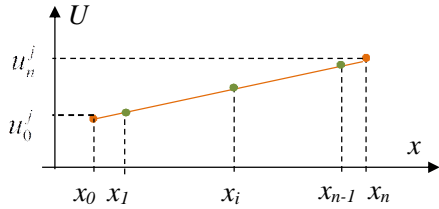
16.1 Краевые условия первого рода.

Рассмотрим случай, когда на всей границе задано краевое условие 1 рода ($\varphi_1 = \varphi_3 = \varphi_5 = \varphi_7 = 0$, $\varphi_2 = \varphi_4 = \varphi_6 = \varphi_8 = 1$). Применим разностно-итерационный метод Либмана.

Перед началом итераций построим стартовую (нулевую) сеточную поверхность $u(x_i, y_j)$.

На границах вычисляем $u(x_0, y_j)$, $u(x_n, y_j)$, $u(x_i, y_0)$, $u(x_i, y_m)$, по формулам

$$\begin{aligned} u(x_0, y_j) &= u_0^j = f_1(y_j), \quad u(x_n, y_j) = u_n^j = f_3(y_j), \quad \text{для } j=1, 2, \dots, m-1. \\ u(x_i, y_0) &= u_i^0 = f_2(x_i), \quad u(x_i, y_m) = u_i^m = f_4(x_i), \quad \text{для } i=0, 1, 2, \dots, n. \end{aligned} \quad (16.7)$$



Для вычисления $u(x_i, y_j)$ во внутренних точках применим линейную интерполяцию. Фиксируем y_j ($j=1, \dots, m-1$) и вычисляем u_i^j ($i=1, \dots, n-1$), по формуле:

$$u_i^j = u_0^j + (u_n^j - u_0^j) \cdot i \cdot h_x \quad (16.8)$$

Теперь на всей сетке найдены u_i^j (получена начальная поверхность). Из (16.6) выразим u_i^j , зависящую от соседних u_{i+1}^j , u_{i-1}^j , u_i^{j+1} и u_i^{j-1} . По этой формуле вычислим (уточним) $\hat{u}_i^j = u^1(x_i, y_j)$ на новой (первой итерационной) поверхности:

$$\begin{aligned} \hat{u}_i^j &= \frac{1}{\alpha_4 h_x^2 - 2 \left(\frac{h_y^2 + \alpha_1 h_x^2}{h_y^2} \right)} \left[h_x^2 \cdot f(x_i, y_j) - (u_{i+1}^j + u_{i-1}^j) - \frac{\alpha_1 h_x^2}{h_y^2} (u_i^{j+1} + u_i^{j-1}) - \right. \\ & \quad \left. - \frac{\alpha_2 h_x}{2} (u_{i+1}^j - u_{i-1}^j) - \frac{\alpha_3 h_x^2}{2 h_y} (u_i^{j+1} - u_i^{j-1}) \right] \quad (i=1, \dots, n-1, j=1, \dots, m-1) \end{aligned} \quad (16.9)$$

Значения $u^1(x_i, y_j) = \hat{u}_i^j$ на границе вычисляются аналогично формулам (16.7):

$$\begin{aligned} u^1(x_0, y_j) &= \hat{u}_0^j = f_1(y_j), \quad u^1(x_n, y_j) = \hat{u}_n^j = f_3(y_j), \quad \text{для } j=1, 2, \dots, m-1. \\ u^1(x_i, y_0) &= \hat{u}_i^0 = f_2(x_i), \quad u^1(x_i, y_m) = \hat{u}_i^m = f_4(x_i), \quad \text{для } i=0, 1, 2, \dots, n. \end{aligned} \quad (16.10)$$

Аналогично по формулам (16.9) и (16.10) вычисляем вторую итерационную поверхность $u^2(x_i, y_j)$ по уже известной первой $u^1(x_i, y_j)$.

Итерации можно прекращать, если $\max_{i,j} (|u^k(x_i, y_j) - u^{k-1}(x_i, y_j)|) < \varepsilon$, для заранее заданного $\varepsilon > 0$.

16.2 Смешанные краевые условия.

Рассмотрим уравнение (16.1) с краевыми условиями

$$\begin{aligned} & \text{для } x=a_x \quad u = f_1(y), \\ & \text{для } x=b_x \quad u = f_3(y), \\ & \text{для } y=a_y \quad \varphi_3 \frac{\partial u}{\partial y} + \varphi_4 u = f_2(x), \\ & \text{для } y=b_y \quad \varphi_7 \frac{\partial u}{\partial y} + \varphi_8 u = f_4(x), \end{aligned}$$

Накладываем на область $a_x \leq x \leq b_x$, $a_y \leq y \leq b_y$ ту же сетку $x_0=a_x$, $x_i = x_{i-1} + h_x$, $x_n=b_x$, $y_0=a_y$, $y_i = y_{i-1} + h_y$, $y_m=b_y$. На левой и правой границах вычислим

$$u_0^j = f_1(y_j) \quad \text{и} \quad u_n^j = f_3(y_j), \quad j=0, 1, 2, \dots, m. \quad (16.11)$$

Точки нулевой поверхности вычисляем также с помощью линейной интерполяции по (16.8).

Имея нулевую поверхность, по формулам (16.9) вычисляем первую итерационную поверхность $u^1(x_i, y_j) = \hat{u}_i^j$ во внутренних точках. Для граничных точек пользуемся формулами (16.11) и (16.12) или (16.12.1).

Нижние и верхние точки поверхности вычислим, используя (16.12).

$$\begin{aligned} \frac{\partial u}{\partial y} &= \frac{u_i^1 - u_i^0}{h_y} + O(h_y) \quad \text{для } y=a_y \quad \text{и} \quad \frac{\partial u}{\partial y} = \frac{u_i^m - u_i^{m-1}}{h_y} + O(h_y) \quad \text{для } y=b_y \quad \text{и} \quad i=1, \dots, n-1: \\ u_i^0 &= \frac{1}{\varphi_3 - h_y \varphi_4} (\varphi_3 \cdot u_i^1 - h_y \cdot f_2(x_i)), \quad u_i^m = \frac{1}{\varphi_7 + h_y \varphi_8} (\varphi_7 \cdot u_i^{m-1} - h_y \cdot f_4(x_i)). \end{aligned} \quad (16.12)$$

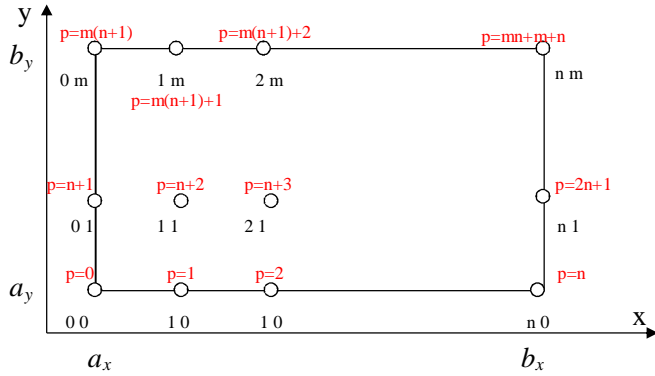
А можно использовать трёхточечные формулы второго порядка точности

$$\text{Для } y=a_y: \quad \frac{\partial u}{\partial y} = \frac{-3u_i^0 + 4u_i^1 - u_i^2}{2h_y} + O(h_y^2). \quad (16.12.1)$$

$$\text{И для } y=b_y: \quad \frac{\partial u}{\partial y} = \frac{u_i^{m-2} - 4u_i^{m-1} + 3u_i^m}{2h_y} + O(h_y^2).$$

Затем вычисляется вторая итерационная поверхность $u^2(x_i, y_j)$ по формулам (16.9) и т.д. до выполнения условий прекращения итераций.

16.3 Краевые условия второго и третьего рода.



Пронумеруем все точки прямоугольной сетки, например, так: $v_p = u_i^j$, где $p = i + (n+1)j$. Тогда $p = 0, 1, \dots, n + m + nm$. Для каждой внутренней точки, а их $(n-1)(m-1)$, запишем (16.6), приведя подобные члены для пяти

неизвестных $u_i^j, u_{i-1}^j, u_{i+1}^j, u_i^{j-1}, u_i^{j+1}$ и введя новую нумерацию:

$$\left(\frac{\alpha_1 h_x^2}{h_y^2} - \frac{\alpha_3 h_x^2}{2h_y} \right) v_{i+(n+1)(j-1)} + \left(1 - \frac{\alpha_2 h_x}{2} \right) v_{i-1+(n+1)j} + \left(h_x^2 \alpha_4 - 2 - \frac{2\alpha_1 h_x^2}{h_y^2} \right) v_{i+(n+1)j} + \left(1 + \frac{\alpha_2 h_x}{2} \right) v_{i+1+(n+1)j} + \left(\frac{\alpha_1 h_x^2}{h_y^2} + \frac{\alpha_3 h_x^2}{2h_y} \right) v_{i+(n+1)(j+1)} = h_x^2 f(x_i, y_j) \quad (16.13)$$

Краевые условия (16.2)-(16.5) дадут ещё $2(n+m)+2$ уравнений и в каждом по 2 неизвестных (или по 3 в случае трёхточечной аппроксимации производных в граничных условиях).

Решив систему из $nm+n+m+1$ уравнений с разреженной матрицей, перейдём с исходной нумерации узлов.

17. Многомерное уравнение параболического типа.

Рассмотрим задачу:

$$\frac{\partial u}{\partial t} = \alpha_1 \cdot \frac{\partial^2 u}{\partial x^2} + \alpha_2 \frac{\partial^2 u}{\partial y^2} + \alpha_3 \frac{\partial u}{\partial x} + \alpha_4 \frac{\partial u}{\partial y} + \alpha_5 u + f(x, y, t). \quad (17.1)$$

Для $a_x \leq x \leq b_x, a_y \leq y \leq b_y, 0 \leq t \leq T_{end}, h_x = (b_x - a_x)/N_x, h_y = (b_y - a_y)/N_y, \tau = T_{end}/N_t, u(x, y, 0) = f_1(x, y), u(a_x, y, t) = f_2(y, t), u(b_x, y, t) = f_3(y, t), u(x, a_y, t) = f_4(x, t), u(x, b_y, t) = f_5(x, t). \alpha_1 > 0, \alpha_2 > 0.$ (17.2)

17.1 Метод переменных направлений.

Полный шаг τ по времени разделим пополам. С первым полушагом $\tau/2$ будем аппроксимировать u'_x и u''_{xx} не явно, а u'_y и u''_{yy} - явно:

$$\frac{u_{i,j}^{k+1/2} - u_{i,j}^k}{\tau/2} = \alpha_1 \frac{u_{i+1,j}^{k+1/2} - 2u_{i,j}^{k+1/2} + u_{i-1,j}^{k+1/2}}{h_x^2} + \alpha_2 \frac{u_{i,j+1}^k - 2u_{i,j}^k + u_{i,j-1}^k}{h_y^2} + \alpha_3 \frac{u_{i+1,j}^{k+1/2} - u_{i-1,j}^{k+1/2}}{2h_x} + \alpha_4 \frac{u_{i,j+1}^k - u_{i,j-1}^k}{2h_y} + \alpha_5 u_{i,j}^{k+1/2} + f(x_i, y_j, t_{k+1/2}) + O(\tau^2 + h_x^2 + h_y^2) \quad (17.3)$$

При фиксированном y_j получим трёхдиагональную систему. Первое и последнее уравнение из краевых условий запишутся:

$$u_{0,j}^{k+1/2} = f_2(y_j, t_{k+1/2}) \text{ и } u_{N_x,j}^{k+1/2} = f_3(y_j, t_{k+1/2}). \quad (17.4)$$

Остальные $N_x - 1$ уравнения получаются из (17.3): для $i = 1, 2, 3, \dots, N_x - 1$

$$\left(\alpha_1 - \frac{h_x \alpha_3}{2} \right) u_{i-1,j}^{k+1/2} + \left(h_x^2 - \frac{2h_x^2}{\tau} - 2\alpha_1 \right) u_{i,j}^{k+1/2} + \left(\alpha_1 + \frac{h_x \alpha_3}{2} \right) u_{i+1,j}^{k+1/2} = -\frac{2h_x^2}{\tau} u_{i,j}^k - \frac{\alpha_2 h_x^2}{h_y^2} (u_{i,j+1}^k - 2u_{i,j}^k + u_{i,j-1}^k) - \frac{\alpha_4 h_x^2}{2h_y^2} (u_{i,j+1}^k - u_{i,j-1}^k) - h_x^2 f(x_i, y_j, t_{k+1/2}) \quad (17.5)$$

С помощью метода прогонки вычислим все $u_{i,j}^{k+1/2}$. Так применяем прогонку для каждого y_j ($j = 1, 2, \dots, N_y - 1$), т.е. $N_y - 1$ раз.

На границах используем (17.2):

$$u_{0,j}^{k+1/2} = f_2(y_j, t_{k+1/2}), u_{N_x,j}^{k+1/2} = f_3(y_j, t_{k+1/2}), u_{i,0}^{k+1/2} = f_4(x_i, t_{k+1/2}), u_{i,N_y}^{k+1/2} = f_5(x_i, t_{k+1/2})$$

После этого вся промежуточная поверхность $u^{k+1/2}$ вычислена.

На втором полушаге $\tau/2$ будем аппроксимировать u'_x и u''_{xx} явно, а u'_y и u''_{yy} - не явно:

$$\frac{u_{i,j}^{k+1} - u_{i,j}^{k+1/2}}{\tau/2} = \alpha_1 \frac{u_{i+1,j}^{k+1/2} - 2u_{i,j}^{k+1/2} + u_{i-1,j}^{k+1/2}}{h_x^2} + \alpha_2 \frac{u_{i,j+1}^{k+1} - 2u_{i,j}^{k+1} + u_{i,j-1}^{k+1}}{h_y^2} + \alpha_3 \frac{u_{i+1,j}^{k+1/2} - u_{i-1,j}^{k+1/2}}{2h_x} + \alpha_4 \frac{u_{i,j+1}^{k+1} - u_{i,j-1}^{k+1}}{2h_y} + \alpha_5 u_{i,j}^{k+1/2} + f(x_i, y_j, t_{k+1/2}) + O(\tau^2 + h_x^2 + h_y^2) \quad (17.6)$$

Теперь при каждом фиксированном x_i получим трёхдиагональную систему. Таких систем будет N_x-1 . Их решения вместе с краевыми условиями даст все точки новой поверхности u_{ij}^{k+1} для $t=t_{k+1}$.

17.2 Метод дробных шагов.

Для решения (17.1) и (17.2) будем дробить шаг τ по t на два. На первом шаге с помощью N_y-1 прогонок и краевых условий вычислим промежуточную сеточную функцию v_{ij} , а на втором - u_{ij}^{k+1} . На обоих этих шагах будем использовать неявную аппроксимацию, на одном по x , на другом по y :

$$\frac{v_{ij} - u_{ij}^k}{\tau} = \alpha_1 \frac{v_{i+1j} - 2v_{ij} + v_{i-1j}}{h_x^2} + \alpha_3 \frac{v_{i+1j} - v_{i-1j}}{2h_x} \quad (17.7)$$

$$\frac{u_{ij}^{k+1} - v_{ij}}{\tau} = \alpha_2 \frac{v_{ij+1} - 2v_{ij} + v_{ij-1}}{h_y^2} + \alpha_4 \frac{v_{ij+1} - v_{ij-1}}{2h_y} \quad (17.8)$$

$$\frac{u_{ij}^{k+1/2} - u_{ij}^k}{\tau} = \alpha_1 \frac{u_{i+1j}^{k+1/2} - 2u_{ij}^{k+1/2} + u_{i-1j}^{k+1/2}}{h_x^2} + \alpha_3 \frac{u_{i+1j}^{k+1/2} - u_{i-1j}^{k+1/2}}{2h_x} \quad (17.7)$$

$$\frac{u_{ij}^{k+1} - u_{ij}^{k+1/2}}{\tau} = \alpha_2 \frac{u_{ij+1}^{k+1} - 2u_{ij}^{k+1} + u_{ij-1}^{k+1}}{h_y^2} + \alpha_4 \frac{u_{ij+1}^{k+1} - u_{ij-1}^{k+1}}{2h_y} \quad (17.8)$$

Примеры заданий для расчётно-графических работ

N_Γ = номер группы (1,2,3,...), N_c = номер студента (от 1 до 40)

1. Методом Гаусса вычислить определитель, обратную матрицу и решение СЛАУ $A\bar{x} = \bar{b}$:

$$\text{а) } A = \begin{pmatrix} N_c & 5 & 2 \\ 5 & N_c & -N_\Gamma \\ 2 & -N_\Gamma & N_c \end{pmatrix}, \quad \bar{b} = \begin{pmatrix} 1 \\ N_c \\ -N_\Gamma \end{pmatrix};$$

$$\text{б) } A = \begin{pmatrix} N_c+5 & 0 & N_\Gamma & N_\Gamma \\ 3 & 3 & 0 & N_\Gamma \\ 2 & 5 & 3 & N_c \\ N_c & N_c+N_\Gamma & N_c & 1 \end{pmatrix}, \quad \bar{b} = \begin{pmatrix} 2 \\ 2 \\ 2 \\ N_c \end{pmatrix}.$$

2. Методом прогонки вычислить решение СЛАУ с расширенной матрицей порядка а) $n=5$; б) $n=10$:

$$\left(\begin{array}{cccccc|cccc|c} b_1 & c_1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & d_1 \\ a_2 & b_2 & c_2 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & d_2 \\ 0 & a_3 & b_3 & c_3 & 0 & \dots & 0 & 0 & 0 & 0 & d_3 \\ 0 & 0 & a_4 & b_4 & c_4 & \dots & 0 & 0 & 0 & 0 & d_4 \\ 0 & 0 & 0 & a_5 & b_5 & \dots & 0 & 0 & 0 & 0 & d_5 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & a_{n-2} & b_{n-2} & c_{n-2} & 0 & d_{n-2} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & a_{n-1} & b_{n-1} & c_{n-1} & d_{n-1} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & a_n & b_n & d_{n-1} \end{array} \right);$$

$a_i = i \cdot N_c + N_\Gamma$, $b_i = N_c \cdot i \cdot i + N_\Gamma$, $c_i = N_\Gamma - N_c \cdot i$, $d_i = N_c + N_\Gamma \cdot i$.

3. Методом простых итераций и методом Зейделя решить СЛАУ $B\bar{x} = \bar{c}$:

$$\text{а) } B = \begin{pmatrix} N_c+10 & N_\Gamma & 1 \\ N_\Gamma & N_c+10 & 3 \\ 1 & 3 & N_\Gamma+4 \end{pmatrix}, \quad \bar{c} = \begin{pmatrix} N_\Gamma \\ N_c+10 \\ 0 \end{pmatrix};$$

$$\text{б) } B = \begin{pmatrix} N_c+10 & N_\Gamma & 1 & 1 \\ N_\Gamma & N_c+10 & 2 & 2 \\ 1 & 2 & N_\Gamma+4 & 1 \\ 1 & 1 & 1 & N_c+N_\Gamma \end{pmatrix}, \quad \bar{c} = \begin{pmatrix} N_\Gamma \\ N_c+10 \\ 0 \\ 1 \end{pmatrix}.$$

4. Степенным методом вычислить спектральный радиус и ему соответствующий собственный вектор; методом вращений Якоби найти

все собственные числа и собственные векторы матриц B из предыдущего задания.

5. С заданной точностью $\varepsilon=10^{-3}$:

а) методом дихотомии найти отрицательный корень уравнения

$$\left(\frac{x}{Nc}\right)^3 - \frac{x}{Nz} + \frac{x}{Nz} = 0 \quad ;$$

б) методом итераций решить уравнение

$$\frac{x^3}{Nc \cdot Nz} + 2 \cdot x - (Nc + Nz) = 0 \quad ;$$

в) методом Ньютона (касательных) найти корень уравнения

$$\frac{x^3}{10} + \frac{x}{Nc} + Nc \cdot Nz = 0 \quad ;$$

г) методом хорд (секущих) найти корень уравнения

$$\frac{x^3}{5} + \frac{x}{Nc} - Nc \cdot Nz = 0 \quad ;$$

6. Методом линеаризации (Ньютона) найти с заданной точностью $\varepsilon=10^{-3}$ все решения системы

$$\begin{cases} Nz \cdot (x^3 - y^3) - Nc = 0 \\ x \cdot y - y - Nc \cdot Nz = 0 \end{cases} .$$

7. По заданной таблице построить интерполяционный многочлен

X	-2	-1	0	1	2
Y	Nc	Nr	-1	Nc	Nr

а) по первым четырём узлам;

б) по последним четырём узлам;

в) по всем пяти узлам.

8. По заданной таблице методом наименьших квадратов построить многочлены первой и второй степени, аппроксимирующие таблицу

X	-2,5·a	-1,5·a	-0,5·a	0,5·a	1,5·a
Y	84-Nr	73-Nr·2	63-Nr·3	55-Nr·4	47-Nr·5

$$a = Nc + Nr.$$

9. Аппроксимировать табличную функцию кубическими сплайнами. Переменная x изменяется на отрезке $[a; b]$. Этот отрезок разбивается на N частей. Значения табличной функции на отрезке $[a; b]$ вычислять по формуле $f(x) = Nr \cdot x + \sin(x)$.

$a = Nc$; $b = Nc + N$; $N = 4$. Сделать проверку: вычислить значения аппроксимируемой функции и сплайна в узловых точках и посередине между узловыми точками. Нарисовать сплайн и табличную функцию.

10. а) Вычислить значение определённого интеграла по формуле трапеций с числом узлов 11, 21 и 41 (шагами 0,1; 0,05 и 0,025)

$$\int_0^1 (x)^{Nr} \cdot (\ln(x))^{Nc} dx = ?$$

Результат уточнить по формуле Рунге.

б) Вычислить значение определённого интеграла по формуле Симпсона с числом узлов 11 и 21. Результат уточнить по формуле Рунге–Ромберга.

$$\int_0^{\frac{\pi}{2}} (\sin(x))^{2Nr} \cdot (\cos(x))^{2Nc} dx = ?$$

11. Решить задачу Коши: $Y'' + Y = a \cdot x$,

$$Y(0)=0; \quad Y'(0)=a+b; \quad 0 \leq x \leq \pi; \quad b = Nr; \quad a = \sqrt{Nc};$$

а) методом Эйлера с числом узлов 5, 10, 15; уточнить последнее значение Y по формуле Рунге;

б) методом Рунге–Кутты с числом узлов 5 и 10; уточнить последнее значение Y по формуле Рунге–Ромберга.

12. Решить краевую задачу

$$Y'' + Y' + Y = a \cdot (x + 1) + b \cos(x);$$

$$Y(0) = 0; \quad Y(\pi) = a \cdot \pi; \quad 0 \leq x \leq \pi; \quad b = Nr, \quad a = Nc$$

методом конечных разностей и методом прогонки, разбив отрезок $[a; b]$ на N отрезков одинаковой длины для:

а) $N = 4$, б) $N = 8$, в) $N = 16$.

13. Методом сеток решить уравнение теплопроводности

$$\frac{\partial y}{\partial t} = \beta^2 \cdot \frac{\partial^2 y}{\partial x^2} + f(t, x),$$

$$x \in [a; b], \quad t \in [0; 3/40], \quad y(t; a) = t \cdot Nc, \quad y(t; b) = t^2 \cdot Nr, \quad y(0; x) = \sin(x), \\ \beta = Nr + 1/Nc, \quad f(t; x) = (x - Nr) \cdot (t - Nc), \quad a = 0, \quad b = \pi, \quad \text{шаг по } x: h_x = \pi/4, \quad \text{шаг по } t: h_t = 1/40$$

а) по явной схеме,

б) по неявной схеме,

в) по явной- неявной схеме.

Библиографический список

1. Численные методы и алгоритмы решения задач ракетной техники: Учеб. пособие / Ю.Н. Бадлевский, П.В.Бирюков, Т.Я.Кошелева и др. Под ред. Т.Я. Кошелевой. – М.: Изд-во МАИ, 1990.
2. Архангельский Н.А., Чернов Н.И. Интерполяция и её приложения в примерах и задачах. – М.: Изд-во МАИ, 1976.
3. Бахвалов Н.С. Численные методы. – М.: Наука, 1975.
4. Бахвалов Н.С., Жидков Н.П., Кобельков А.А. Численные методы. – М.: Наука, 1987.
5. Березин И.С., Жидков Н.П. Методы вычислений. – М. Наука, 1962.
6. Бабенко К.И. Основы численного анализа. – М.: Наука, 1986.
7. Бут Э.Д. Численные методы. – М.: Государственное издательство физико-математической литературы, 1959.
8. Воеводин В.В., Кузнецов Ю.А. Матрицы и вычисления. – М.: Наука, 1984.
9. Воеводин В.В. Вычислительные основы линейной алгебры. – М.: Наука, 1977.
10. Гантмахер Ф.Р. Теория матриц. – М.: Наука, 1988.
11. Годунов С.К., Антонов А.Г., Кирилук О.П., Костин В.И. Гарантированная точность решения систем линейных уравнений в евклидовых пространствах. – Новосибирск.: Наука, 1988.
12. Гутер Р.С., Овчинский Б.В. Элементы численного анализа и математической обработки результатов опыта. – М.: Наука, 1970.
13. Данилина Н.И., Дубровская Н.С., Кваша О.П., Смирнов Г.Л. Вычислительная математика. М.: Высшая школа, 1985.
14. Демидович Б.П., Марон И.А. Основы вычислительной математики. – М.: Наука, 1970.
15. Джонсон К. Численные методы в химии. Пер. с английского. – М.: Наука, 1983.
16. Джордж А. Численное решение больших разреженных систем уравнений. – М.: Мир, 1984.
17. Ильин В.П., Кузнецов Ю.И. Трёхдиагональные матрицы и их приложения. – М.: Наука, 1985.
18. Икрамов Х.Д. Численное решение матричных уравнений. – М.: Наука, 1984.
19. Калиткин Н.Н. Численные методы. – М.: Наука, 1978.
20. Киреев В.И., Пантелеев А.В. Численные методы в примерах и задачах: Учеб. пособие. – М.: Высш. шк., 2004.
21. Копчёнова Н.В., Марон И.А. Вычислительная математика в примерах и задачах. – М.: Наука, 1972.

22. Крылов В.И., Бобков В.В., Монастырский П.В. Вычислительные методы. – М.: Наука, 1977.
23. Ланкастер П. Теория матриц. – М.: Наука, 1982.
24. Лоусон Ч., Хенсон Р. Численное решение задач методом наименьших квадратов. – М.: Наука, 1986.
25. Ланцош К. Практические методы прикладного анализа. – М.: Наука, 1961.
26. Марчук Г.И. Методы вычислительной математики. – М.: Наука, 1989.
27. Мудров А.Е. Численные методы для ПЭВМ на языках бейсик, фортран и паскаль. – Томск: МП «РАСКО», 1991.
28. Парлетт Б. Симметричная проблема собственных значений. Численные методы. – М.: Мир, 1983.
29. Пирумов У.Г. Численные методы: Учеб. пособие для студ. вузов. – М.: Дрофа, 2003.
30. Райс Дж. Матричные вычисления и математическое обеспечение. – М.: Мир, 1984.
31. Ракитин В.И., Первушин В.Е. Практическое руководство по методам вычислений с приложением программ для персональных компьютеров: Учеб. пособие. – М.: Высш. шк., 1998.
32. Самарский А.А., Гулин А.В. Численные методы. – М.: Наука, 1989.
33. Самохин А.Б., Самохина А.С. Фортран и вычислительные методы. – М.: Русина, 1994.
34. Уилкинсон Дж.Х. Алгебраическая проблема собственных значений. – М.: Наука, 1970.
35. Уилкинсон, Райнш. Справочник алгоритмов на языке АЛГОЛ. Линейная алгебра. Пер. с английского. – М.: Машиностроение, 1976.
36. Формалёв В.Ф., Ревизников Д.Л. Численные методы. – М.: Физматлит, 2004.
37. Форсайт Дж., Малькольм М., Моулер К. Машинные методы математических вычислений. Пер с английского. М.: Мир, 1980.
38. Хор Р., Джонсон Ч. Матричный анализ. М.: Мир, 1989.
39. Эстербю О., Златев З. Прямые методы для разреженных матриц. – М.: Мир, 1987.
40. Хемминг Р.В. Численные методы. – М.: Наука, 1972.
41. Черкасов М.А. Практический курс программирования на паскале: Учеб. пособие. – М.: Изд-во МАИ, 2005.
42. Черкасов М.А. Методы программирования на паскале в примерах и задачах: Учеб. пособие. – М.: Изд-во МАИ, 2005.
43. Черкасов М.А. Практический курс программирования на паскаль-АВС: Учеб. пособие. – М.: Изд-во Буки-Веди, 2015с.: ил.

Оглавление

Предисловие. О погрешностях.....	4
1. Прямые методы решения первой основной задачи линейной алгебры	7
1.1. Метод Гаусса решения СЛАУ	7
1.2. Вычисление обратной матрицы методом Гаусса	11
1.3. <i>UL</i> разложение Дулитла	13
1.4. <i>LU</i> разложение Краута	15
1.5. Метод Холецкого (квадратного корня) для симметричных матриц.....	16
1.6. Решение трёхдиагональной СЛАУ методом прогонки	18
2. Итерационные методы решения СЛАУ	21
2.1. Метод простых итераций решения СЛАУ	22
2.2. Метод Зейделя решения СЛАУ	26
3. Решение второй задачи линейной алгебры.....	28
3.1. Степенной метод вычисления спектрального радиуса.....	28
3.2. Метод вращений Якоби для симметричной матрицы	33
4. Методы решения трансцендентных уравнений	38
4.1. Метод половинного деления (дихотомии, бисекций)	39
4.2. Метод хорд (секущих)	41
4.3. Метод касательных (Ньютона)	43
4.4. Метод итераций	44
4.5. Метод парабол (Мюллера)	45
5. Решение систем нелинейных уравнений.....	47
5.1. Метод Ньютона (линеаризации)	47
5.2. Метод простой итерации	49
5.3. Метод Зейделя.....	51
6. Интерполяция.....	53
7. Аппроксимация методом наименьших квадратов	57
7.1. Квадратичная аппроксимация	57
7.2. Линейная аппроксимация	61
8. Интерполяция сплайнами	63
9. Численное дифференцирование	69
10. Численное интегрирование	70
10.1. Формула трапеций	70
10.2. Формула Симпсона	71
11. Формулы Рунге и Рунге–Ромберга.....	72
12. Решение задачи Коши для ОДУ.....	74
12.1. Решение задачи Коши для ОДУ первого порядка	76

12.1.1. Метод Эйлера для ОДУ первого порядка.....	76
12.1.2. Метод Рунге–Кутты для ОДУ первого порядка.....	77
12.2. Решение задачи Коши для ОДУ второго порядка	78
12.2.1. Метод Эйлера для ОДУ второго порядка.....	79
12.2.2. Метод Рунге–Кутты для ОДУ второго порядка.....	82
13. Решение краевой задачи для ОДУ	84
13.1. Метод конечных разностей.....	84
13.2. Метод стрельбы	88
14. Решение уравнения теплопроводности. Метод сеток.....	91
14.1. Явная схема	91
14.2. Неявная схема	94
14.3. Явно-неявная схема	98
15. Решение уравнения гиперболического типа.	100
15.1. Явная схема.....	100
15.2. Неявная схема.....	103
16. Решение уравнения эллиптического типа.	106
16.1 Краевые условия первого рода.	107
16.2 Смешанные краевые условия.	108
16.3 Краевые условия второго и третьего рода.....	109
17. Многомерное уравнение параболического типа.....	110
17.1 Метод переменных направлений.....	110
17.2 Метод дробных шагов.....	111
Библиографический список	115
Оглавление.....	117

Тем. План 2007 поз 38

Черкасов Михаил Анатольевич

Численные методы. Решение задач