

ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ И МЕТОДЫ ВЫЧИСЛЕНИЙ

Крыловецкий Александр Абрамович
каф. цифровых технологий

Литература

1. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы
2. Волков Е.А. Численные методы
3. Самарский А.А. Введение в численные методы
4. Березин И.С., Жидков Н.П. Методы вычислений
5. Сборник задач по методам вычислений. Под ред. П.И. Монастырского
6. Дьяконов В.П. Maple 10/9.5 в математике, физике, образовании

1 Введение в численные методы

◆ **Численные методы** – методы приближенного решения задач прикладной математики, основанные на реализации алгоритмов, соответствующих определенным математическим моделям. Раздел математики, изучающий численные методы, называется численным анализом, или вычислительной математикой.

Вычислительные методы начали развиваться еще в античном мире. Так, например, по итерационной формуле Герона тогда умели вычислять квадратный корень:

$$x_{n+1} = \frac{1}{2} \left(x_n + \frac{a}{x_n} \right), \quad \text{причем} \quad \lim_{n \rightarrow \infty} x_n = \sqrt{a},$$

x_1 – любое положительное число.

Развитие численных методов связано с именами **Л. Эйлера, И. Ньютона, О.Л. Коши, Ж.Л. Лагранжа, А.М. Лежандра, П.С. Лапласа, А. Пуанкаре, П.Л. Чебышева, Карла Рунге** и многих других известных математиков.

В наше время импульсом к развитию вычислительной математики стали задачи ядерной физики, космонавтики, динамики атмосферы, термогидрографии, физики плазмы, синергетики, биомеханики, теории оптимизации, экономики и др.

◆ **Специфические особенности вычислительной математики**

- Замена непрерывных объектов дискретными: сетки, сеточные функции, конечные разности и т.д
- В машинных вычислениях используются числа с ограниченным числом знаков после запятой из-за конечности длины мантиссы числа при представлении числа в памяти ЭВМ. Т.е. в вычислениях присутствует машинная погрешность округления.

- Большое значение имеет обусловленность задачи, т.е. чувствительность решения к малым изменениям входных данных.
- В отличие от классической математики выбор вычислительного алгоритма влияет на результат.
- Существенную роль играет экономичность вычислительного алгоритма, т.е. минимизация числа операций на ЭВМ.

1.1 Числа с плавающей точкой

Вещественные числа, представленные в форме с плавающей точкой, чаще всего используются при решении расчетных задач на компьютерах

$$A = \pm m \cdot \gamma^n \quad (1)$$

Здесь m — **мантисса** числа, γ — основание системы счисления, n — порядок числа.

Если $\frac{1}{\gamma} \leq m < 1$, то число имеет нормализованную форму. Запись числа в виде $XXX.XX$ – форма с фиксированной точкой.

В случае системы счисления с основанием γ произвольное число G будет иметь вид:

$$G = \pm 0.g_1g_2g_3\dots g_k \cdot \gamma^n \quad (2)$$

где g_i – целые числа, причем $0 \leq g_i \leq \gamma - 1$

Таким образом, подмножество действительных чисел, с которым работает компьютер, является конечным: оно определяется разрядностью k , а также минимальной и максимальной границами порядка n_1 и n_2 . Можно показать, что это подмножество содержит

$$N = 2(\gamma - 1)(n_2 - n_1 + 1)\gamma^{k-1} + 1 \quad (3)$$

чисел.

Наименьшее число называется машинным нулем и равно

$$G_0 = (\gamma - 1)\gamma^{n_1-1}, \quad (4)$$

наибольшее число называется машинной бесконечностью и равно

$$G_{\infty} = (1 - \gamma^{-k})\gamma^{n_2}, \quad (5)$$

В настоящее время общепринятым для арифметических операций с двоичными числами с плавающей точкой является стандарт IEEE 754, предусматривающий форматы числа с одинарной и двойной точностью.

Для этих форматов

Точность	Байты	k	n_1	n_2	G_0	G_{∞}
одинарная	4	24	-125	128	$1.2 \cdot 10^{-38}$	$3.4 \cdot 10^{38}$
двойная	8	53	-1021	1024	$2.2 \cdot 10^{-308}$	$1.8 \cdot 10^{308}$

Соответствующие типы данных в языке Си: float и double; в языке Паскаль: single и double; в языке Фортран: real и double precision.

1.2 Вычислительные погрешности

✓✓ Абсолютная погрешность некоторого числа равна разности между его приближенным значением, полученным в результате вычислений или измерений, и истинным значением.

✓✓ Относительная погрешность – отношение абсолютной погрешности к точному (или приближенному) значению числа.

$$\Delta a = a - a_0, \quad \delta a = \frac{\Delta a}{a} \quad (6)$$

Для приближенного числа, полученного в результате округления, абсолютная погрешность Δa принимается равной половине единицы последнего разряда числа. Т.е., если $a = 0.467$, то $\Delta a = 0.0005$.

✓✓ Значащими цифрами данного числа считаются все цифры, начиная с первой ненулевой цифры.

При изменении формы записи числа число значащих цифр не должно меняться. Например, $4300 = 0.4300 \cdot 10^4$.

Правила оценки предельных погрешностей

При вычислении абсолютных погрешностей надо использовать правила вычисления производных (с заменой $-$ на $+$):

$$\begin{aligned}\Delta(a \pm b) &= \Delta a + \Delta b \\ \delta(ab) &= \delta a + \delta b \\ \delta(a/b) &= \delta a + \delta b \\ \delta(a^k) &= k\delta a\end{aligned}\tag{7}$$

Найдем относительную погрешность разности двух чисел:

$$\delta(a - b) = \frac{\Delta(a - b)}{a - b} = \frac{\Delta a + \Delta b}{a - b}$$

Пример: $a = 5462$, $b = 5460$. Найти $\delta(a - b)$.

Очевидно, что $\Delta a = \Delta b = 0.5$. Тогда

$$\delta(a - b) = \frac{0.5 + 0.5}{2} = 0.5 \rightarrow 50\%.$$

□ Упражнение. Найти относительную погрешность выражения:

$$\sqrt{\frac{a+b}{c^2(1-c)}}.$$

Источники погрешностей

1. Математическая модель, если в ней не учитываются важные элементы рассматриваемой задачи, и исходные данные приводят к *неустранимым погрешностям*, т.к. они не могут быть уменьшены вычислителем.

2. *Погрешность численного метода* возникает из-за замены интеграла суммой, интерполирования данных и т.д. Погрешность численного метода как правило может быть сделана как угодно малой. Обычно ее доводят до величины в несколько раз меньшей неустранимой погрешности.

3. *Погрешности округлений* являются неизбежными из-за ограниченности разрядной сетки машины и могут быть найдены по фор-

муле:

$$\delta_{max} = 0.5\gamma^{1-k} \quad (8)$$

Здесь γ – основание системы счисления, k – количество разрядов мантиссы числа.

Способы уменьшения вычислительных погрешностей:

1. При вычислении суммы многих слагаемых избегать складывать большие по величине слагаемые с маленькими, а производить сложение по мере возрастания слагаемых.
2. При наличии алгебраических выражений производить их упрощение для более точных вычислений. Например, если $a \gg x$, то $(a+x)^2 - a^2 \rightarrow 0$, поэтому надо преобразовать к виду $(a+x)^2 - a^2 = 2ax + x^2$.
3. Организация постоянного контроля погрешностей для предотвращения получения неправильного результата.

1.3 Устойчивость, корректность и сходимость

В расчетных задачах возникают погрешности в исходных данных (не зависящие от вычислителя). Возникает вопрос об их влиянии на точность результатов расчетов. Оказывается, что некоторые задачи весьма чувствительны к таким погрешностям.

Предположим, что в результате решения задачи по исходному значению некоторой величины x находится некоторая величина y .

✓✓ Задача называется устойчивой по исходному параметру x , если решение y непрерывно от него зависит, т.е. малое приращение исходной величины Δx вызывает малое приращение расчетной величины Δy .

Отсутствие устойчивости означает, что даже незначительные погрешности в исходных данных приводят к большим погрешностям в решении.

Пример неустойчивой задачи – квадратное уравнение с парамет-

ром:

$$x^2 - 2x + \operatorname{sign} a = 0. \quad (9)$$

Его решение при $a \geq 0$ $x_1 = x_2 = 1$, при $a \leq 0$ $x_{1,2} = 1 \pm \sqrt{2}$. Очевидно, что при $a = 0$ малая отрицательная погрешность в задании a приведет к конечной погрешности в результате.

✓✓ Задача называется поставленной корректно, если для любых значений исходных данных ее решение существует и является единственным и устойчивым. В настоящее время существуют методы решения некорректных задач – методы регуляризации.

Метод численного решения может оказаться неустойчивым при накоплении погрешностей, поэтому для корректности численного алгоритма требуется его устойчивость.

✓ Сходимость численного метода означает близость получаемого численного решения задачи к истинному решению. Например, в случае итерационного процесса (метода последовательных приближений) получаем последовательность значений $x_1, x_2, \dots, x_n, \dots$. Го-

воят, что эта последовательность сходится к точному решению A , если при неограниченном возрастании числа итераций предел этой последовательности существует и равен A . В этом случае наш численный метод — сходящийся.

При использовании методов дискретизации (метод сеток) под сходимостью метода понимается стремления численного решения к точному при стремлении к нулю параметра дискретизации.

2 Численное решение нелинейных уравнений

2.1 Постановка задачи

Пусть задана функция $f(x)$ и нам надо найти корни уравнения

$$f(x) = 0 \quad (10)$$

Мы ограничимся поиском только действительных корней. Решение нашей задачи можно разделить на два этапа:

- 1 – локализация корней, т.е. нахождение на оси x таких отрезков, каждому из которых принадлежит не более одного корня;
- 2 – вычисление с требуемой точностью корня или корней, которые принадлежат найденным на первом этапе отрезкам.

Выполнение этапа 1 проводится средствами математического анализа. Из компьютерных методов можно использовать построение таблицы значений функции, построение графика и т.д.

Задача 2 этапа – вычисление с заданной точностью ε корня уравнения, принадлежащего отрезку $[a, b]$. Т.е., найденное значение корня

x должно отличаться от точного $x^{(T)}$ не более чем на ε :

$$|x - x^{(T)}| \leq \varepsilon \quad (11)$$

2.2 Метод деления отрезка пополам

Метод деления отрезка пополам также называют методом бисекции. В качестве начального приближения принимаем середину отрезка $[a, b]$:

$$x_0 = \frac{a + b}{2}. \quad (12)$$

Далее исследуем знак функции в точках a , b и x_0 . Тот из отрезков, на концах которого функция имеет разные знаки, принимаем в качестве нового отрезка $[a_1, b_1]$. Далее находим

$$\tilde{x}_1 = \frac{a_1 + b_1}{2}, \quad (13)$$

в результате n -ое приближение имеет вид

$$\tilde{x}_n = \frac{a_n + b_n}{2}, \quad (14)$$

причем

$$f(a_n)f(b_n) < 0, \quad (n = 1, 2, \dots) \quad (15)$$

$$b_n - a_n = \frac{b - a}{2^n}. \quad (16)$$

Так как левые концы $a_1, a_2, \dots, a_n, \dots$ образуют монотонную неубывающую ограниченную последовательность, а правые концы $b_1, b_2, \dots, b_n, \dots$ – монотонную невозрастающую ограниченную последовательность, то из (16) следует, что существует общий предел

$$\tilde{x} = \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n$$

Переходя к пределу $n \rightarrow \infty$ в (15) получаем

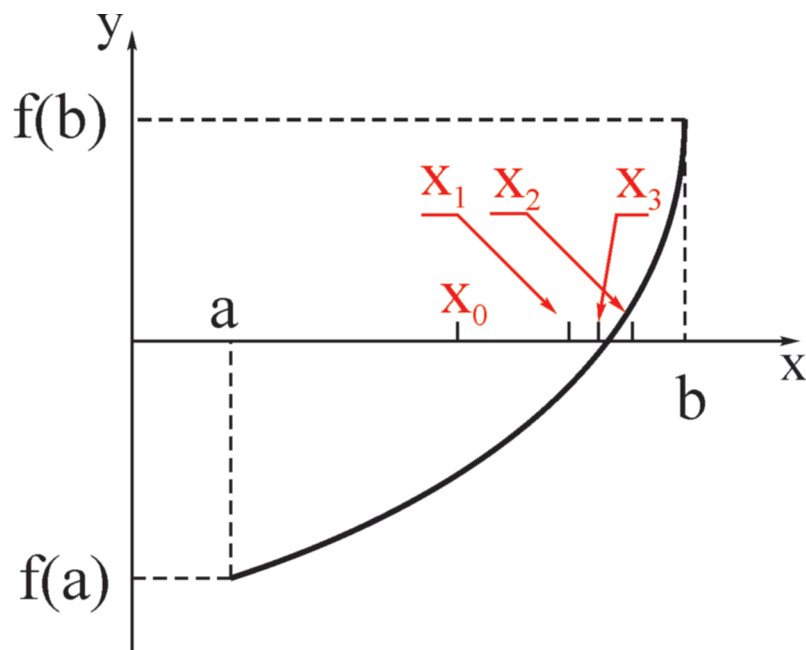
$$|f(\tilde{x})|^2 \leq 0.$$

Отсюда, $f(\tilde{x}) = 0$, т.е. \tilde{x} является корнем исходного уравнения.

С учетом (11) получаем, что вычисления надо продолжать до тех пор, пока не будет выполнено условие

$$b_n - a_n < 2\varepsilon. \quad (17)$$

Если корни уравнения не отделены на отрезке $[a, b]$, то таким способом можно найти один из корней.



На рисунке показан метод деления отрезка пополам. Отличительными особенностями данного метода являются:

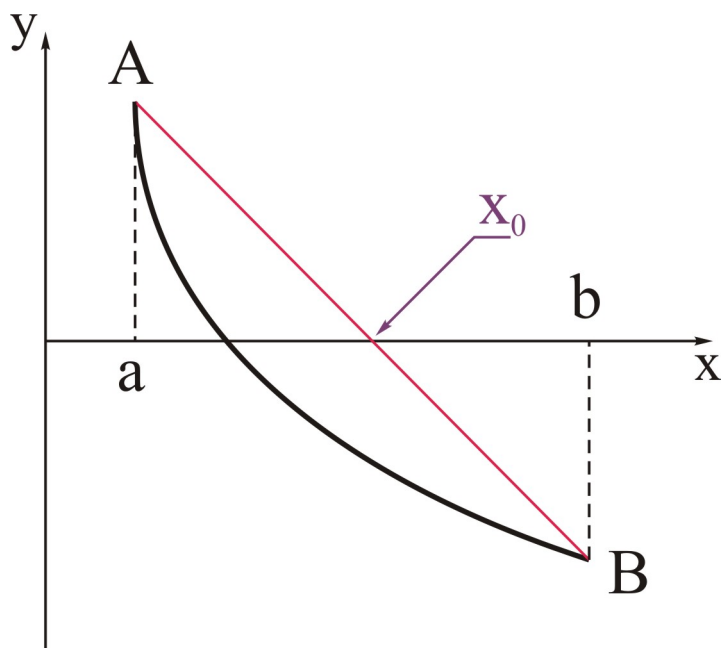
- метод всегда сходится, и полученный ответ всегда будет иметь любую наперед заданную точность;
- метод весьма медленный, необходимое число итераций

$$N > \log_2 \frac{b - a}{2\varepsilon}.$$

Данный метод практически удобно применять для грубого нахождения корня, т.к. при увеличении точности значительно возрастает объем вычислений.

2.3 Метод пропорциональных частей (метод хорд)

Более быстрым, чем метод бисекций, является метод пропорциональных частей. Он состоит в нахождении корня на отрезке $[a, b]$ таком, что $f(a)f(b) < 0$. Причем отрезок делится при итерациях не пополам, а на пропорциональные части, относящиеся как $f(a) : f(b)$. Геометрически это соответствует тому, что за приближение к корню принимается точка пересечения хорды, проходящей через точки $(a, f(a))$ и $(b, f(b))$, с осью x .

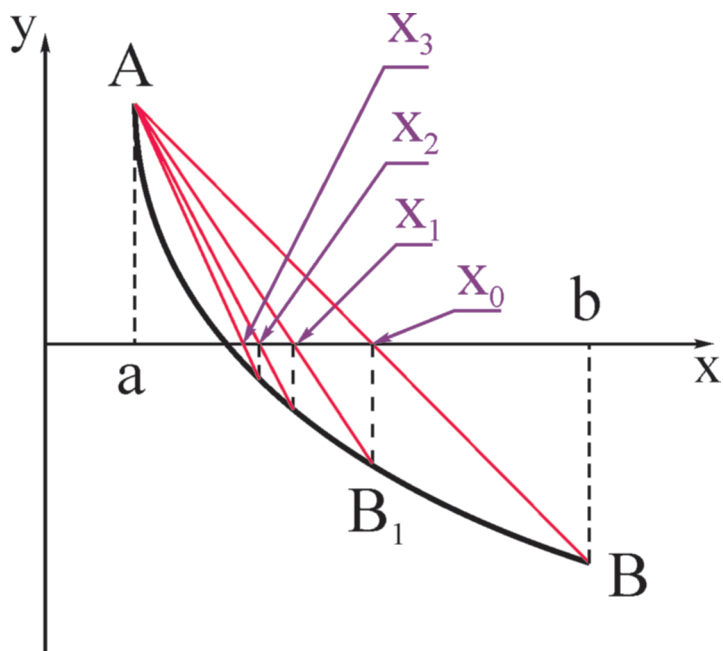


Пусть для определенности $f(a) > 0$, $f(b) < 0$. Запишем уравнение хорды AB:

$$\frac{y - f(a)}{f(b) - f(a)} = \frac{x - a}{b - a} \quad (18)$$

Найдем точку ее пересечения с осью $x \rightarrow x = x_0, y = 0$:

$$x_0 = a - \frac{b - a}{f(b) - f(a)} f(a) \quad (19)$$



На следующем шаге сравниваем знаки $f(a)$, $f(x_0)$ и $f(b)$, делаем вывод, что корень принадлежит отрезку $[a, x_0]$. Строим хорду AB_1 , находим ее точку пересечения с осью x , т.е. следующее приближение x_1 , и т.д. Условие прекращения итераций – условие бли-

зости двух последовательных приближений:

$$x_n - x_{n-1} < \varepsilon. \quad (20)$$

В общем случае выпуклость кривой определяется знаком второй производной. Нашему случаю соответствует $f''(x) > 0$ на отрезке $[a, b]$. Если $f''(x) < 0$, то мы можем свести этот случай к нашему, рассматривая уравнение $-f(x) = 0$. Если, как было выбрано выше, $f(a) > 0$, $f(b) < 0$, то при последовательных приближениях конец a остается неподвижным, а “движется” конец b , т.е. получаем

$$x_{n+1} = x_n - \frac{x_n - a}{f(x_n) - f(a)} f(x_n). \quad (21)$$

Если же $f(a) < 0$, $f(b) > 0$, то “двигаться” будет уже конец a , и мы получаем

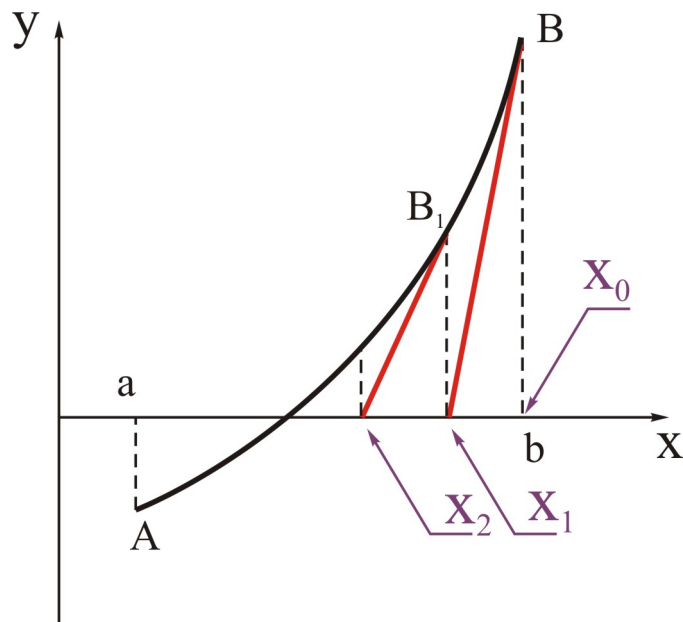
$$x_{n+1} = x_n - \frac{b - x_n}{f(b) - f(x_n)} f(x_n). \quad (22)$$

Замечание. Уравнение хорды АВ может быть переписано в виде

$$\frac{y - f(b)}{f(a) - f(b)} = \frac{b - x}{b - a}. \quad (23)$$

Метод хорд как правило обеспечивает более быструю сходимость, по сравнению с методом деления отрезка пополам. Поэтому можно сочетать оба метода: после нескольких шагов метода бисекций, когда корень уже в достаточной степени локализован, переходить к методу хорд.

2.4 Метод Ньютона



Метод Ньютона также называют методом касательных. Он отличается от метода хорд тем, что вместо хорды к кривой проводится касательная в точке, абсцисса которой является полученным приближением корня на предыдущем шаге. Начинается решение с выбора некоторого нулевого приближения x_0 . Положим $x_0 = b$, причем $f(x_0)f''(x_0) > 0$. Уравнение касательной, проведенной к кривой в точке B с координатами $(x_0, f(x_0))$ записывается в виде:

$$y = f'(x_0)(x - x_0) + f(x_0). \quad (24)$$

Следующее приближение – абсцисса точки пересечения касательной с осью x :

$$y = 0 \quad \Rightarrow \quad x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Формула для n -го приближения имеет вид:

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}, \quad n = 1, 2, \dots$$

Если в нашем случае взять $x_0 = a$, то на следующем шаге мы выйдем за пределы отрезка $[a, b]$, что является непрактичным. При этом можно доказать общее правило:

в качестве исходной точки x_0 выбирается тот конец отрезка $[a, b]$, чтобы

$$f(x_0)f''(x_0) > 0.$$

Условие прекращения итераций

$$x_n - x_{n-1} < \varepsilon. \tag{25}$$

Самым существенным отличием метода Ньютона от предыдущих — необходимость вычисления производной функции.

2.5 Метод итерации

Метод итерации является одним из наиболее важных способов численного решения уравнения.

Перепишем наше уравнение в виде:

$$x = \varphi(x)$$

Если у нас есть начальное приближение x_0 , то следующее приближение находим как

$$x_1 = \varphi(x_0)$$

Повторяя этот процесс, получим последовательность чисел

$$x_n = \varphi(x_{n-1}).$$

Предел этой последовательности (если он существует) является корнем уравнения.

3 Аппроксимация функций

3.1 Постановка задачи

Пусть некоторая величина y является функцией аргумента x , но явная связь между y и x неизвестна (либо известная зависимость $y = f(x)$ слишком громоздка для численных расчетов). Допустим, что в результате экспериментов получена таблица значений $\{x_i, y_i\}$, требуется же получить значения y в других точках, отличных от узлов x_i . Эта проблема решается в задаче о приближении (аппроксимации) функции: функцию $f(x)$, явный вид которой неизвестен, требуется приближенно заменить некоторой функцией $\varphi(x)$ (наз. аппроксимирующей), так чтобы отклонение от $f(x)$ в заданной области было наименьшим. Построенная таким образом аппроксимация называется точечной (примеры: интерполирование, среднеквадратичное приближение и т.д.)

Одним из основных типов точечной аппроксимации является ин-

терполирование: для заданной функции $f(x)$ строится интерполирующая функция $\varphi(x)$, принимающая в заданных точках x_i те же значения y_i , что и функция $f(x)$:

$$\varphi(x_i) = y_i, \quad i = 0, 1, \dots, n \quad (26)$$

причем $x_i \neq x_k$ при $i \neq k$, x_i – узлы интерполяции.

Интерполирующая функция может строиться сразу для всего рассматриваемого интервала x – глобальная интерполяция, или отдельно для разных частей этого интервала – кусочная (локальная) интерполяция. Если полученная функция $\varphi(x)$ применяется для нахождения значения функции $f(x)$ за пределами отрезка, содержащего узлы, то говорят об экстраполяции.

Рассмотрим использование в качестве функции $\varphi(x)$ интерполяционного многочлена

$$\varphi(x) = P_m(x) = a_0 + a_1x + a_2x^2 + \dots + a_mx^m. \quad (27)$$

При глобальной интерполяции мы будем использовать все $n+1$ уравнений системы (26), что позволяет найти $n+1$ коэффициент, откуда

следует, что максимальная степень интерполяционного многочлена $- m = n$:

$$P_n(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n. \quad (28)$$

Подставляя (28) в (26) получаем:

$$\begin{aligned} a_0 + a_1x_0 + a_2x_0^2 \dots + a_nx_0^n &= y_0 \\ a_0 + a_1x_1 + a_2x_1^2 \dots + a_nx_1^n &= y_1 \\ \dots & \\ a_0 + a_1x_n + a_2x_n^2 \dots + a_nx_n^n &= y_n \end{aligned} \quad (29)$$

(29) – система линейных алгебраических уравнений относительно неизвестных коэффициентов a_i . Определитель такой системы – определитель Вандермонда – отличен от нуля, если среди узлов x_i нет совпадающих. Следовательно, в этом случае система (29) имеет единственное решение. Решив систему (29) построим интерполяционный многочлен – метод неопределенных коэффициентов.

Недостатки такого метода:

– при большом количестве узлов получается высокая степень мно-

гочлена,

– привязка к узлам интерполяции, которые, если они получены в результате измерений, могут содержать случайные погрешности.

Другой способ – подбор наиболее простой аппроксимирующей функции, график которой проходит максимально близко от узлов.

Мера отклонения функции $\varphi(x)$ от заданной функции $f(x)$:

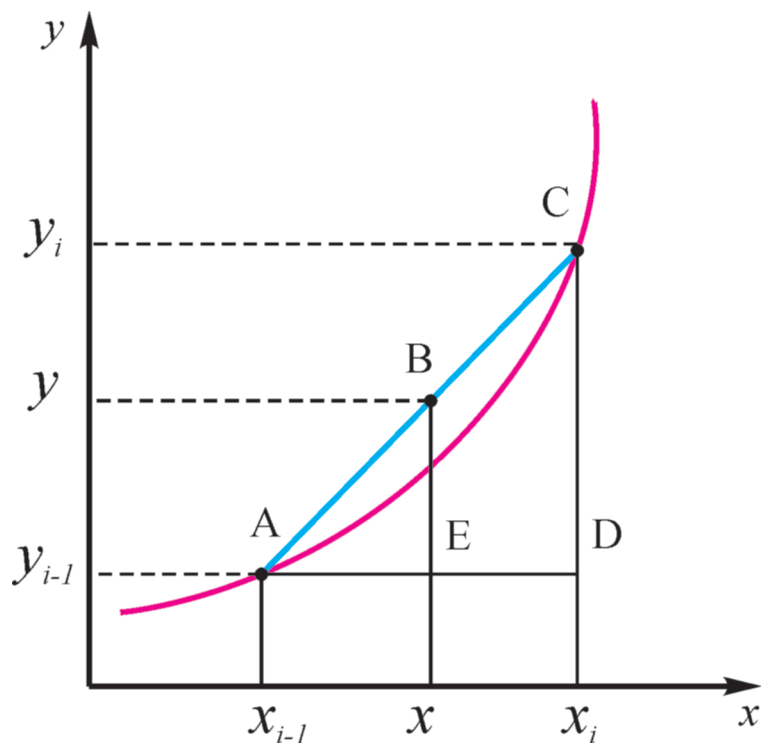
$$S = \sum_{i=0}^n |\varphi(x_i) - y_i|^2. \quad (30)$$

Метод наименьших квадратов состоит в подборе аппроксимирующей функции так, чтобы S было наименьшим.

3.2 Интерполирование

Кусочно-линейная интерполяция

КЛИ состоит в том, что заданные точки (x_i, y_i) соединяются прямолинейными отрезками, и функция $f(x)$ приближается ломанной с вершинами в узлах. Всего имеется n интервалов (x_{i-1}, x_i) , для каждого из них интерполяционным многочленом является уравнение прямой, проходящей через две точки.



Например, для i -го интервала уравнение прямой, проходящей через точки (x_{i-1}, y_{i-1}) и (x_i, y_i) имеет вид:

$$\frac{y - y_{i-1}}{y_i - y_{i-1}} = \frac{x - x_{i-1}}{x_i - x_{i-1}} \quad (31)$$

Отсюда находим

$$\begin{aligned} y &= a_i x + b_i, \\ x_{i-1} &\leq x \leq x_i \end{aligned} \quad (32)$$

$$a_i = \frac{y_i - y_{i-1}}{x_i - x_{i-1}},$$

$$b_i = y_{i-1} - a_i x_{i-1}$$

Многочлен Лагранжа

Будем строить интерполяционный многочлен, единый для всего отрезка $[x_0, x_n]$, в виде линейной комбинации многочленов степени n :

$$L(x) = y_0 l_0(x) + y_1 l_1(x) + \dots + y_n l_n(x) \quad (33)$$

так, чтобы многочлены $l_i(x)$ обращались в нуль во всех узлах интерполяции, кроме i - го, где он должен равняться единице. Этим условиям при $i = 0$ отвечает многочлен вида:

$$l_0(x) = \frac{(x - x_1)(x - x_2)(x - x_3) \dots (x - x_n)}{(x_0 - x_1)(x_0 - x_2)(x_0 - x_3) \dots (x_0 - x_n)} \quad (34)$$

Аналогично,

$$\begin{aligned}
l_1(x) &= \frac{(x - x_0)(x - x_2)(x - x_3)\dots(x - x_n)}{(x_1 - x_0)(x_1 - x_2)(x_1 - x_3)\dots(x_1 - x_n)} \\
l_2(x) &= \frac{(x - x_0)(x - x_1)(x - x_3)\dots(x - x_n)}{(x_2 - x_0)(x_2 - x_1)(x_2 - x_3)\dots(x_2 - x_n)} \\
&\dots \\
l_i(x) &= \frac{(x - x_0)\dots(x - x_{i-1})(x - x_{i+1})\dots(x - x_n)}{(x_i - x_0)\dots(x_i - x_{i-1})(x_i - x_{i+1})\dots(x_i - x_n)} \\
&\dots \\
l_n(x) &= \frac{(x - x_0)(x - x_1)(x - x_2)\dots(x - x_{n-1})}{(x_n - x_0)(x_n - x_1)(x_n - x_2)\dots(x_n - x_{n-1})}
\end{aligned} \tag{35}$$

Подставляя (34),(35) в (33) получаем интерполяционный многочлен Лагранжа:

$$L(x) = \sum_{i=0}^n y_i \frac{(x - x_0)\dots(x - x_{i-1})(x - x_{i+1})\dots(x - x_n)}{(x_i - x_0)\dots(x_i - x_{i-1})(x_i - x_{i+1})\dots(x_i - x_n)} \tag{36}$$

Единственность найденного решения следует из единственности решения системы (29). Если положить $n = 1$ в (36), то получим рассмотренный ранее случай линейной интерполяции, при $n = 2$ — случай квадратичной интерполяции:

$$L(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)}y_0 + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)}y_1 + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}y_2 \quad (37)$$

□ Упражнение. Проверить, что из (36) при $n = 1$ следует случай линейной интерполяции.

Многочлен Ньютона

При построении интерполяционного многочлена Лагранжа не накладывалось никакого требования на распределение узлов интерполяции. Рассмотрим случай равноотстоящих по оси x узлов интерпо-

ляции. Введем $h = x_i - x_{i-1}$ — шаг интерполяции, $h = \text{const}$.

Конечные разности

Разности первого порядка (первые разности)

$$\Delta y_0 = y_1 - y_0 = f(x_0 + h) - f(x_0)$$

$$\Delta y_1 = y_2 - y_1 = f(x_0 + 2h) - f(x_0 + h)$$

...

$$\Delta y_{n-1} = y_n - y_{n-1} = f(x_0 + nh) - f(x_0 + (n-1)h)$$

Разности второго порядка (вторые разности)

$$\Delta^2 y_0 = \Delta y_1 - \Delta y_0$$

$$\Delta^2 y_1 = \Delta y_2 - \Delta y_1$$

...

Разности порядка k

$$\Delta^k y_i = \Delta^{k-1} y_{i+1} - \Delta^{k-1} y_i, \quad i = 0, 1, \dots, n-1$$

Конечные разности выражаются через значения функции

$$\Delta^2 y_0 = \Delta y_1 - \Delta y_0 = (y_2 - y_1) - (y_1 - y_0) = y_2 - 2y_1 + y_0$$

$$\Delta^3 y_0 = \Delta^2 y_1 - \Delta^2 y_0 = \dots = y_3 - 3y_2 + 3y_1 - y_0$$

В общем случае

$$\Delta^k y_i = \sum_{j=0}^k (-1)^j C_k^j y_{i+k-j} \quad (38)$$

где

$$C_k^j = \frac{k!}{j!(k-j)!}$$

□ Упражнение. Используя метод математической индукции доказать формулу (38).

Интерполяционный многочлен Ньютона будем искать в следующем виде:

$$N(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \dots + \\ + a_n(x - x_0)(x - x_1)\dots(x - x_{n-1}) \quad (39)$$

График многочлена должен проходить через заданные узлы, т.е $N(x_i) = y_i$ $i = 0, 1, \dots, n$. Для нахождения коэффициентов многочлена получаем систему

$$\begin{aligned} N(x_0) &= a_0 = y_0, \\ N(x_1) &= a_0 + a_1(x_1 - x_0) = a_0 + a_1h = y_1, \\ N(x_2) &= a_0 + a_1(x_2 - x_0) + a_2(x_2 - x_0)(x_2 - x_1) = \\ &= a_0 + 2a_1h + 2a_2h^2 = y_2, \\ &\dots \end{aligned}$$

Отсюда находим коэффициенты a_i :

$$a_0 = y_0, \quad a_1 = \frac{y_1 - a_0}{h} = \frac{y_1 - y_0}{h} = \frac{\Delta y_0}{h}$$

$$a_2 = \frac{y_2 - a_0 - 2a_1h}{2h^2} = \frac{y_2 - a_0 - 2\Delta y_0}{2h^2} = \frac{\Delta^2 y_0}{2h^2}$$

...

$$a_k = \frac{\Delta^k y_0}{k!h^k}, \quad k = 0, 1, 2, \dots, n.$$

Подставляя в (39), получаем

$$\begin{aligned} N(x) = y_0 + \frac{\Delta y_0}{h}(x - x_0) + \frac{\Delta^2 y_0}{2!h^2}(x - x_0)(x - x_1) + \dots + \\ + \frac{\Delta^n y_0}{n!h^n}(x - x_0)(x - x_1)\dots(x - x_{n-1}) \quad (40) \end{aligned}$$

Если ввести переменную $t = \frac{x - x_0}{h}$, то

$$x = x_0 + th, \quad \frac{x - x_1}{h} = \frac{x - x_0 - h}{h} = t - 1$$

$$\frac{x - x_2}{h} = t - 2, \quad \dots, \quad \frac{x - x_{n-1}}{h} = t - n + 1$$

В результате получаем

$$N(x) = N(x_0 + th) = y_0 + t\Delta y_0 + \frac{t(t-1)}{2!}\Delta^2 y_0 + \dots + \\ + \frac{t(t-1)\dots(t-n+1)}{n!}\Delta^n y_0 \quad (41)$$

Полученное выражение называется первым интерполяционным многочленом Ньютона для интерполирования вперед. Оно справедливо на всем отрезке $[x_0, x_n]$, но для уменьшения ошибок округления разумно использовать его только для левой половины рассматриваемого отрезка.

Полученная формула для интерполирования вперед практически неудобна для интерполирования функции вблизи конца отрезка. В этом случае используют формулу для многочлена Ньютона для интерполирования назад, которую мы сейчас и получим.

Интерполирующий полином запишем в следующем виде:

$$N(x) = a_0 + a_1(x - x_n) + a_2(x - x_n)(x - x_{n-1}) + \dots \\ + a_n(x - x_n)(x - x_{n-1})\dots(x - x_1). \quad (42)$$

Из условия совпадения значения многочлена и функции в узлах находим

$$N(x_n) = y_n \Rightarrow a_0 = y_n \\ N(x_{n-1}) = y_{n-1} \Rightarrow y_n + a_1(-h) = y_{n-1} \Rightarrow \\ a_1 = \frac{y_n - y_{n-1}}{h} = \frac{\Delta y_{n-1}}{h}$$

Аналогично находим

$$a_2 = \frac{y_n - 2y_{n-1} + y_{n-2}}{2h^2} = \frac{\Delta^2 y_{n-2}}{2h^2}$$

и в общем случае, применяя метод математической индукции можно строго доказать, что

$$a_k = \frac{\Delta^k y_{n-k}}{k!h^k}.$$

Подставляя найденные коэффициенты в (42), получаем

$$N(x) = y_n + \frac{\Delta y_{n-1}}{1!h}(x - x_n) + \frac{\Delta^2 y_{n-2}}{2!h^2}(x - x_n)(x - x_{n-1}) + \dots \\ + \frac{\Delta^n y_0}{n!h^n}(x - x_n)\dots(x - x_1). \quad (43)$$

Делая замену $t = \frac{x - x_n}{h}$, окончательно находим

$$N(x) = N(x_n + th) = y_n + t\Delta y_{n-1} + \frac{t(t+1)}{2!}\Delta^2 y_{n-2} + \dots + \\ + \frac{t(t+1)\dots(t+n-1)}{n!}\Delta^n y_0 \quad (44)$$

Соотношение (44) – второй интерполяционный многочлен Ньютона для интерполирования назад.

Отметим, что существует один и только один интерполяционный многочлен при заданном наборе узлов интерполяции. Формулы Лагранжа, Ньютона и др. порождают один и тот же многочлен, разница

состоит в алгоритме их построения.

3.3 Точность интерполяции

Значения интерполяционного многочлена $y = \varphi(x)$ и рассматриваемой функции $y = f(x)$ в узлах $x = x_i$, ($i = 0, 1, \dots, n$) в точности совпадают. Если исследуемая функция — многочлен степени n , то $f(x) \equiv \varphi(x)$. В остальных случаях разность

$$R(x) = f(x) - \varphi(x) \neq 0$$

Очевидно, что $R(x)$ есть погрешность интерполяции и носит название — остаточный член интерполяционной формулы. Можно показать, что остаточный член интерполяционного многочлена Лагранжа имеет вид

$$R_L(x) = \frac{(x - x_0)(x - x_1) \dots (x - x_n)}{(n + 1)!} f^{(n+1)}(x') \quad (45)$$

В этой формуле $f^{(n+1)}(x')$ – производная $(n+1)$ - го порядка функции $f(x)$ в некоторой точке $x' \in [x_0, x_n]$. Из анализа (45) следует, что погрешность интерполяции тем выше, чем ближе точка x лежит к концам отрезка $[x_0, x_n]$. Если же использовать интерполяционный многочлен вне отрезка $[x_0, x_n]$, то погрешность возрастает очень заметно.

Остаточный член интерполяционного многочлена Ньютона для случая равноотстоящих узлов следует из (45):

$$R_N(x) = \frac{t(t-1)\dots(t-n)}{(n+1)!} f^{(n+1)}(x') h^{n+1}, \quad t = \frac{x - x_0}{h}. \quad (46)$$

Из вида остаточного члена следует, что повышение степени интерполяционного многочлена уменьшает погрешность, однако из-за неясного поведения $f^{(n+1)}(x)$ возможны проблемы. Поэтому на практике для повышения точности целесообразно уменьшать шаг и выбирать специальное расположение узлов (например, сгущая их к концам отрезка). При этом как правило стараются использовать много-

члены малой степени.

3.4 Сплайны

В настоящее время для интерполяции широко используются кубические сплайн-функции (сплайны), представляющие собой специальным образом построенные многочлены третьей степени. Они представляют собой математическую модель гибкого тонкого стержня, закрепленного между узлам интерполяции при условии минимума его потенциальной энергии. Общий вид сплайна:

$$S_i(x) = a_i + b_i(x - x_{i-1}) + c_i(x - x_{i-1})^2 + d_i(x - x_{i-1})^3 \quad (47)$$

на отрезке $x_{i-1} \leq x \leq x_i$. Незвестных коэффициентов a_i, b_i, c_i, d_i возникает $4n$ штук, и для их нахождения надо столько же уравнений. Условия прохождения графика S_i через узлы

$$S_i(x_{i-1}) = y_{i-1}, \quad S_i(x_i) = y_i, \quad i = 1, 2, \dots, n$$

дают $2n$ уравнений. Еще $2n - 2$ возникают из условий непрерывности первых и вторых производных во внутренних узлах интерполяции:

$$S'_i(x_i) = S'_{i+1}(x_i), \quad S''_i(x_i) = S''_{i+1}(x_i), \quad i = 1, 2, \dots, n - 1$$

И последние два уравнения получают из условия закрепления концов сплайна

$$S'(x_0) = k_1, \quad S'(x_n) = k_2$$

или

$$S''(x_0) = m_1, \quad S''(x_n) = m_2$$

причем k_1 и k_2 определяют угол наклона сплайна в точках закрепления, а m_1 и m_2 — кривизну в точках закрепления. При свободном закреплении

$$m_1 = m_2 = 0$$

и полученная таким образом функция называется свободным кубическим сплайном.

4 Численное дифференцирование

Производной функции $y = f(x)$ называется предел отношения приращения функции Δy к приращению аргумента Δx при $\Delta x \rightarrow 0$:

$$y' = f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} \quad (48)$$

В случае если функция задана в виде таблицы, для вычисления производной используется приближенное равенство

$$y' \approx \frac{\Delta y}{\Delta x} \quad (49)$$

Эта формула называется аппроксимацией производной с помощью конечных разностей.

С помощью левых разностей производная может быть записана следующим образом

$$\Delta y_1 = y_1 - y_0, \quad \Delta x = h, \quad y'_1 \approx \frac{y_1 - y_0}{h}, \quad (50)$$

с помощью правых разностей

$$\nabla y_1 = y_2 - y_1, \quad \Delta x = h, \quad y'_1 \approx \frac{y_2 - y_1}{h}, \quad (51)$$

с помощью центральных

$$\delta y_1 = y_2 - y_0, \quad \Delta x = 2h, \quad y'_1 \approx \frac{y_2 - y_0}{2h}, \quad (52)$$

Аналогично можно вычислить старшие производные

$$y''_1 = (y'_1)' \approx \frac{y'_2 - y'_1}{h} \approx \frac{(y_2 - y_1)/h - (y_1 - y_0)/h}{h} = \frac{y_2 - 2y_1 + y_0}{h}$$

Для вычисления производных могут использоваться интерполяционные формулы. Например, в случае интерполяционного многочлена Ньютона:

$$\frac{dN}{dx} = \frac{dN}{dt} \frac{dt}{dx} = \frac{1}{h} \frac{dN}{dt} \quad (53)$$

Найдя в явном виде производную $\frac{dN}{dt}$, получим формулу для вычисления производной функции. Также можно для этих целей воспользоваться многочленом Лагранжа.

5 Численное интегрирование

Пусть на отрезке $[a, b]$ задана функция $y = f(x)$. Разобьем отрезок на n элементарных отрезков $[x_{i-1}, x_i]$ ($i = 1, 2, \dots, n$). На каждом из этих отрезков выберем произвольную точку ξ_i и составим сумму произведений значения функции $f(\xi_i)$ на длину отрезка Δx_i :

$$S_n = \sum_{i=1}^n f(\xi_i) \Delta x_i \quad (54)$$

S_n – интегральная сумма.

Определенным интегралом от функции $f(x)$ на отрезке $[a, b]$ называется предел интегральной суммы при неограниченном увеличении числа точек разбиения, так что длина наибольшего отрезка стремится к нулю:

$$\int_a^b f(x) dx = \lim_{\max \Delta x_i \rightarrow 0} \sum_{i=1}^n f(\xi_i) \Delta x_i \quad (55)$$

Во многих случаях, если подынтегральная функция задана в аналитическом виде, интеграл вычисляется непосредственно по формуле Ньютона-Лейбница:

$$\int_a^b f(x)dx = F(b) - F(a) \quad (56)$$

Но реально воспользоваться этой формулой получается не всегда. Основные причины:

вид функции $f(x)$ не позволяет выразить первообразную в элементарных функциях

аналитический вид функции $f(x)$ неизвестен, известны значения функции в фиксированных точках. В этих случаях используются приближенные методы интегрирования и, прежде всего, численные. Задача численного интегрирования функции заключается в вычислении значения определенного интеграла на основании ряда значений подынтегральной функции. Численное вычисление однократно-

го интеграла называют квадратурой, двойного — кубатурой.

В случае однократного интеграла подынтегральную функцию $f(x)$ заменяют на рассматриваемом отрезке $[a, b]$ интерполирующей или аппроксимирующей функцией $\varphi(x)$, а затем приближенно полагают

$$\int_a^b f(x)dx \approx \int_a^b \varphi(x)dx \quad (57)$$

причем интеграл в правой части равенства (57) вычисляется непосредственно. В результате находим

$$\int_a^b \varphi(x)dx = \sum_{i=0}^n \alpha_i y_i. \quad (58)$$

Т.о., мы получили, что

$$\int_a^b f(x)dx \approx \sum_{i=0}^n \alpha_i y_i. \quad (59)$$

Здесь y_i – значения функции в узлах интерполяции, α_i – числовые коэффициенты. Формула (59) носит название квадратурная формула, правая часть – квадратурная сумма.

В зависимости от способа ее вычисления возникают разные методы численного интегрирования – прямоугольника, трапеций, парабол, сплайнов и т.д.

Квадратурная формула (59) может быть записана в другом виде

$$\int_a^b f(x)dx \approx \sum_{i=0}^n \sigma_i \quad (60)$$

где σ_i – приближенное значение площади элементарной криволинейной трапеции, соответствующей отрезку $[x_{i-1}, x_i]$.

В случае, если в качестве $\varphi(x)$ взять многочлен Лагранжа, построенный на $n + 1$ равноотстоящих узлах, то полученные с его помощью значения коэффициентов α_i в (59) вместе с самой формулой (59) образуют т.н. формулы Ньютона-Котеса.

5.1 Метод прямоугольников

Метод прямоугольников предполагает замену интеграла интегральной суммой

$$\int_a^b f(x)dx \approx \sum_{i=1}^n f(\xi_i)\Delta x_i \quad (61)$$

Под ξ может пониматься как левая, так и правая граница элементарных отрезков: $\xi_i = x_{i-1}$ или $\xi_i = x_i$.

Соответствующие формулы метода прямоугольников

$$\int_a^b f(x)dx \approx h_1 y_0 + h_2 y_1 + \dots + h_n y_{n-1} \quad (62)$$

$$\int_a^b f(x)dx \approx h_1 y_1 + h_2 y_2 + \dots + h_n y_n \quad (63)$$

где $h_i = x_i - x_{i-1}$.

Более точным метод прямоугольников становится, если использовать значения функции в средних точках элементарных отрезков

$$\int_a^b f(x)dx \approx \sum_{i=1}^n h_i f(x_{i-1/2}), \quad (64)$$

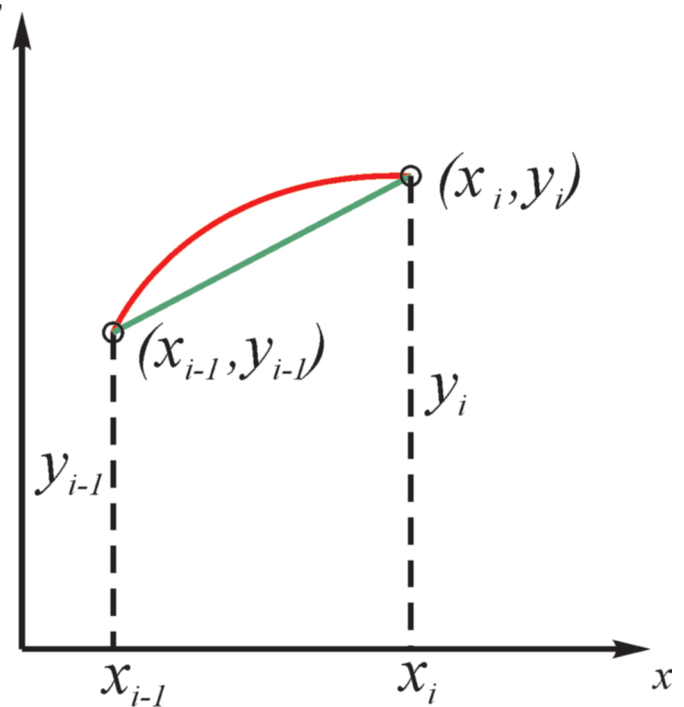
где

$$x_{i-1/2} = \frac{x_{i-1} + x_i}{2} = x_{i-1} + \frac{h_i}{2}$$

Последняя формула отражает метод средних.

5.2 Метод трапеций

Метод прямоугольников соответствует кусочно постоянной интерполяции: на каждом отрезке функция заменяется постоянной. Метод трапеций использует линейную интерполяцию — на каждом отрезке функция заменяется линейной функцией, в результате получаем ломаную, соединяющую узлы. В результате значение интеграла представляется как сумма площадей элементарных трапеций. Площадь каждой трапеции



$$\sigma_i = \frac{y_{i-1} + y_i}{2} h_i, \quad i = 1, \dots, n.$$

В результате формула трапеций для численного интегрирования имеет вид

$$\int_a^b f(x)dx \approx \frac{1}{2} \sum_{i=1}^n h_i(y_{i-1} + y_i). \quad (65)$$

В случае постоянного шага формула трапеций принимает вид

$$\int_a^b f(x)dx \approx h \frac{y_0 + y_n}{2} + h \sum_{i=1}^{n-1} y_i. \quad (66)$$

Очевидно, что погрешность формул прямоугольников и трапеций определяется шагом интегрирования. С уменьшением шага увеличивается точность вычисления интеграла. Если увеличение числа точек невозможно (функция задана в табличном виде), повышения точности можно достичь увеличивая степень используемых интерполяционных многочленов.

5.3 Метод Симпсона

Разобьем отрезок интегрирования $[a, b]$ на четное число равных частей с шагом h : $[x_0, x_2], [x_2, x_4], \dots, [x_{i-1}, x_{i+1}], \dots, [x_{n-2}, x_n]$, n – четное число. На каждом отрезке подынтегральную функцию заменим интерполяционным многочленом второй степени:

$$f(x) \approx \psi_i(x) = a_i x^2 + b_i x + c_i$$

В качестве $\psi_i(x)$ можно взять интерполяционный многочлен Лагранжа, проходящий через точки (x_{i-1}, y_{i-1}) , (x_i, y_i) , (x_{i+1}, y_{i+1}) :

$$\begin{aligned} \psi_i(x) &= \frac{(x - x_i)(x - x_{i+1})}{(x_{i-1} - x_i)(x_{i-1} - x_{i+1})} y_{i-1} + \\ &+ \frac{(x - x_{i-1})(x - x_{i+1})}{(x_i - x_{i-1})(x_i - x_{i+1})} y_i + \frac{(x - x_{i-1})(x - x_i)}{(x_{i+1} - x_{i-1})(x_{i+1} - x_i)} y_{i+1} = \\ &= \frac{1}{2h^2} \{ (x - x_i)(x - x_{i+1}) y_{i-1} - 2(x - x_{i-1})(x - x_{i+1}) y_i + \\ &\quad + (x - x_{i-1})(x - x_i) y_{i+1} \} \quad (67) \end{aligned}$$

Площадь криволинейной трапеции соответствующей отрезку $[x_{i-1}, x_{i+1}]$ можно найти с помощью определенного интеграла

$$s_{i-1,i+1} = \int_{x_{i-1}}^{x_{i+1}} \psi_i(x) dx = \frac{h}{3}(y_{i-1} + 4y_i + y_{i+1}) \quad (68)$$

Его вычисление можно провести в системе Maple (см. пример).

Аналогично можно вычислить такие интегралы для каждого элементарного отрезка и результаты просуммировать. В итоге получим

$$S = \frac{h}{3}(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \dots + 2y_{n-2} + 4y_{n-1} + y_n)$$

Таким образом, мы получили формулу Симпсона (или формулу парабол):

$$\int_a^b f(x) dx \approx \frac{h}{3} \{y_0 + 4(y_1 + y_3 + \dots + y_{n-1}) + \\ + 2(y_2 + y_4 + \dots + y_{n-2}) + y_n\} \quad (69)$$

5.4 Метод Гаусса

Метод Гаусса не предполагает разбиения отрезка интегрирования на равные промежутки. Формулы численного интегрирования интерполяционного типа ищутся такими, чтобы они обладали наивысшим порядком точности при заданном числе узлов.

Т.е., задача ставится таким образом: нужно подобрать узлы t_1, t_2, \dots, t_n и коэффициенты $\alpha_1, \alpha_2, \dots, \alpha_n$ таким образом, чтобы квадратурная формула

$$\int_{-1}^1 f(t)dt = \sum_{i=1}^n \alpha_i f(t_i) \quad (70)$$

была точной для всех полиномов $f(t)$ наивысшей возможной степени N . Т.к. мы имеем $2n$ постоянных t_i, α_i , а полином степени $2n - 1$ определяется $2n$ коэффициентами, то эта наивысшая степень будет равна $N = 2n - 1$.

В частности, необходимо, чтобы равенство (70) было верным при

$$f(t) = 1, t, t^2, \dots, t^{2n-1}. \quad (71)$$

Подставляя последовательно $f(t)$ из (71) в (70) и учитывая, что

$$\int_{-1}^1 t^k dt = \frac{1 - (-1)^{k+1}}{k+1} = \begin{cases} \frac{2}{k+1}, & \text{при } k \text{ четном;} \\ 0, & \text{при } k \text{ нечетном} \end{cases} \quad (72)$$

получаем систему $2n$ уравнений для нахождения t_i и α_i :

$$\begin{aligned}
 \sum_{i=1}^n \alpha_i &= 2, \\
 \sum_{i=1}^n \alpha_i t_i &= 0, \\
 &\dots\dots\dots \\
 \sum_{i=1}^n \alpha_i t_i^{2n-2} &= \frac{2}{2n-1}, \\
 \sum_{i=1}^n \alpha_i t_i^{2n-1} &= 0.
 \end{aligned} \tag{73}$$

Для того, чтобы не находить значения коэффициентов и узлов каждый раз, их значения табулированы для отрезка интегрирования $[-1, 1]$. Простой заменой переменных произвольный отрезок $[a, b]$ мо-

жет быть приведен к $[-1, 1]$:

$$\int_a^b \dots dx \rightarrow \int_{-1}^1 \dots dt,$$

$$x = \frac{b+a}{2} + \frac{b-a}{2}t, \quad dx = \frac{b-a}{2}dt,$$

$$\int_a^b f(x) dx = \frac{b-a}{2} \int_{-1}^1 f\left(\frac{b+a}{2} + \frac{b-a}{2}t\right) dt.$$

Например, для $n = 4$ узлы и коэффициенты формулы Гаусса

$$-t_1 = t_4 = 0.861\,136\,311\,594\,0492$$

$$-t_2 = t_3 = 0.339\,981\,043\,584\,8646$$

$$\frac{1}{2}\alpha_1 = \frac{1}{2}\alpha_4 = 0.173\,927\,422\,568\,7284$$

$$\frac{1}{2}\alpha_2 = \frac{1}{2}\alpha_3 = 0.326\,072\,577\,431\,2716$$

5.5 Точность численного интегрирования

Можно показать, что остаточный член метода прямоугольников на отрезке $[a, b]$ имеет вид

$$|R_{\text{пр}}| \leqslant \frac{(b-a)h^2}{24} \max_{a \leqslant x \leqslant b} |f''(x)|$$

для метода трапеций

$$|R_{\text{тр}}| \leq \frac{(b-a)h^2}{12} \max_{a \leq x \leq b} |f''(x)|$$

для метода Симпсона

$$|R_{\text{С}}| \leq \frac{(b-a)h^4}{180} \max_{a \leq x \leq b} |f^{(4)}(x)|$$

для n – точечного метода Гаусса

$$|R_{\Gamma}| \leq \frac{2^{2n+1}(n!)^4}{[(2n)!]^3(2n+1)} \max_{a \leq x \leq b} |f^{(2n)}(x)|$$

для $n = 4$ – метода Гаусса

$$|R_{\Gamma}| \leq \frac{1}{3\,472\,875} \max_{a \leq x \leq b} |f^{(8)}(x)|$$

5.6 Особые случаи численного интегрирования

1. Подынтегральная функция разрывна на отрезке интегрирования

Если подынтегральная функция на отрезке интегрирования терпит разрыв, то интеграл вычисляется численно для каждого отрезка непрерывности и результаты складывают. Например, в случае одной точки разрыва $x = c$ ($a < c < b$)

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

2. Несобственные интегралы.

Несобственными интегралами называются такие интегралы, которые имеют хотя бы одну бесконечную границу интегрирования или подынтегральную функцию, обращающуюся в бесконечность хотя бы в одной точке отрезка интегрирования.

Интеграл с бесконечным пределом

$$\int_a^{\infty} f(x)dx$$

называется сходящимся, если существует конечный предел

$$\lim_{b \rightarrow \infty} \int_a^b f(x)dx$$

и по определению

$$\int_a^{\infty} f(x)dx = \lim_{b \rightarrow \infty} \int_a^b f(x)dx.$$

Для вычисления сходящегося несобственного интеграла с заданной

точностью ε его представляют в виде

$$\int_a^\infty f(x)dx = \int_a^b f(x)dx + \int_b^\infty f(x)dx.$$

В случае сходимости интеграла число b можно выбрать столь большим, чтобы

$$\left| \int_b^\infty f(x)dx \right| < \frac{\varepsilon}{2}.$$

В случае, если функция обращается в бесконечность в некоторой точке $x = c$ конечного отрезка интегрирования, то можно попробовать выделить особенность, представив подынтегральную функцию в виде суммы двух функций:

$$f(x) = \varphi(x) + \psi(x)$$

так чтобы $\varphi(x)$ была ограничена, а несобственный интеграл от $\psi(x)$ вычислялся бы аналитически.

Возникающие в ряде задач сингулярные интегралы вида

$$\int_a^b \frac{f(x)}{x - c} dx$$

понимаются в смысле главного значения

$$\lim_{\varepsilon \rightarrow 0} \left(\int_a^{c-\varepsilon} \frac{f(x)}{x - c} dx + \int_{c+\varepsilon}^b \frac{f(x)}{x - c} dx \right)$$

Этот интеграл может быть записан как сумма интеграла по отрезку, симметричному относительно точки c и интеграла от гладкой функции по оставшейся части.

5.7 Кратные интегралы

Будем рассматривать двойные интегралы

$$\iint_G f(x, y) dx dy \quad (74)$$

Самым простым методом вычисления данного интеграла является метод ячеек. Пусть областью интегрирования является прямоугольник:

$$a \leq x \leq b, \quad c \leq y \leq d.$$

По теореме о среднем среднее значение функции $f(x, y)$:

$$\bar{f}(x, y) = \frac{1}{S} \iint_G f(x, y) dx dy, \quad S = (b - a)(d - c). \quad (75)$$

Предположим, что среднее значение приближенно равно значению функции в центре прямоугольника, т.е.

$$f(\bar{x}, \bar{y}) \approx \bar{f}(x, y),$$

тогда для вычисления интеграла получаем

$$\iint_G f(x, y) dx dy \approx S f(\bar{x}, \bar{y})$$

$$\bar{x} = \frac{a + b}{2}, \quad \bar{y} = \frac{c + d}{2}.$$

Точность формулы можно повысить, если разбить область G на прямоугольные ячейки ΔG_{ij} , то для каждой ячейки

$$\iint_{\Delta G_{ij}} f(x, y) dx dy \approx f(\bar{x}_i, \bar{y}_j) \Delta x_i \Delta y_j \quad (76)$$

и для интеграла

$$\iint_G f(x, y) dx dy \approx \sum_{i=1}^M \sum_{j=1}^N f(\bar{x}_i, \bar{y}_j) \Delta x_i \Delta y_j \quad (77)$$

В правой части последнего равенства стоит интегральная сумма, поэтому при уменьшении периметров ячеек эта сумма стремится к

значению интеграла для любой непрерывной функции $f(x, y)$. Погрешность метода ячеек имеет второй порядок малости по Δx и Δy . Для дальнейшего повышения точности можно применить метод сгущения узлов сетки. При этом шаг по каждой переменной уменьшают в одинаковое число раз, т.е. отношение M/N остается постоянным.

В случае, если область G прямоугольная, ее может быть целесообразно привести к прямоугольному виду путем соответствующей замены переменных.

Рассмотрим область – криволинейный четырехугольник: $a \leq x \leq b$, $\varphi_1(x) \leq y \leq \varphi_2(x)$. Для приведения этой области к прямоугольной, надо использовать замену:

$$t = \frac{y - \varphi_1(x)}{\varphi_2(x) - \varphi_1(x)}, \quad 0 \leq t \leq 1. \quad (78)$$

Другой возможный способ вычисления многомерных интегралов состоит в сведении их к последовательному вычислению одномерных интегралов.

В векторно-матричном виде система (79) запишется в виде

$$A\mathbf{x} = \mathbf{b} \quad (81)$$

где \mathbf{x} и \mathbf{b} – вектор-столбцы неизвестных и правых частей.

Иногда системы уравнений имеют некоторые специальные виды матриц:

- симметричная матрица $a_{ij} = a_{ji}$,
- треугольная матрица с равными нулю элементами, расположенными ниже или выше диагонали,
- клеточная матрица,
- ленточная матрица (пример – трехдиагональная),
- единичная матрица

Определителем матрицы A n -го порядка называется число D , рав-

ное

$$D = \det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum (-1)^k a_{1\alpha} a_{2\beta} \dots a_{n\omega}, \quad (82)$$

где индексы $\alpha, \beta, \dots, \omega$ пробегают все возможные $n!$ перестановок номеров $1, 2, \dots, n$, k – число инверсий (обмен двух соседних индексов местами) в данной перестановке.

Необходимым и достаточным условием существования единственного решения системы лин. уравнений является $D \neq 0$. В случае, если $D = 0$ – матрица называется вырожденной, система может не иметь решений, или иметь их бесконечное множество.

Для иллюстрации рассмотрим систему двух уравнений

$$a_1x + b_1y = c_1 \quad (83)$$

$$a_2x + b_2y = c_2. \quad (84)$$

Возможны три случая:

1) прямые пересекаются, т.е.

$$\frac{a_1}{a_2} \neq \frac{b_1}{b_2}$$

2) прямые параллельны, т.е.

$$\frac{a_1}{a_2} = \frac{b_1}{b_2} \neq \frac{c_1}{c_2}$$

3) прямые совпадают, т.е.

$$\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}$$

Определитель нашей системы имеет вид

$$D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$$

В практических вычислениях из-за округлений редко удастся получить точное равенство определителя нулю. При $D \approx 0$ прямые могут оказаться почти параллельными, и координаты точки пересечения могут быть весьма чувствительными к малым изменениям коэффициентов системы. Т.е., малые погрешности вычислений или исходных данных могут привести к существенным погрешностям в решении. В этом случае система уравнений называется плохо обусловленной.

6.2 Методы решения линейных систем

Методы решения линейных систем делятся на два класса:

1) **точные (прямые) методы**, которые представляют из себя конечные алгоритмы для вычисления корней системы; примеры: пра-

вило Крамера, метод Гаусса, метод главных элементов, метод квадратных корней и др.

2) **итерационные методы**, позволяющие получать корни системы с заданной точностью путем сходящихся бесконечных процессов; примеры: метод итераций, метод Гаусса-Зейделя, метод релаксации и т.д.

Точные методы используют конечные формулы для вычисления неизвестных. Плюсы:

- решение получается после заранее известного числа операций,
- универсальность и простота метода.

Минусы:

- в памяти машины должна находиться вся матрица, поэтому при больших матрицах это требует большого объема памяти.
- накопление ошибок округления, поскольку на любом этапе используются все предыдущие результаты. Поэтому даже точные методы на самом деле являются приближенными, при этом затруднительно

оценить погрешность.

Итерационные методы – методы последовательного приближения. Суть состоит в задании начального приближения и последовательности операций, после выполнения которой будут получаться следующие приближения. Плюсы:

- нет необходимости хранить в памяти машины всю матрицу,
- нет накопления погрешностей, поскольку точность вычисления в каждой итерации определяется лишь результатом предыдущей итерации.

Минусы:

- эффективность существенно зависит от удачного выбора начального приближения и быстроты сходимости метода.

6.3 Точные (прямые) методы

1. Формулы Крамера.

Для системы n линейных уравнений с n неизвестными

$$A\mathbf{x} = \mathbf{b} \quad (85)$$

с невырожденной матрицей A

$$\Delta = \det A \neq 0$$

единственное решение находится умножением (85) слева на матрицу A^{-1} , т.е.

$$A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} \quad \Rightarrow \quad \mathbf{x} = A^{-1}\mathbf{b}.$$

В курсе линейной алгебры показывается, что для неизвестных справедливы формулы (Крамера)

$$x_1 = \frac{\Delta_1}{\Delta}, \quad x_2 = \frac{\Delta_2}{\Delta}, \quad \dots, \quad x_n = \frac{\Delta_n}{\Delta} \quad (86)$$

где Δ_i – определители, получающиеся из определителя системы Δ заменой его i -го столбца столбцом свободных членов \mathbf{b} .

Например, для системы

$$a_{11}x + a_{12}y = b_1 \quad (87)$$

$$a_{21}x + a_{22}y = b_2. \quad (88)$$

неизвестные записываются следующим образом:

$$x = \Delta_1/\Delta, \quad y = \Delta_2/\Delta, \quad (89)$$

где

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \Delta_1 = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}$$

Т.о., решение системы из n уравнений с n неизвестными сводится к вычислению $n + 1$ определителя порядка n . Т.к. вычисление определителей большого порядка требует выполнения большого объема арифметических операций, то применение данного метода ограничено.

2. Метод Гаусса.

Алгоритм последовательного исключения неизвестных носит название метода Гаусса. В основе метода лежит приведение матрицы системы к треугольному виду. Этого можно достичь, последовательно исключая неизвестные из уравнений:

- с помощью первого уравнения исключаем x_1 из второго и всех последующих уравнений системы;
- с помощью второго уравнения исключаем x_2 из третьего и всех последующих уравнений системы;

.....

- продолжаем до тех пор, пока в левой части последнего (n -го) уравнения не останется одно слагаемое с x_n . В результате матрица будет приведена к треугольному виду. Данный процесс носит название прямого хода метода Гаусса.

Далее решая последнее уравнение, находим x_n и подставляем его в предыдущее уравнение. Затем решаем его, находим x_{n-1} и т.д. В

этом состоит обратный ход метода Гаусса.

Рассмотрим систему:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \quad (90)$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \quad (91)$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3. \quad (92)$$

Для того, чтобы исключить x_1 из второго уравнения, прибавим к нему первое, умноженное на

$$-\frac{a_{21}}{a_{11}}.$$

Далее, умножим первое уравнение на

$$-\frac{a_{31}}{a_{11}}$$

и сложим с третьим. В результате у нас получится система уравне-

ний

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \quad (93)$$

$$a'_{22}x_2 + a'_{23}x_3 = b'_2, \quad (94)$$

$$a'_{32}x_2 + a'_{33}x_3 = b'_3, \quad (95)$$

где

$$a'_{ij} = a_{ij} - \frac{a_{i1}}{a_{11}}a_{1j}, \quad i, j = 2, 3, \quad (96)$$

$$b'_i = b_i - \frac{a_{i1}}{a_{11}}b_1, \quad i = 2, 3. \quad (97)$$

Далее исключаем из третьего уравнения x_2 . Для этого надо умножить второе уравнение на

$$-\frac{a'_{32}}{a'_{22}}$$

и сложить с третьим. В результате

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \quad (98)$$

$$a'_{22}x_2 + a'_{23}x_3 = b'_2, \quad (99)$$

$$a''_{33}x_3 = b''_3, \quad (100)$$

где

$$a''_{ij} = a'_{ij} - \frac{a'_{i2}}{a'_{22}}a'_{2j}, \quad i, j = 3, \quad (101)$$

$$b''_i = b'_i - \frac{a'_{i2}}{a'_{22}}b'_2, \quad i = 3. \quad (102)$$

Далее начинается обратный ход метода.

В процессе исключения неизвестных необходимо выполнять операцию деления на коэффициенты a_{11} , a'_{22} (ведущие элементы). Они должны быть отличны от нуля. Если это не так, то необходимо соответствующим образом переставить уравнения системы, что должно быть предусмотрено в вычислительном алгоритме.

Модификацией метода Гаусса является схема с выбором ведущего элемента. В ней требование неравенства нулю диагональных элементов a_{ii} заменяется более жестким: из всех оставшихся в i -ом столбце элементов нужно выбрать наибольший по модулю и переставить уравнения так, чтобы этот элемент оказался на месте элемента a_{ii} . Благодаря этому уменьшаются множители, используемые для преобразования уравнений, что уменьшает погрешности вычислений. Модифицированный метод Гаусса обеспечивает приемлемую точность для систем с $n \lesssim 1000$. Число арифметических действий в методе Гаусса $\lesssim n^3$, т.е. время, необходимое для решения, $\sim n^3$.

Уточнение корней

Полученные методом Гаусса приближенные значения корней можно уточнить. Пусть для системы

$$A\mathbf{x} = \mathbf{b}$$

найденно приближенное решение \mathbf{x}_0 . Положим

$$\mathbf{x} = \mathbf{x}_0 + \boldsymbol{\delta},$$

где

$$\boldsymbol{\delta} = \begin{pmatrix} \delta_1, \\ \delta_2, \\ \dots \\ \delta_n \end{pmatrix}.$$

Тогда находим

$$A(\mathbf{x}_0 + \boldsymbol{\delta}) = \mathbf{b}$$

или

$$A\boldsymbol{\delta} = \boldsymbol{\varepsilon},$$

где $\boldsymbol{\varepsilon} = \mathbf{b} - A\mathbf{x}_0$ – невязка для решения \mathbf{x}_0 . Т.о. для нахождения $\boldsymbol{\varepsilon}$ нужно решить линейную систему с прежней матрицей и другой правой частью.

3. Метод прогонки.

Метод прогонки представляет собой модификацию метода Гаусса для систем уравнений с трехдиагональной матрицей. Запишем такую систему уравнений:

$$b_1x_1 + c_1x_2 = d_1, \quad (103)$$

$$a_2x_1 + b_2x_2 + c_2x_3 = d_2, \quad (104)$$

$$a_3x_2 + b_3x_3 + c_3x_4 = d_3, \quad (105)$$

$$\dots\dots\dots (106)$$

$$a_{n-1}x_{n-2} + b_{n-1}x_{n-1} + c_{n-1}x_n = d_{n-1}, \quad (107)$$

$$a_nx_{n-1} + b_nx_n = d_n. \quad (108)$$

Прямая прогонка состоит в нахождении прогоночных коэффициентов A_i и B_i , с помощью которых неизвестное x_i выражается через x_{i+1} :

$$x_i = A_ix_{i+1} + B_i, \quad i = 1, 2, \dots, n - 1. \quad (109)$$

Из первого уравнения системы находим

$$x_1 = -\frac{c_1}{b_1}x_2 + \frac{d_1}{b_1}$$

с другой стороны

$$x_1 = A_1x_2 + B_1.$$

Сравнивая, находим

$$A_1 = -\frac{c_1}{b_1}, \quad B_1 = \frac{d_1}{b_1}. \quad (110)$$

Подставим во второе уравнение системы выражение для x_1 :

$$a_2(A_1x_2 + B_1) + b_2x_2 + c_2x_3 = d_2$$

Отсюда,

$$x_2 = \frac{-c_2x_3 + d_2 - a_2B_1}{a_2A_1 + b_2}$$

или

$$x_2 = A_2x_3 + B_2$$

$$A_2 = -\frac{c_2}{e_2}, \quad B_2 = \frac{d_2 - a_2 B_1}{e_2}, \quad e_2 = a_2 A_1 + b_2.$$

Последние соотношения можно обобщить для произвольного номера i :

$$A_i = -\frac{c_i}{e_i}, \quad B_i = \frac{d_i - a_i B_{i-1}}{e_i}, \quad e_i = a_i A_{i-1} + b_i, \quad i = 2, \dots, n-1 \quad (111)$$

На этом заканчивается прямая прогонка. При обратной прогонке последовательно вычисляются неизвестные x_i . Для этого сначала запишем

$$x_{n-1} = A_{n-1} x_n + B_{n-1}$$

и последнее уравнение системы

$$a_n x_{n-1} + b_n x_n = d_n.$$

Из двух последних уравнений находим x_n :

$$x_n = \frac{d_n - a_n B_{n-1}}{b_n + a_n A_{n-1}}$$

Далее по прогоночным формулам находим все остальные неизвестные.

6.4 Итерационные методы

1. Метод простой итерации.

Исходная система уравнений

$$A\mathbf{x} = \mathbf{b}$$

Отсюда

$$0 = \mathbf{b} - A\mathbf{x}, \quad \mathbf{x} = \mathbf{b} - A\mathbf{x} + \mathbf{x}, \quad \mathbf{x} = (\mathbf{b} - A\mathbf{x})\tau + \mathbf{x}$$

$$\mathbf{x} = (E - \tau A)\mathbf{x} + \tau\mathbf{b},$$

$$\mathbf{x} = B\mathbf{x} + \tau\mathbf{b}, \quad \text{где} \quad B = E - \tau A \quad (112)$$

Получившаяся система (112) является основой метода простой итерации. Выбираем некоторое начальное приближение $\mathbf{x}^{(0)}$, подставля-

ем в правую часть (112), получаем следующее (1-ое) приближение:

$$\mathbf{x}^{(1)} = B\mathbf{x}^{(0)} + \tau\mathbf{b}$$

В результате

$$\mathbf{x}^{(k+1)} = B\mathbf{x}^{(k)} + \tau\mathbf{b}, \quad k = 0, 1, 2, \dots \quad (113)$$

Формула (113) выражает метод простой итерации. От выбора параметра τ зависит его сходимость и скорость сходимости.

2. Метод Гаусса-Зейделя.

Метод Г-З представляет собой итерационный метод. Рассмотрим его на примере системы

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1, \quad (114)$$

$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2, \quad (115)$$

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3. \quad (116)$$

Выразим неизвестные из уравнений

$$x_1 = \frac{1}{a_{11}}(b_1 - a_{12}x_2 - a_{13}x_3), \quad (117)$$

$$x_2 = \frac{1}{a_{22}}(b_2 - a_{21}x_1 - a_{23}x_3), \quad (118)$$

$$x_3 = \frac{1}{a_{33}}(b_3 - a_{31}x_1 - a_{32}x_2), \quad (119)$$

Далее зададим некоторое нулевой приближение для неизвестных $(x_1^{(0)}, x_2^{(0)}, x_3^{(0)})$. Подставляя в формулу для x_1 находим $x_1^{(1)}$ (первое приближение для x_1). При вычислении $x_2^{(1)}$ вместо $x_1^{(0)}$ будем использовать $x_1^{(1)}$. В результате найдем $x_2^{(1)}$. Далее находим $x_3^{(1)}$. И далее повторяем вычисления. В общем случае k -ое приближение бу-

дет выражаться через $k - 1$ приближение следующим образом

$$x_1^{(k)} = \frac{1}{a_{11}}(b_1 - a_{12}x_2^{(k-1)} - a_{13}x_3^{(k-1)}), \quad (120)$$

$$x_2^{(k)} = \frac{1}{a_{22}}(b_2 - a_{21}x_1^{(k)} - a_{23}x_3^{(k-1)}), \quad (121)$$

$$x_3^{(k)} = \frac{1}{a_{33}}(b_3 - a_{31}x_1^{(k)} - a_{32}x_2^{(k)}), \quad (122)$$

Итерационный процесс необходимо продолжать до тех пор, пока значения неизвестных в двух последовательных приближениях не станут близкими с заданной погрешностью.

6.5 Задачи на собственные значения

Рассмотрим квадратную матрицу n -го порядка:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \quad (123)$$

Вектор

$$\mathbf{x} = \{x_1, x_2, \dots, x_n\}$$

называется собственным вектором матрицы A , соответствующим собственному значению λ , если он удовлетворяет системе уравнений:

$$A\mathbf{x} = \lambda\mathbf{x}. \quad (124)$$

Характеристической матрицей C данной матрицы A называется матрица

$$C = A - \lambda E = \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix} \quad (125)$$

где E – единичная матрица. В результате мы получаем, что систему (124) можно записать в виде

$$(A - \lambda E)\mathbf{x} = 0 \quad \text{или} \quad C\mathbf{x} = 0 \quad (126)$$

Полученная система является однородной системой n линейных уравнений с n неизвестными. Ненулевые решения наша система имеет только в случае

$$\det C = 0.$$

Определитель $\det C$ является многочленом n -й степени относительно λ :

$$\det C = c_0\lambda^n + c_1\lambda^{n-1} + \dots + c_{n-1}\lambda + c_n.$$

Этот многочлен называется характеристическим многочленом и его корни являются собственными значениями матрицы A .

7 Обыкновенные дифференциальные уравнения

ОДУ называются такие уравнения, которые содержат одну или несколько производных от искомой функции $y = y(x)$:

$$F(x, y, y', \dots, y^{(n)}) = 0 \quad (127)$$

Наивысший порядок n входящей в уравнение производной называется порядком дифференциального уравнения. Уравнение первого порядка

$$F(x, y, y') = 0$$

второго порядка

$$F(x, y, y', y'') = 0$$

Если из уравнения удастся выразить старшую производную и привести его к виду

$$y^{(n)} = F(x, y, y', \dots, y^{(n-1)})$$

то такая форма записи называется уравнением, разрешенным относительно старшей производной.

Линейным дифференциальным уравнением называется уравнение, линейное относительно искомой функции и ее производных.

Решением дифференциального уравнения называется всякая n раз дифференцируемая функция $y = \varphi(x)$, которая после подстановки в уравнение превращает его в тождество.

Общее решение уравнения n -го порядка содержит n произвольных постоянных C_1, C_2, \dots, C_n :

$$y = \varphi(x, C_1, C_2, \dots, C_n) \quad (128)$$

Причем любое решение нашего уравнения можно представить в виде (128) при определенных значениях констант.

Для уравнения первого порядка общее решение зависит от одной произвольной постоянной:

$$y = \varphi(x, C)$$

Если постоянной придать определенное значение $C = C_0$, то получаем частное решение

$$y = \varphi(x, C_0).$$

Геометрическая интерпретация общего решения уравнения первого порядка состоит в том, что оно описывает бесконечное семейство интегральных кривых $y = y(x, C)$ с параметром C , а частному решению соответствует одна кривая этого семейства, причем через каждую точку (x_0, y_0) проходит одна и только одна интегральная кривая.

Для уравнений высших порядков – через каждую точку (x_0, y_0) проходит не одна интегральная кривая, поэтому для выделения некоторого частного решения надо задать дополнительные условия по числу произвольных постоянных.

7.1 Приближенные методы решения ОДУ. Метод Эйлера

Наиболее универсальным методом решения ОДУ является метод конечных разностей. Он заключается в том, что область непрерывного изменения аргумента заменяется дискретным множеством точек – узлами, образующими разностную сетку. Искомая функция непрерывного аргумента заменяется функцией дискретного аргумента на данной сетке – сеточной функцией. Решение ОДУ сводится к отысканию значений сеточной функции в узлах сетки.

Например, дифференциальное уравнение (задача Коши)

$$\frac{dy}{dx} = f(x, y), \quad y(x_0) = A$$

вводя равномерную сетку с шагом h и приняв в качестве узлов сетки $x_0, x_1 = x_0 + h, \dots$, можно свести к разностному, приближая производную в i -ом узле с помощью правых разностей

$$\frac{y_{i+1} - y_i}{h} = f_i, \quad i = 0, 1, \dots, \quad y_0 = A$$

Здесь $f_i = f(x_i, y_i)$. Таким образом мы определили разностную схему. Разностной схемой называется замкнутая система разностных уравнений вместе с дополнительными условиями – начальными или краевыми.

Выражая y_{i+1} , получаем

$$y_{i+1} = y_i + hf_i = y_i + hf(x_i, y_i), \quad i = 0, 1, \dots \quad (129)$$

В результате приходим к методу Эйлера:

$$\begin{aligned} y_1 &= y_0 + hf(x_0, y_0), \\ y_2 &= y_1 + hf(x_1, y_1), \\ &\dots\dots\dots \\ y_n &= y_{n-1} + hf(x_{n-1}, y_{n-1}), \\ &\dots\dots\dots \end{aligned}$$

Очевидно, что в данном случае искомая интегральная кривая приближается ломаной, наклон которой на элементарном участке $[x_i, x_{i+1}]$ определяется наклоном интегральной кривой исходного уравнения,

выходящей из точки (x_k, y_k) .

Неявный метод Эйлера состоит в приближении производной в окрестности i -го узла с помощью левой разности

$$\frac{y_i - y_{i-1}}{h} = f_i, \quad i = 0, 1, \dots, \quad y_0 = A$$

Очевидно, что в этом методе искомая величина y_i входит в уравнение в общем случае нелинейным образом. Поэтому необходимо применять известные методы решения нелинейных уравнений.

Можно показать, что явный и неявный метод – методы первого порядка точности. Т.е., погрешность в точке x_i

$$\delta_i = y(x_i) - y_i$$

равная разности между точным значением искомой функции $y(x_i)$ и значением сеточной функции в узле

$$\delta_i \lesssim h.$$

Можно построить другие методы решения задачи Коши. Например, приблизим производную в окрестности i -го узла с помощью

центральных разностей:

$$\frac{y_{i+1} - y_{i-1}}{2h} = f_i, \quad i = 0, 1, \dots, \quad y_0 = A$$

Полученная система уравнений не замкнута – необходимо доопределить y_1 , что можно сделать с помощью метода Эйлера

$$y_1 = y_0 + hf(x_0, y_0).$$

Можно показать, что построенная схема имеет второй порядок точности.

Рассмотрим модифицированный метод Эйлера. Запишем систему уравнений

$$\frac{y_{i+1} - y_i}{h} = f_{i+1/2}, \quad i = 0, 1, \dots, \quad y_0 = A$$

где

$$f_{i+1/2} = f(x_{i+1/2}, y_{i+1/2}) = f\left(x_i + \frac{h}{2}, y_{i+1/2}\right),$$

а значение $y_{i+1/2}$ вычисляем по методу Эйлера

$$y_{i+1/2} = y_i + \frac{h}{2}f_i = y_i + \frac{h}{2}f(x_i, y_i), \quad i = 0, 1, \dots$$

В результате получаем следующую разностную схему:

$$\frac{y_{i+1} - y_i}{h} = f\left(x_i + \frac{h}{2}, y_i + \frac{h}{2}f(x_i, y_i)\right), \quad i = 0, 1, \dots, \quad y_0 = A$$

которая, как можно показать, имеет второй порядок точности.

7.2 Метод Эйлера с пересчетом (предиктор-корректор)

Предиктор – предсказание результата, корректор – уточнение результата. Заменяем правую часть нашего уравнения на среднее значений в соседних узлах:

$$\frac{y_{i+1} - y_i}{h} = \frac{1}{2} [f(x_i, y_i) + f(x_{i+1}, y_{i+1})]. \quad (130)$$

Как видно, наклон интегральной кривой посередине отрезка $[x_i, x_{i+1}]$ приближенно заменяется средним арифметическим наклонов на гра-

ницах этого отрезка. Отсюда,

$$y_{i+1} = y_i + \frac{h}{2} [f(x_i, y_i) + f(x_{i+1}, y_{i+1})]. \quad (131)$$

Искомое значение y_{i+1} входит и в правую сторону тоже, и его нельзя в общем случае выразить явно. Но его можно найти по формуле метода Эйлера:

$$y_{i+1} = y_i + hf(x_i, y_i) \quad - \text{ предиктор} \quad (132)$$

и подставляя в правую часть, получаем

$$y_{i+1} = y_i + \frac{h}{2} [f(x_i, y_i) + f(x_{i+1}, (y_i + hf(x_i, y_i)))] \quad - \text{ корректор.} \quad (133)$$

7.3 Методы Рунге-Кутты

Рассмотренные ранее метод Эйлера и его модификация относятся к классу методов Рунге – Кутты. Суть методов – для вычисления y_{i+1}

используется y_i , а также значения функции $f(x, y)$ в некоторых специальных точках. Широко распространен метод Рунге-Кутты четвертого порядка. Алгоритм этого метода имеет вид

$$y_{i+1} = y_i + \frac{h}{6}(k_0 + 2k_1 + 2k_2 + k_3), \quad i = 0, 1, \dots$$

$$k_0 = f(x_i, y_i), \quad k_1 = f\left(x_i + \frac{h}{2}, y_i + \frac{hk_0}{2}\right)$$

$$k_2 = f\left(x_i + \frac{h}{2}, y_i + \frac{hk_1}{2}\right)$$

$$k_3 = f(x_i + h, y_i + hk_2)$$

7.4 Многоточечные методы

Многоточечные или многошаговые методы решения задачи Коши отличаются тем, что решение в текущем узле зависит от данных не только в одном предыдущем узле, но и в ряде предшествующих.

Такие методы можно строить, используя метод неопределенных коэффициентов. Для этого запишем разностное соотношение для i -го узла в виде

$$\frac{a_0 y_i + a_1 y_{i-1} + \dots + a_m y_{i-m}}{h} = b_0 f_i + b_1 f_{i-1} + \dots + b_l f_{i-l}. \quad (134)$$

Коэффициенты a_i и b_i подбираются так, чтобы получить аппроксимацию с нужным порядком точности. Если $a_0 = -a_1 = 1$, $a_2 = \dots = a_m = 0$, то получаем методы Адамса. Уравнение $y' = f(x, y)$ эквивалентно интегральному соотношению

$$y_{i+1} - y_i = \int_{x_i}^{x_{i+1}} f \, dx. \quad (135)$$

Если решение в узлах вплоть до i -го уже вычислено, то по известным значениям $f_i = f(x_i, y_i)$ можно интерполировать подинтегральную функцию полиномами различной степени. Далее вычисляя интеграл от выбранного интерполяционного полинома, получаем фор-

мулы Адамса.

Приближая функцию f на отрезке $[x_i, x_{i+1}]$ полиномом в форме Ньютона

$$f = f_i + \frac{f_i - f_{i-1}}{h}(x - x_i) + \frac{f_i - 2f_{i-1} + f_{i-2}}{2h^2}(x - x_i)(x - x_{i-1}) + \dots \quad (136)$$

и учитывая первые два слагаемых при вычислении интеграла, получаем метод Адамса второго порядка точности

$$\frac{y_{i+1} - y_i}{h} = \frac{3}{2}f_i - \frac{1}{2}f_{i-1}. \quad (137)$$

Учитывая в (136) три слагаемых приходим к методу третьего порядка точности

$$\frac{y_{i+1} - y_i}{h} = \frac{1}{12}(23f_i - 16f_{i-1} + 5f_{i-2}). \quad (138)$$

и случае четырех слагаемых в (136) метод Адамса четвертого поряд-

ка

$$\frac{y_{i+1} - y_i}{h} = \frac{1}{24}(55f_i - 59f_{i-1} + 37f_{i-2} - 9f_{i-3}). \quad (139)$$

Метод Адамса по сравнению с методом Рунге-Кутты требует меньших затрат при нахождении очередного значения y_{i+1} , т.к. требуется найти лишь f_i , а f_{i-1}, \dots, f_{i-3} уже известны к этому моменту, а по формула Рунге-Кутты на каждом шаге надо находить четыре значения f .

Формулы Рунге-Кутты позволяют проводить вычисления с переменным шагом. Ну и при использовании многошаговых методов может неконтролируемо возрастать погрешность вычислений, т.е. возникать неустойчивость.