

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

Государственное образовательное учреждение
высшего профессионального образования
«Оренбургский государственный университет»

В.Н. КОЗЛОВ, А.И. АКИМОВ, В.В. ТУГОВ

ОСНОВЫ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ

Рекомендовано к изданию Ученым советом
государственного образовательного учреждения
высшего профессионального образования
«Оренбургский государственный университет»
в качестве учебного пособия для студентов
направления 220100 – Системный анализ и управление

Оренбург 2008

УДК 519.6(075.8)
ББК 22.19я73
К 59

Рецензент:

кандидат технических наук, доцент Глинская Н.Ю.

Козлов В.Н.
К 59 **Основы вычислительной математики [Текст]: учебное пособие/ В.Н. Козлов, А.И. Акимов, В.В. Тугов. - Оренбург : ГОУ ОГУ, 2008. - 111 с.**

В учебном пособии рассматриваются основные понятия и методы вычислительной математики. Излагаемый материал сопровождается примерами и алгоритмами вычислений, а также заданиями для самостоятельного решения.

Учебное пособие предназначено для студентов направления 220100 – Системный анализ и управление, а также для 230100 - [Информатика и вычислительная техника и др.](#)

Е 160212000

ББК 22.19я73

© Козлов Н.И
Акимов А.И.,
Тугов В.В., 2008
© ГОУ ОГУ, 2008

Содержание

Введение.....	9
1 Численные методы математического анализа.....	10
1.1 Погрешность результата численного решения задач.....	10
1.2 Вычисление значений функций.....	11
1.2.1 Вычисление полиномов. Схема Горнера.....	11
1.2.2 Вычисление значения функции с помощью рядов.....	12
1.3 Приближенное решение уравнений.....	15
1.3.1 Метод дихотомии (деления отрезка пополам).....	16
1.3.2 Метод хорд.....	17
1.3.3 Метод Ньютона (метод касательных).....	18
1.3.4 Метод секущих.....	21
1.3.5 Метод простых итераций (последовательных приближений).....	21
1.4 Интерполирование и аппроксимация функций (приближение функций).....	25
1.4.1 Понятие сетки.....	26
1.4.2 Конечные разности и разностные отношения.....	27
1.4.3 Понятие интерполирования функций.....	29
1.4.3.1 Полиномиальная интерполяция.....	31
1.4.3.2 Интерполяционный полином Лагранжа.....	32
1.4.3.3 Интерполяционная формула Эйткина.....	34
1.4.3.4 Интерполяционная формула Ньютона.....	38
1.4.4 Обратное интерполирование.....	40
1.5 Численное дифференцирование.....	49
1.6 Численное интегрирование.....	52
1.6.1 Понятие квадратурной формулы.....	52
1.6.2 Интерполяционные квадратуры.....	53
1.6.3 Вычисление несобственных интегралов.....	59
2 Решение систем алгебраических уравнений.....	62
2.1 Решение систем линейных уравнений.....	62
2.1.1 Основные понятия.....	62
2.1.2 Метод Крамера.....	63
2.1.3 Метод исключения Гаусса.....	64
2.1.4 Метод исключения Гаусса с выбором главного элемента.....	67
2.1.5 Матричная формулировка метода Гаусса.....	68
2.1.6 Вычисление обратной матрицы и значения определителя.....	69
2.1.7 Метод исключения для ленточных матриц. Прогонка.....	70
2.1.8 Метод простых итераций.....	72
2.1.9 Метод Зейделя.....	73
2.2 Решение систем нелинейных уравнений.....	77
2.2.1 Метод простых итераций.....	78
2.2.2 Метод Зейделя.....	80
2.2.3 Метод Ньютона-Рафсона.....	81
2.2.4 Метод минимизации квадратичной функции.....	83
3 Минимизация функций.....	85
3.1 Одномерная минимизация.....	85
3.1.1 Метод сканирования.....	85
3.1.2 Метод золотого сечения.....	87
3.1.3 Метод парабол.....	89
3.2 Многомерная минимизация.....	91
3.2.1 Метод сканирования.....	92
3.2.2 Спуск по координатам.....	94
3.2.3 Метод градиента.....	94

3.2.4 Слепой поиск.....	98
3.2.5 Метод случайных направлений.....	98
3.2.6 Условная минимизация. Метод штрафных функций.....	98
4 Решение обыкновенных дифференциальных уравнений.....	101
4.1 Метод Эйлера (Рунге-Кутта первого порядка).....	101
4.2 Модифицированный метод Эйлера (Рунге-Кутта второго порядка).....	103
4.3 Метод Рунге-Кутта четвертого порядка.....	104
4.4 Метод прогноза и коррекции.....	106
4.5 Решение краевых задач. Метод стрельб.....	106
5 Решение уравнений в частных производных	109
5.1 Построение аппроксимирующих разностных схем.....	109
5.2 Решение задачи Коши.....	110
5.3 Решение уравнения теплопроводности (диффузии).....	115
Список использованных источников.....	118

Введение

Анализ инженерно-технических задач нередко сводится к решению математических задач различной сложности. Для решения математических задач используют графические, аналитические и численные методы.

При графическом анализе задачи на основе геометрических построений получают оценочные решения. Эти решения являются грубыми, содержащими большую погрешность. Поэтому графические методы используют зачастую лишь для предварительного определения области существования решения.

Аналитические методы позволяют выразить решение задачи в виде конечных математических формул. Если такие формулы для решения задачи имеются, то, подставив в них соответствующие исходные данные, после конечного числа математических операций получают числовое решение задачи. Однако на практике такие случаи крайне редки.

Наиболее часто при решении указанных задач используют численные методы, рассматриваемые в курсе вычислительной математики. Вычислительная математика, называемая также прикладной, специализируется на разработке алгоритмов решения математических задач различной сложности, сопоставительной оценке эффективности алгоритмов, определения условий достижения решений и др.

По типу численные методы подразделяются на прямые, итерационные и вероятностные. *Прямые методы* - методы, в которых решение получают за конечное число арифметических действий. При этом ошибки округления, накапливаясь, могут исказить результат. В *итерационных методах* решение получают в виде предела сходящейся бесконечной последовательности. В этих методах ошибки округления почти не влияют на результат, но имеется проблема обеспечения сходимости решения. С развитием вычислительной техники итерационные методы находят все более широкое применение. *Вероятностные методы* - это в основном методы случайного поиска. Они эффективно используются при решении задач больших размерностей.

В учебном пособии рассматриваются в основном методы первых двух типов. Для большинства методов приведены примеры и схемы алгоритмов. Причем с целью обеспечения наглядности часть схем представлена в упрощенной форме.

1 Численные методы математического анализа

1.1 Погрешность результата численного решения задач

При математическом моделировании можно выделить 4 источника погрешности результата решения задач:

1) неточная математическая модель. Она связана с неточностью информации о решаемой задаче или с грубостью принятых физических допущений при постановке задачи;

2) неточные исходные данные. Часто исходные данные являются экспериментальными, определенными с некоторой погрешностью. Данная погрешность зависит от чистоты подготовки и проведения эксперимента.

3) погрешность используемого численного метода решения задачи. Она связана с заменой исходной задачи другой, более простой.

4) ошибки округления. Они накапливаются в ходе вычислений и могут сильно исказить результат.

Различают абсолютную и относительную погрешности.

Абсолютная погрешность: $\Delta = |a^* - a|$, где a^* , a - точное и приближенное значения числа.

Относительная погрешность: $\delta = |(a^* - a) \div a^*|$.

Связь между абсолютной и относительной погрешностями $\Delta = |a^*| \delta$.

Для оценки погрешности результата можно использовать следующие приемы:

а) решить задачу двумя методами и результаты сопоставить;

б) незначительно изменить исходные данные и повторно осуществить расчет. Результаты также сопоставить.

Корректность задач

Задача $y = A(x)$ называется *корректной*, если для любых входных данных x решение y существует, единственно и устойчиво по входным данным.

В результате погрешности входных данных $x + \Delta x$ имеем погрешность решения $\Delta y = A(x + \Delta x) - A(x)$. Если решение непрерывно зависит от входных данных, т.е. при $\Delta x \rightarrow 0$ следует $\Delta y \rightarrow 0$, то задача называется *устойчивой по входным данным*, в противном случае - неустойчивой по входным данным.

Отсутствие устойчивости (некорректность задачи) означает, что даже небольшой погрешности Δx соответствует большое значение Δy . Для таких задач применять численные методы бессмысленно - в ходе вычисления погрешности будут катастрофически возрастать.

Возможны случаи слабой устойчивости задачи (плохая обусловленность): $\Delta y \leq C \Delta x$, причем C - большая константа. В таких случаях численное решение задачи может существенно отличаться от истинного решения.

Аналогичным образом можно говорить о корректности алгоритма.

Вопросы и упражнения для самоконтроля.

1. Назовите основные источники погрешности результата решения задачи.
2. Что такое абсолютная и относительная погрешности? Какая связь между ними?
3. Какая задача называется корректной?
4. Найти абсолютные и относительные погрешности числа $\pi = 3,141592654\dots$, заданного тремя и двумя цифрами после запятой.
5. Дана функция $f(a,b,c)$. Оценить погрешности результата вычислений заданного двумя цифрами после запятой для:

а) $f(a,b,c) = \frac{a}{a^2 + bc}$, $a = 0,0034 \cdot N$, $b = 0,456 \cdot N$, $c = 0,0213 \cdot N$;

б) $f(a,b,c) = \frac{a\bar{n} - bc}{a^2 + b^2}$, $a = 4,345 \cdot N$, $b = 4,748 \cdot N$, $c = 7,89 \cdot N$;

в) $f(a,b,c) = \frac{a + c}{a^2 - b^2}$, $a = 23,567 \cdot N$, $b = 25,897 \cdot N$, $c = 17,432 \cdot N$;

где N – параметр, соответствующий списку группы.

1.2 Вычисление значений функций

Часто при вычислениях используют значения элементарных функций. К элементарным функциям относят степенные функции, многочлены, алгебраические функции, логарифмические, тригонометрические функции и др. При ручных вычислениях можно пользоваться различными таблицами, которых в настоящее время в достаточном количестве. При расчетах на ЭВМ расчет простых элементарных функций производится по стандартным подпрограммам, заложенным в память машины. Однако при расчетах на микрокалькуляторах возникает проблема использования по возможности точного и экономного метода их расчета.

1.2.1 Вычисление полиномов. Схема Горнера

Рассмотрим полином

$$P(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n \quad (1.1)$$

Из курса алгебры известна теорема Безу.

Теорема. Остаток от деления полинома $P(x)$ на двучлен $(x-x_0)$ равен значению полинома $P(x)$ при $(x = x_0)$:

$$P(x) = Q(x)(x-x_0) + B_n,$$

где $B_n = P(x_0)$.

Пусть $Q(x) = b_0 x^{n-1} + \dots + b_{n-2} x + b_{n-1}$, тогда

$$P(x) = (b_0 x^{n-1} + \dots + b_{n-2} x + b_{n-1})(x-x_0) + B_n.$$

Раскрывая скобки и сравнивая коэффициенты при одинаковых степенях x , получим:

$$\begin{array}{lcl}
 B_0 = a_0 & & B_0 = a_0 \\
 B_1 - B_0 x_0 = a_1 & \Rightarrow & B_1 = a_1 + B_0 x_0 \\
 B_2 - B_1 x_0 = a_2 & & B_2 = a_2 + B_1 x_0 \\
 \dots & & \dots \\
 B_{n-1} - B_{n-2} x_0 = a_{n-1} & & B_{n-1} = a_{n-1} + B_{n-2} x_0 \\
 B_n - B_{n-1} x_0 = a_n & & B_n = a_n + B_{n-1} x_0.
 \end{array} \quad (1.2)$$

Из последней системы, последовательно вычисляя b_i находим значение b_n . Если в формулах (1.2) подставить каждую предыдущую в последующую формулу, то получим *схему Горнера*. Алгоритм метода приведен на рисунке 1.1.

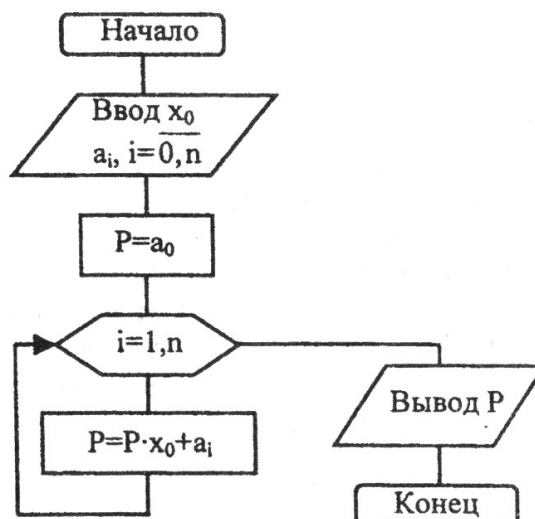


Рисунок 1.1 - Алгоритм схемы Горнера

$$P(x_0) = B_n = (\dots((a_0 x_0 + a_1) x_0 + a_2) x_0 + \dots a_{n-1}) x_0 + a_n. \quad (1.3)$$

Этот метод вычисления точный и ограничивается лишь применением конечного числа арифметических действий.

Пример. Вычислить значение полинома $P(x) = 3x^3 + 2x^2 - 5x + 7$ при $x_0 = 3$:
 $P(3) = ((3 \cdot 3 + 2) \cdot 3 - 5) \cdot 3 + 7 = (11 \cdot 3 - 5) \cdot 3 + 7 = 84 + 7 = 91.$

1.2.2 Вычисление значения функции с помощью рядов

Использование степенных рядов позволяет свести задачу вычисления значений функции к задаче вычисления значений полинома и, таким образом, ограничиться лишь выполнением арифметических операций.

Разложение функций в степенные ряды можно осуществить с помощью ряда Тейлора:

$$f(x) = f(a) + \frac{f^{(1)}(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \dots$$

Данное разложение точное. На практике обычно суммирование ряда осуществляют до некоторого n -го члена, отбрасывая члены малого порядка, начиная с $n+1$ -го. При этом разность

$$R_n(x) = f(x) - \sum_{k=0}^n \frac{f^{(k)}(a)}{k!}(x-a)^k \quad (1.4)$$

называется *остаточным членом* и представляет собой ошибку замены функции $f(x)$ конечным рядом. Это погрешность метода.

Из математического анализа известно, что

$$R_n(x) = \frac{f^{(n+1)}(a + \theta(x-a))}{(n+1)!}(x-a)^{n+1}, \quad 0 < \theta < 1. \quad (1.5)$$

Оценка (1.4) или (1.5) называется *априорной оценкой* точности (априори - до опыта). Такую оценку можно проводить до начала вычислений. Оценка же по первому отброшенному члену близка к оценке (1.5), но менее точна. Её называют *апостериорной оценкой* (апостериори - из опыта):

$$\tilde{R}_n(x) = \frac{f^{(n+1)}(a)}{(n+1)!}(x-a)^{n+1}. \quad (1.6)$$

Оценку (1.5) используют при теоретических исследованиях, оценку (1.6) - при практических вычислениях.

Пример. Вычислить значение e^2 с точностью 0,01.

Разложим функцию e^x в ряд в точке, $a=0$:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots$$

Для оценки точности вычислений используем апостериорную оценку:

$$\tilde{R}_n(x) = \frac{x^{n+1}}{(n+1)!}.$$

Так как $|\tilde{R}_n(x)| < 0,01$, то должно выполняться неравенство $|x^{n+1}| < 0,01 \cdot (n+1)!$.

При $n=6$ имеем $\tilde{R}_6 = 2^7/7! = 0,026$. Заданная точность не обеспечивается.

При $n=7$ $\tilde{R}_7(2) = 2^8/8! = 0,006$. Заданная точность обеспечивается.

Следовательно, для вычисления значения функции с точностью 0,01 необходимо просуммировать первые 8 членов ряда. Априорная же оценка будет определяться остаточным членом ряда:

$$R_7(2) = \frac{2^8}{8!} + \frac{2^9}{9!} + \dots$$

Графическая иллюстрация оценок примера приведена на рисунке 1.2.

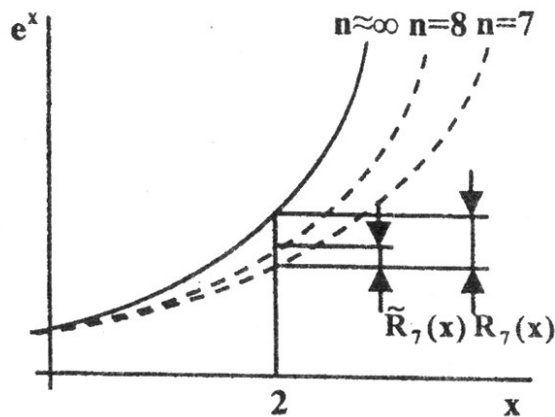


Рисунок 1.2 - Графическая иллюстрация оценок

Существуют различные способы вычисления значения функций. Анализ этих методов посвящена специальная литература. Почти все методы основаны на вычислении частичной суммы рядов, которые сводятся к выполнению простых арифметических операций. В настоящее время, вообще-то, нет необходимости знать эти методы — алгоритмы вычисления простых элементарных функций заложены в трансляторах, достаточно в программе привести название функции. Расчет же более сложных функций всегда можно осуществить через элементарные функции.

Вопросы и упражнения для самоконтроля.

1. Какие функции относятся к элементарным?
2. Чем ограничивается схема Горнера?
3. Что понимают под остаточным членом?
4. Какую оценку точности называют априорной?
5. Какую оценку точности называют апостериорной? _
6. Вычислить значение полинома:
 - а) $P(x) = x^4 - 2x^3 + x^2 - 12x + 20$ при $x_0 = 7 \cdot N$;
 - б) $P(x) = x^4 - x^3 - 29x^2 - 71x - 140$ при $x_0 = 1 \cdot N$;
 - в) $P(x) = x^4 - 5x^3 + x^2 - 15x + 50$ при $x_0 = 0,5 \cdot N$;
 - г) $P(x) = x^4 + 10x^3 + 36x^2 + 70x + 75$ при $x_0 = 3 + 0,5 \cdot N$;
 где N — параметр, соответствующий списку группы.
7. Вычислить значение функции $f(x)$ с точностью 10^{-4} :
 - а) $f(x) = e^{-x}$ при $x = 2 \cdot N$;
 - б) $f(x) = \ln(1+x)$ при $x = N$;
 где N — параметр, соответствующий списку группы.

1.3 Приближенное решение уравнений

Далеко не каждое уравнение может быть решено точно. В первую очередь это замечание относится к трансцендентным уравнениям. Доказано также, что нельзя построить формулу расчета корней произвольного полинома степени выше четвертой. С другой стороны, не всегда точное решение является необходимым. Порой достаточно определить корни уравнений приближенно, с необходимой точностью.

Пусть дано уравнение

$$f(x)=0, \quad (1.7)$$

где функция $f(x)$ определена и непрерывна на отрезке $x \in [a, b]$.

Определение. Всякое значение x^* , обращающее значение функции $f(x)$ в нуль

$$f(x^*)=0,$$

называется корнем уравнения (1.7) или нулем функции $f(x)$.

При нахождении приближенных значений корней уравнения решают две задачи:

- 1) отделение корней, т.е. определение достаточно малых промежутков, в каждом из которых заключен только один корень (простой или кратный);
- 2) уточнение корней с наперед заданным числом верных знаков.

Корни можно отделить графически. При аналитическом отделении корней полезны следующие утверждения.

Утверждение 1. Если функция $f(x)$ непрерывна на отрезке $[a, b]$ и на его концах принимает разные знаки, т.е. $f(a) \cdot f(b) < 0$, то на данном отрезке, имеется, по крайней мере, один корень.

Утверждение 2. Если функция $f(x)$ на концах отрезка $[a, b]$ принимает разные знаки, то число корней на отрезке нечетно, если принимает одинаковые знаки, то число корней четно или корней нет вообще.

Утверждение 3. Если $f(a) \cdot f(b) < 0$ и производная $f'(x)$ сохраняет знак ($f'(x) > 0$ или $f'(x) < 0$), то на отрезке имеется единственный корень.

При отделении корней полинома полезна также теорема Безу, согласно которой остаток от деления полинома $P(x)$ на двучлен $(x-a)$ равен $P(a)$. Следствием этой теоремы является равенство $P(x) = Q(x)(x-x^*)$. Если x^* - корень $P(x)$, то остаток равен нулю, т.к. $P(x^*) = 0$. Тогда поиск остальных корней можно проводить по полиному $Q(x)$, степень которого на единицу ниже степени полинома $P(x)$. Следовательно, согласно основной теореме алгебры, и число корней на единицу меньше.

Всеобщих рекомендаций по отделению корней нет. Обычно строят примерный график функции и визуально отделяют корни.

Перейдем к рассмотрению проблемы уточнения корней, заключающейся в решении задачи

Задача. Определить значение x^* , удовлетворяющее уравнению (1.7) с точностью ε .

1.3.1 Метод дихотомии (деления отрезка пополам)

Пусть мы нашли такие точки a и b , что $f(a) \cdot f(b) < 0$ и на данном отрезке располагается только один корень.

Разделим отрезок пополам точкой $x = (a+b)/2$ и вычислим значение функции $f(x)$. Из двух половинок отрезка выберем ту, для которой выполняется неравенство $f(x_{\text{гран}}) \cdot f(x) < 0$. Здесь $x_{\text{гран}}$ - точка a или b (рисунок 1.3). Пусть, например, в качестве $x_{\text{гран}}$ выступает точка a . Тогда, если $f(a) \cdot f(x) < 0$, то смещается точка b ($b=x$). Если $f(a) \cdot f(x) > 0$, то смещается точка a ($a=x$).

Новый отрезок вновь разделим пополам и выберем ту половину, на концах которой функция имеет разные знаки, и т.д. продолжая процесс половинного деления, можно прийти до сколь угодно малого отрезка, содержащего корень уравнения.

Имеем итерационный процесс.

Алгоритм метода приведен на рисунке 1.4.

Критерием останова итерационного процесса служит неравенство

$$|f(x)| < \varepsilon \quad \text{или} \quad |b-a| < \varepsilon,$$

где ε - малая положительная величина (точность).

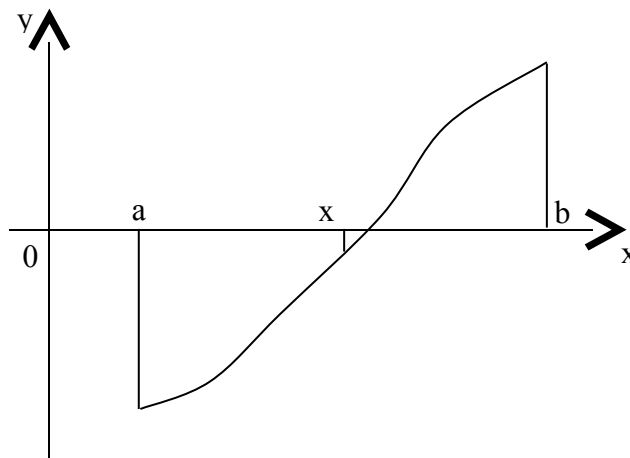


Рисунок 1.3 – К объяснению метода половинного деления

Достоинство метода. Метод абсолютно устойчив. При условии $f(a) \cdot f(b) < 0$ всегда сходится к решению.

Недостаток метода. Требуется много итераций.

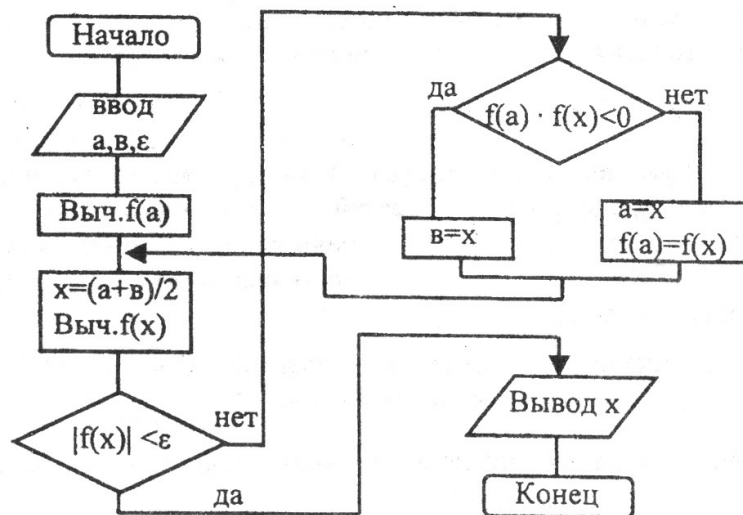


Рисунок 1.4 - Алгоритм метода половинного деления

Пример. Определить положительный корень уравнения

$$f(x) = x^2 + 5x - 6 = 0.$$

Нетрудно убедиться, что корнями уравнения являются $x_1^* = -6$ и $x_2^* = 1$.
Определим приближенное значение x_2^* методом дихотомии.

Пусть $x^* \in [0; 3]$. Так как $f(0) \cdot f(3) = (-6) \cdot 18 < 0$, то на данном отрезке располагается корень уравнения.

Шаг 1. $x_1 = (a+b)/2 = (0+3)/2 = 1,5$; $f(x_1) = 1,5^2 + 5 \cdot 1,5 - 6 = 3,75$.

Так как $f(a) \cdot f(x_1) = (-6) \cdot 3,75 < 0$, то выбираем отрезок $[0; 1,5]$.

Шаг 2. $x_2 = (0+1,5)/2 = 0,75$; $f(x_2) = 0,75^2 + 5 \cdot 0,75 - 6 = -1,69$.

Так как $f(a) \cdot f(x_2) = (-6) \cdot (-1,69) > 0$, то выбираем отрезок $[0,75; 1,5]$.

Шаг 3. $x_3 = (0,75+1,5)/2 = 1,125$ и т.д.

1.3.2 Метод хорд

Вычисление корня можно ускорить, если делить отрезок $[a, b]$ не пополам, а в соотношении $f(a) : f(b)$.

Координаты точки x :

$$x = a - \Delta,$$

где $\Delta = (b-a) f(a) / (f(b) - f(a))$.

Этот способ эквивалентен замене кривой $f(x)$ хордой, проходящей через точки А и В (рисунок 1.5).

Искомое уравнение выводится из уравнения прямой:

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

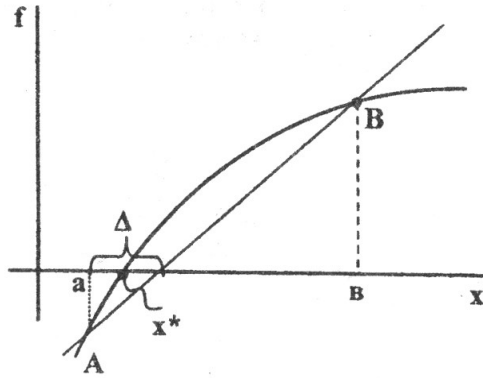


Рисунок 1.5 - Иллюстрация метода хорд

В предыдущем алгоритме (рисунок 1.4) изменится блок расчета x и в соответствующих местах необходимо рассчитать значение $f(x)$.

Пример. По методу хорд определить положительный корень уравнения

$$f(x) = x^2 + 5x - 6 = 0.$$

Примем $x^* \in [0; 3]$.

Шаг1. $f(0) = -6$; $f(3) = 18$.

Согласно методу исходный отрезок делится на две части точкой

$$x_1 = 0 - (3 - 0) \cdot (-6) / (18 - (-6)) = 0,75;$$

$$f(x_1) = 0,75^2 + 5 \cdot 0,75 - 6 = -1,69.$$

Так как $f(a) \cdot f(x_1) = (-6) \cdot (-1,69) > 0$, то выбираем отрезок $[0,75; 3]$.

Шаг2. $x_2 = 0,75 - (3 - 0,75)(-1,69) / (18 - (-1,69)) = 0,94$;

$$f(x_2) = 0,94^2 + 5 \cdot 0,94 - 6 = -0,42.$$

Так как $f(a) \cdot f(x_2) > 0$, то $x^* \in [0,94; 3]$ и т.д.

1.3.3 Метод Ньютона (метод касательных)

Рассмотрим понятие о *линеаризации функции*. Пусть функция $f(x)$ разложена в ряд Тейлора

$$f(x) = f(a) + \frac{f'(a)}{1!} (x - a) + \dots + \frac{f^{(k)}(a)}{k!} (x - a)^k + \dots$$

Выделим частичную сумму ряда до k -го члена включительно. При $k \approx \infty$ значение суммы ряда близко к значению функции $f(x)$; при $k \ll \infty$ приближение грубое. При $k=1$ получим прямую, касающуюся функции $f(x)$ в точке A - линеаризовали функцию (рисунок 1.6).

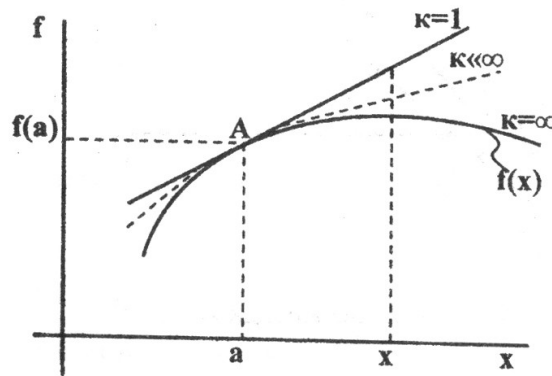


Рисунок 1.6 - Графическая интерпретация линеаризации функции

Итак, *линеаризация функции* - это разложение функции в ряд Тейлора до первой производной включительно.

Пусть найдено некоторое n -е приближение корня, т.е. x_n - точка, достаточно близкая к корню x^* .

Линеаризуя функцию в точке x_n , имеем

$$f(x) \approx f(x_n) + f'(x_n)(x - x_n).$$

Пусть точка x близка к корню x^* . Тогда $f(x) \approx 0$. Имеем

$$x - x_n = -f(x_n)/f'(x_n) \Rightarrow x = x_n - f(x_n)/f'(x_n)$$

Обозначим переменную x символом x_{n+1} . Тогда

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, \dots \quad (1.8)$$

Данный итерационный процесс называется *методом Ньютона* или *методом касательных*.

Рассмотрим геометрическую иллюстрацию метода (рисунок 1.7). Через точку M_n , проведем касательную к функции $f(x)$. Уравнение касательной

$$y - f(x_n) = f'(x_n)(x - x_n).$$

В качестве точки x возьмем точку пересечения касательной с осью Ox .

Тогда $y=0$. Зная значения $x_n, f(x_n), f'(x_n)$, по уравнению касательной можем найти координату точки пересечения с осью Ox :

$$x = x_n - \frac{f(x_n)}{f'(x_n)}$$

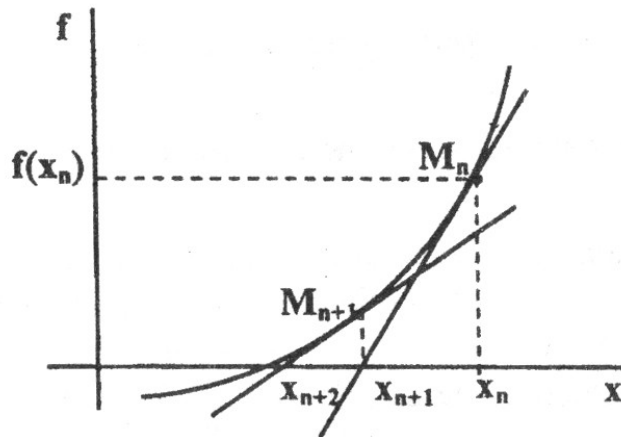


Рисунок 1.7- Иллюстрация метода Ньютона

Положив $x = x_{n+1}$, получим уравнение (1.8).

На следующем шаге метода через точку M_{n+1} проводим другую касательную и определяем координаты точки x_{n+2} и т.д. Алгоритм метода приведен на рисунке 1.8.

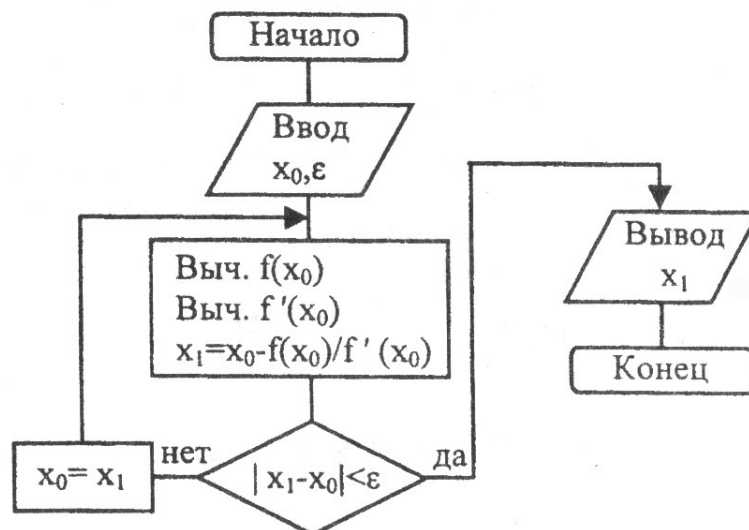


Рисунок 1.8 - Алгоритм метода Ньютона

В методе x_0 является начальным приближением, задаваемым самим вычислителем и берется тот конец отрезка $[a, b]$, на котором функция $f(x)$ и ее вторая производная имеют одинаковые знаки, т.е. выполняется неравенство $f(x_0) \cdot f''(x_0) > 0$.

При применении метода требуется выполнение условий:

- 1) все x_n должны принадлежать области определения функции $f(x)$;
- 2) $f'(x) \neq 0$, $n = 0, 1, 2, \dots$
- 3) знакопостоянство $f'(x_n) f''(x_n)$ для всех x_n принадлежащих области определения функции $f(x)$.

Достоинство метода. При отсутствии перегибов функции и выполнении условий 1), 2) и 3) метод сходится быстро (квадратично) и допускает различные модификации.

Недостаток метода. Для успешной сходимости метода требуется достаточно хорошее начальное приближение, в связи с чем необходимо знать поведение функции на отрезке $[a, b]$.

Пример. Методом Ньютона определить корень уравнения предыдущего примера $f(x)=x^2+5x-6=0$.

Производная функции имеет вид $f'(x)=2x+5$.

В качестве начального приближения возьмем точку $x_0=1,5$.

Согласно уравнению (1.8) на первой итерации имеем

$$x_1 = 1,5 - \frac{1,5 + 5 \cdot 1,5 - 6}{2 \cdot 1,5 + 5} = 1,03.$$

На второй итерации получаем

$$x_2 = 1,03 - \frac{f(1,03)}{f'(1,03)}.$$

1.3.4 Метод секущих

В методе Ньютона на каждом шаге вычисляются три величины x_n , $f(x_n)$ и $f'(x_n)$. Можно упростить вычисления, отказавшись от вычисления $f'(x_n)$.

Пусть

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}$$

Подставим в (1.8), имеем

$$x_{n+1} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}, \quad n = 0, 1, 2, \dots \quad (1.9a)$$

Можно также использовать рекуррентное уравнение

$$x_{n+1} = x_n - \frac{f(x_n)x_n - f(x_{n-1})x_{n-1}}{f(x_n) - f(x_{n-1})}, \quad n = 0, 1, 2, \dots \quad (1.9б)$$

Метод секущих является более экономичным по сравнению с методом Ньютона по количеству функций подлежащих расчету.

1.3.5 Метод простых итераций (последовательных приближений)

Исходную задачу $f(x)=0$ приведем к эквивалентному виду

$$f(x)=x-\varphi(x)=0 \quad \text{или} \quad x=\varphi(x).$$

Выберем некоторое приближение x_0 и вычислим дальнейшие приближения по формуле

$$x_{n+1} = \varphi(x_n), \quad n=0,1,2,\dots \quad (1.10)$$

Утверждение. Если итерационный процесс (1.10) сходится к некоторому пределу x^* , то этот предел есть корень уравнения. Предполагая, $\varphi(x)$ - непрерывная функция, получим

$$\bar{\delta}^* = \lim_{n \rightarrow \infty} x_{n+1} = \varphi \left(\lim_{n \rightarrow \infty} x_n \right) = \varphi(x^*),$$

т.е.

$$\bar{\delta}^* = \varphi(x^*).$$

Рассмотрим геометрическое представление утверждения.

При решении уравнения $x = \varphi(x)$ ищется абсцисса точки пересечения прямой $y=x$ и кривой $y = \varphi(x)$ (см. рисунок 1.9).

По принятому x_0 на кривой $\varphi(x)$ находится точка M_0 . Через нее проводится прямая, параллельная оси Ox , и находится точка пересечения ее с прямой $y=x$. Абсцисса этой точки принимается за следующее приближение x_1 . По ней на кривой $\varphi(x)$ находится точка M_1 . Далее от точки M_1 проводят такое же построение, как и от точки M_0 , и т.д.

Последовательность x_n , $n=0,1,2,\dots$ может быть построена, если все $x_n \in [a,b]$. Из уравнения (1.10) следует, что тогда все $\varphi(x_n) \in [c,d]$, где отрезок $[c,d]$ на оси Oy равен отрезку $[a,b]$ на оси Ox . Последовательность x_n , $n=0,1,2,\dots$ стремится в этом случае к пределу x^* (рисунок 1.8, а, б).

Возможны случаи, когда для некоторого $x_n \in [a,b]$ функция $\varphi(x_n) \notin [c,d]$ (рисунок 1.8, в, г). В этом случае для монотонной функции φ сходимость к корню x^* исключается.

Рассмотрим более подробно условия сходимости метода.

Необходимым условием сходимости является непрерывность функции $f(x)$. Определим *достаточное условие*.

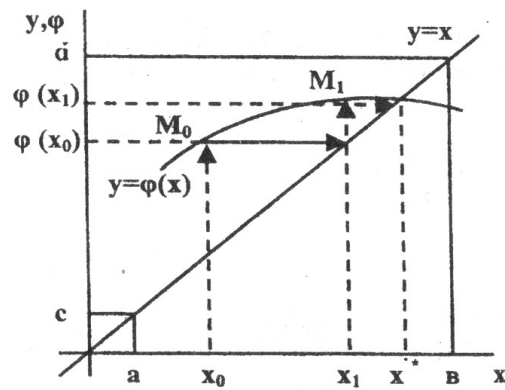
Теорема. Пусть $\varphi(x)$ определена и дифференцируема на отрезке $[a,b]$ и все ее значения $\varphi(x) \in [c,d]$. Тогда, если выполняется неравенство

$$|\varphi'(\bar{\delta})| < 1, \quad \forall \bar{\delta} \in [\hat{a}, \hat{a}]$$

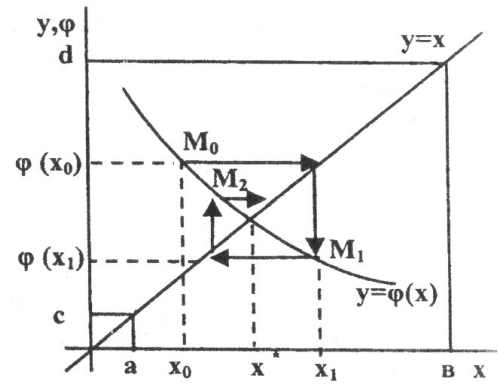
то процесс итерации $x_{n+1} = \varphi(x_n)$ сходится независимо от начального значения $x_n \in [a,b]$.

Доказательство. Из двух соседних приближений имеем

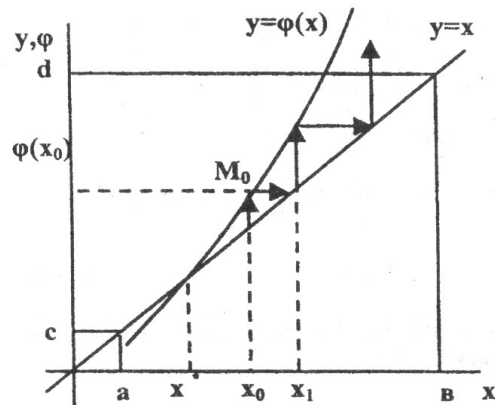
$$\bar{\delta}_{n+1} - x_n = \varphi(x_n) - \varphi(x_{n-1}).$$



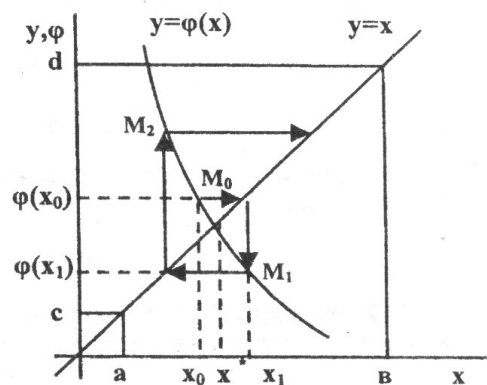
а)



б)



в)



г)

Рисунок 1.9 - Геометрическое представление метода простых итераций

Умножим правую часть на $(x_n - x_{n-1})/(x_n - x_{n-1})$ и, используя теорему Лагранжа о среднем, получаем

$$x_{n+1} - x_n = (x_n - x_{n-1}) \cdot \varphi'(\bar{x}), \quad \bar{x} \in [x_n, x_{n-1}].$$

Пусть q будет максимальным значением $|\varphi'(x)|$ на всем рассматриваемом отрезке $[a, b]$:

$$q = \max_{x \in [a, b]} |\varphi'(x)|.$$

Тогда, $|x_{n+1} - x_n| \leq q|x_n - x_{n-1}| \leq q^2|x_{n-1} - x_{n-2}|$ или

$$|x_{n+1} - x_n| \leq q^n|x_1 - x_0|, \quad n = 0, 1, 2, \dots$$

Ввиду того, что $q < 1$, отрезки $|x_n - x_{n-1}|$ убывают по геометрической прогрессии и последовательность x_n сходится при любом $x_0 \in [a, b]$. Очевидно, что предельное значение x_n есть корень x^* .

В связи с тем, что при выполнении условий теоремы метод сходится при любом $x_0 \in [a, b]$, то метод является самоисправляющимся. То есть, если при вычислении произошел сбой компьютера, получен неверный результат и этот результат не выходит за пределы $[a, b]$, то его можно принять за новое приближение x_0 и продолжить расчеты.

Алгоритм метода приведен на рисунке 1.10.

Пример. Определить положительный корень уравнения предыдущего примера

$$f(x) = x^2 + 5x - 6 = 0.$$

Исходное уравнение можно записать в виде

$$1) \quad x = \sqrt{6 - 5x} \quad \Rightarrow \quad \varphi'_1 = 5/(2\sqrt{6 - 5x});$$

$$2) \quad x = (6 - x^2)/5 \quad \Rightarrow \quad \varphi'_2 = -2x/5.$$

$$\text{При } x=1 \text{ имеем } \varphi'_1 = 2,5 \text{ и } \varphi'_2 = -0,4.$$

Ввиду того, что $|\varphi'_1| > 1$ и $|\varphi'_2| < 1$, то уравнение 1) не обеспечивает сходимость метода. Приемлемым является уравнение 2).

В качестве начального приближения возьмем точку $x_0 = 0$. Исходя из (1.10), следует:

Шаг 1. $x_1 = (6 - 0)/5 = 1,200$

Шаг 2. $x_2 = (6 - 1,2^2)/5 = 0,912$

Шаг 3. $x_3 = (6 - 0,912^2)/5 = 1,034$ и т.д.

Если бы выбрали уравнение 1), то имели:

Шаг 1. $x_1 = \sqrt{6 - 5 \cdot 0} = 2,45$

Шаг 2. $x_2 = \sqrt{6 - 5 \cdot 2,45} = \sqrt{-4,25}$ - нет сходимости

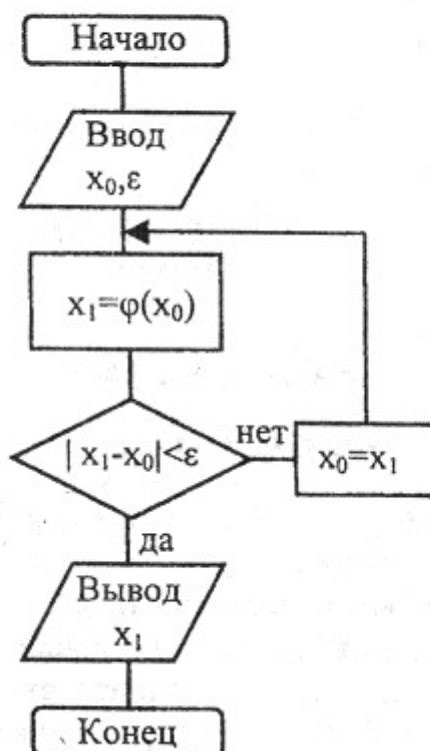


Рисунок 1.10 - Алгоритм метода простых итераций

Вопросы и упражнения для самоконтроля.

1. Что называют нулем функции $f(x)$?
2. Какие решаются задачи при нахождении приближенных значений корней уравнения?
3. Перечислите утверждения которые используются при аналитическом отделении корней.
4. В чем состоит основная идея метода дихотомии?
5. Какое неравенство служит критерием останова итерационного процесса метода дихотомии?
6. Назовите достоинства и недостатки метода дихотомии.
7. В чем заключается отличие метода хорд от метода дихотомии.
8. Что понимают под линеаризацией функции?
9. Какие условия необходимо выполнить для применения метода Ньютона?
10. Достоинства и недостатки метода Ньютона.
11. В чем состоит основная идея метода секущих?
12. Какова последовательность действий при решении уравнений методом простых итераций?
13. Какие условия гарантируют сходимость метода простых итераций к корню уравнения?
14. Методами дихотомии, хорд, Ньютона, секущих простых итераций найти корни уравнений ($\epsilon = 10^{-3}$):
 - а) $2x^3 - 3x^2 + 4x + 9 = 0$;
 - б) $\arctg x + \ln x = 0$;
 - в) $\delta^4 - 2\delta - 4 = 0$ (положительные корни);
 - г) $x - 1,6 \cos x = 0$;
 - д) $-3 + 2x^4 - x^2 + 1,2x = 0$ (отрицательные корни);
 - е) $2x^4 + 3x^3 + 8x^2 + 6x + 5 = 0$;
 - ж) $\lg(1,7x) - 2,8x = 0$;
 - з) $15x^4 - 4x^3 - 6x^2 - 4x - 1 = 0$ (отрицательные корни);
 - и) $4,3x - 5 \cos x = 0$;
 - к) $\ln(10x) = 7x - 1$;
 - л) $100x^3 + 45x^2 - 12x + 2,5 = 0$ (положительные корни);
 - м) $\delta^5 - \delta - 0,4 = 0$ (отрицательные корни);
 - н) $x^2 + 4 - 1/x = 0$;
 - о) $4x - 16 + e^x = 0$;
 - п) $x^3 + 4,05x^2 - 0,03x + 0,02 = 0$;
 - р) $2\delta^3 - 2\delta - 5 = 0$;
 - с) $3 - x - \lg x = 0$;
 - т) $\delta^3 + 2\delta - 9 = 0$.

1.4 Интерполирование и аппроксимация функций (приближение функций)

В инженерной практике довольно часто возникают задачи аналитического описания экспериментальных данных или замены громоздкой функции более простой. Здесь рассматриваются только частные, наиболее широко используемые способы приближения функций путем интерполирования и аппроксимации.

Рассмотрение поставленной задачи начнем с введения понятий сетки, конечных разностей и разностных отношений, используемых в дальнейшем.

1.4.1 Понятие сетки

Пусть на прямой x задан отрезок $[a, b]$. Если зададимся некоторой системой точек $x_i \in [a, b]$, то говорят, что на отрезке $[a, b]$ введена *одномерная сетка* Ω_h с узлами $x_i, i = \overline{0, n}$.

Данное определение сетки не предполагает какой-либо связи между нумерацией узлов x_i и их расположением в заданной области. Если такая связь имеется, то имеем *упорядоченную сетку*.

Пример упорядоченной сетки:

$$a \leq x_0 < x_1 < x_2 < \dots < x_n \leq b, \quad x_i = x_{i-1} + h_i.$$

Величина h_i называется *шагом сетки*:

$$h_i = x_i - x_{i-1}, \quad \text{где } i - \text{номер узла сетки.}$$

Равномерная сетка (рисунок 1.11):

$$x_i = a + i \cdot h, \quad i = \overline{0, n}, \quad h = \frac{b - a}{n}.$$

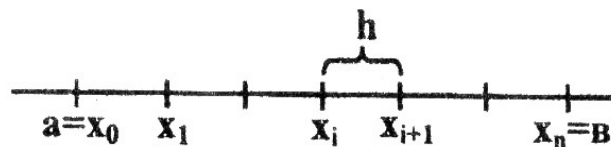


Рисунок 1.11 - Равномерная одномерная сетка

Число узлов: $((b-a)/h+1)=n+1$.

Мы рассмотрели одномерную сетку. Аналогично вводятся *двумерная* и *многомерные* сетки.

Введем сетку $\Omega = \Omega_x \times \Omega_y$:

$$a \leq x_0 < \dots < x_n \leq b; \quad c \leq y_0 < \dots < y_m \leq d,$$

$$h_{x_i} = x_i - x_{i-1}, \quad i = \overline{1, n}; \quad h_{y_j} = y_j - y_{j-1}, \quad j = \overline{1, m}.$$

Эта сетка называется *прямоугольной* (рисунок 1.12), точки (x_i, y_j) называются *узлами сетки*.

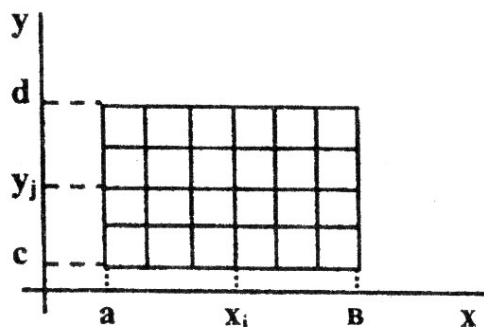


Рисунок 1.12 - Двумерная прямоугольная сетка

Если $x_i = a + i \cdot h_x$, $i = \overline{0, n}$; $y_j = c + j \cdot h_y$, $j = \overline{0, m}$; то имеем *равномерную прямоугольную сетку*.

Число узлов равно $((b-a)/h_x+1)((d-c)/h_y+1)=(n+1)(m+1)$.

Если нет связи между нумерацией узлов и их расположением в области Ω , то имеем *хаотичную сетку* (рисунок 1.13).

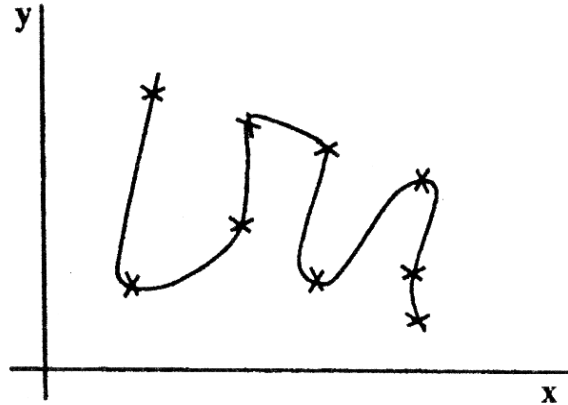


Рисунок 1.13 - Пример хаотичной сетки

Если на $[a, b]$ определена функция $f(x)$ и $f_i = f(x_i)$, $i = \overline{0, n}$, она называется *табличной (сеточной) функцией*.

Пример. Изменение давления в системе в зависимости от температуры представлено таблицей

Давление, Мпа	0,18	0,19	0,22	0,27	0,33	0,44
Температура, °C	150	175	200	225	250	275

Если здесь в качестве функции представлено давление, в качестве аргумента - температура, то можно говорить, что на отрезке $[150; 275]$ задана равномерная сетка с шагом $h=25$. При этом значениями сеточной (табличной) функции являются значения давления.

Если в качестве функции взять температуру и в качестве аргумента давление, то на отрезке $[0,18; 0,44]$ задана упорядоченная сетка с различными шагами.

1.4.2 Конечные разности и разностные отношения

Величины $\Delta f_0 = f_1 - f_0$, $\Delta f_1 = f_2 - f_1, \dots$, $\Delta f_i = f_{i+1} - f_i$, $i = \overline{0, n-1}$ называются *конечными разностями первого порядка*. Здесь $f_i = f(x_i)$ - сеточная функция.

Величины $\Delta^2 f_i = \Delta f_{i+1} - \Delta f_i$ - *конечные разности второго порядка*.

Величины $\Delta^k f_i = \Delta^{k-1} f_{i+1} - \Delta^{k-1} f_i$ - *конечные разности k-го порядка*.

Следует отметить следующую зависимость: $\Delta^k f_i = \Delta (\Delta^{k-1} f_i)$.

Конечные разности можно вычислить непосредственно, используя значения функции, например:

$$\Delta^2 f_0 = \Delta f_1 - \Delta f_0 = (f_2 - f_1) - (f_1 - f_0) = f_2 - 2f_1 + f_0$$

и т.д.

Продолжая вычисления, по индукции можно убедиться, что

$$\Delta^k f_0 = f_k - \frac{k}{1!} f_{k-1} + \frac{k(k-1)}{2!} f_{k-2} \dots + (-1)^k f_0.$$

Мы рассмотрели *правые разности*. Используются также *левые разности*:

$$\nabla f_k = f_k - f_{k-1}$$

или *центральные разности*

$$\delta f_k = f_{k+1/2} - f_{k-1/2}$$

Перейдем к рассмотрению *разностных отношений*.

Пусть на отрезке $[a, b]$ задана сетка Ω_n (не обязательно упорядоченная).

Разностными отношениями 1-го порядка называются величины

$$f(x_0, x_1) = \frac{\Delta f_0}{x_1 - x_0} = \frac{f_1 - f_0}{x_1 - x_0}; \dots f(x_{n-1}, x_n) = \frac{\Delta f_{n-1}}{x_n - x_{n-1}} = \frac{f_n - f_{n-1}}{x_n - x_{n-1}}.$$

Разностные отношения второго порядка

$$f(x_0, x_1, x_2) = \frac{\Delta^2 f_0}{x_2 - x_0} = \frac{f(x_1, x_2) - f(x_0, x_1)}{x_2 - x_1},$$

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

По индукции *разностные отношения k-го порядка*

$$f(x_0, x_1, \dots, x_k) = \frac{f(x_1, \dots, x_k) - f(x_0, \dots, x_{k-1})}{x_k - x_0}.$$

Разностные отношения имеют смысл и размерность производных соответствующих порядков.

Разностные отношения могут быть получены через значения функции:

$$f(x_0, x_1) = f_0 / (x_0 - x_1) + f_1 / (x_1 - x_0)$$

$$f(x_0, x_1, x_2) = \frac{1}{x_2 - x_0} (f(x_1, x_2) - f(x_0, x_1)) = \frac{1}{x_2 - x_0} \left(\frac{f_1}{x_1 - x_2} + \frac{f_2}{x_2 - x_1} - \frac{f_0}{x_0 - x_1} - \frac{f_1}{x_1 - x_0} \right) = \frac{f_0}{(x_0 - x_1)(x_0 - x_2)} + \frac{f_1}{(x_1 - x_0)(x_1 - x_2)} + \frac{f_2}{(x_2 - x_0)(x_2 - x_1)}.$$

По индукции

$$f(x_0, \dots, x_k) = \sum_{i=0}^k \frac{f_i}{(x_i - x_0) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_k)}.$$

Можно показать, что при всяком n имеет место равенство

$$\begin{aligned} f(x_n) = & f(x_0) + (x_n - x_0) f(x_0, x_1) + (x_n - x_0)(x_n - x_1) f(x_0, x_1, x_2) + \dots \\ & + (x_n - x_0) \dots (x_n - x_{n-1}) f(x_0, x_1, \dots, x_n). \end{aligned} \quad (1.11)$$

Доказательство может быть получено по индукции. Ввиду простоты ограничимся лишь проверкой (1.11) при $n=1$ и $n=2$.

$$n=1: \quad f(x_0, x_1) = \frac{f_1 - f_0}{x_1 - x_0} \quad \Rightarrow \quad f_1 = f_0 + (x_1 - x_0) f(x_0, x_1).$$

$$n=2: \quad f(x_0, x_1, x_2) = \frac{f(x_1, x_2) - f(x_0, x_1)}{x_2 - x_0} \Rightarrow$$

$$f_2 = f_0 + (x_2 - x_0) f(x_0, x_1) + (x_2 - x_0)(x_2 - x_1) f(x_0, x_1, x_2).$$

Если имеем равномерную сетку, то $x_k = x_0 + kh$, $k = \overline{0, n}$. В этом случае

$$f(x_0, x_1) = f(x_0, x_0 + h) = \frac{f(x_0 + h) - f(x_0)}{x_0 + h - x_0} = \frac{\Delta f_0}{1!h}.$$

$$f(x_0, x_1, x_2) = f(x_0, x_0 + h, x_0 + 2h) = \frac{1}{2h} \left(\frac{\Delta f_1}{h} - \frac{\Delta f_0}{h} \right) = \frac{\Delta^2 f_0}{2!h^2}$$

и так далее.

Связь между конечными разностями и разностными отношениями

$$f(x_0, \dots, x_k) = \frac{\Delta^k f_0}{k! h^k}.$$

1.4.3 Понятие интерполирования функций

Пусть на отрезке $[a, b]$ задана табличная функция $f(x_i)$, определенная в узлах $x_0, \dots, x_n$ сетки Ω_h (рисунок 1.14). Здесь табличная функция $f(x_i)$ является следом некоторой функции $f(x)$, аналитический вид которой нам не известен или функция очень громоздка, ее анализ затруднителен. Следом функции являются значения некоторой функции $f(x)$ в узлах сетки Ω_h .

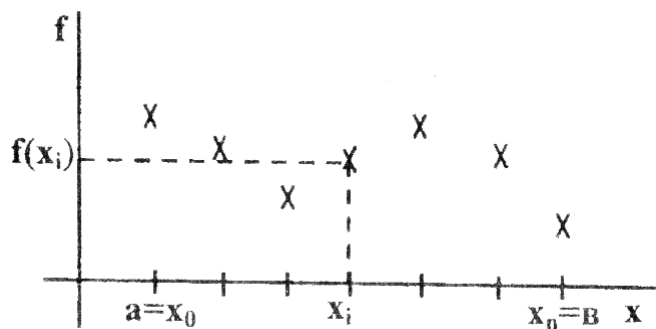


Рисунок 1.14 - График табличной (сеточной) функции

Под *интерполированием* понимают приближение табличной функции $f(x_i)$ некоторой непрерывной функцией $P(x)$, аналитический вид которой строится на основе табличных данных (рисунок 1.15).

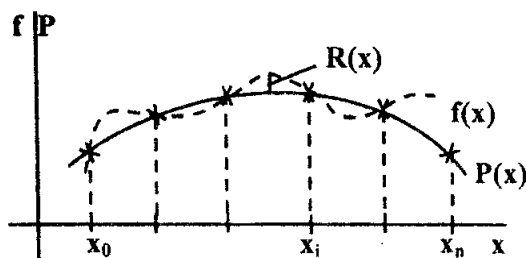


Рисунок 1.15 - Интерполирование табличной функции $f(x_i)$ функцией $P(x)$

При построении интерполирующей функции исходят из того, что в узлах сетки выполняются равенства $P(x_i)=f(x_i)$, $i=\overline{0, n}$.

Известно несколько способов приближения табличной функции: с помощью полиномов, тригонометрических, логарифмических, экспоненциальных функций и др.

Используя интерполирующую функцию $P(x)$, можно решать ряд задач математического анализа: нахождение значения функции в точках, отличных от узлов сетки ($x \neq x_i$), определения корней функции, точек экстремумов и перегибов, производных и других задач. Нередко анализ громоздкой функции проводят с помощью приближенной интерполирующей функции.

Термин *интерполяция* используется при исследовании функции внутри отрезка $[a, b]$. Если значение x выходит за пределы $[a, b]$ ($x \notin [a, b]$), то употребляют термин *экстраполяция*.

1.4.3.1 Полиномиальная интерполяция

Возьмем полином

$$P_n(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n \quad (1.12)$$

и его коэффициенты $a_k, k = \overline{0, n}$ выберем так, чтобы выполнялись условия

$$P_n(x_i) = f(x_i), \quad i = \overline{0, n}, \quad (1.13a)$$

т.е. удовлетворяющие системе уравнений

ряющие системе урав-

$$\begin{cases} P_n(x_0) = f(x_0) \\ \vdots \\ P_n(x_n) = f(x_n) \end{cases} \quad (1.13b)$$

Эти уравнения дают $n+1$ уравнений для нахождения коэффициентов a_k . Можно показать, что если узлы не кратные, то решение системы (1.13) существует, и оно единственно. Но задача решения системы сама по себе является трудоемкой. Обычно её не решают, а строят интерполяционный полином, используя так называемые фундаментальные полиномы.

Один такой фундаментальный полином имеет вид

$$\omega_k(x) = \frac{(x - x_0) \dots (x - x_{k-1})(x - x_{k+1}) \dots (x - x_n)}{(x_k - x_0) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_n)} = \frac{\omega(x)}{(x - x_k) \cdot \omega'(x_k)}.$$

Этот полином - полином степени n . Здесь $\omega(x) = (x - x_0) \dots (x - x_n)$.

Основным свойством $\omega_k(x)$ является (рисунок 1.16):

$$\omega_k(x) = \begin{cases} 1, & x = x_k \\ 0, & x = x_i, \quad i \neq k \end{cases} \quad (1.14)$$

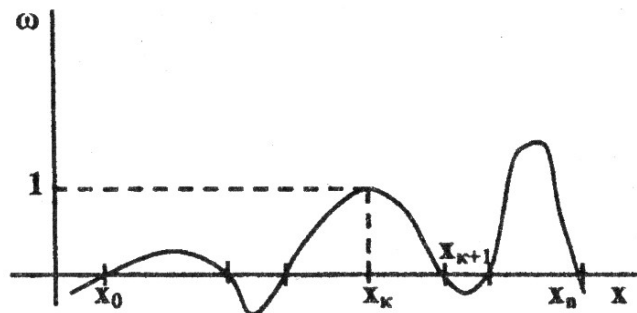


Рисунок 1.16 - График функции $\omega_k(x)$

1.4.3.2 Интерполяционный полином Лагранжа

Построим интерполяционный полином в виде

$$P_n(x) = \sum_{k=0}^n \omega_k(x) f(x_k), \quad (1.15a)$$

$$P_n(x) = \sum_{k=0}^n f(x_k) \frac{(x - x_0) \dots (x - x_{k-1})(x - x_{k+1}) \dots (x - x_n)}{(x_k - x_0) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_n)}. \quad (1.15b)$$

Полином (1.15) называют *интерполяционным полиномом Лагранжа*. Покажем, что этот полином удовлетворяет условию (1.13).

Действительно,

$$P_n(x_i) = \sum_{k=0}^n \omega_k(x_i) f(x_k) \stackrel{(1.14)}{=} \sum_{k=0}^n \delta_{ki} f(x_k) = f(x_i),$$

так как $\delta_{ki} = \begin{cases} 1, & k = i \\ 0, & k \neq i \end{cases}$ - символ Кронекера.

Погрешность интерполяционного полинома Лагранжа

$$R_n(x) = f(x) - P_n(x) = \frac{f^{(n+1)}(\zeta)}{(n+1)!} \omega(x), \quad \zeta \in [a, b].$$

Частные случаи:

а) *линейная интерполяция* ($n=1$) (рисунок 1.17). В качестве начальной точки возьмем $i-1$ -й узел, в качестве конечной - i -й узел:

$$P_1(x) = f(x_{i-1}) \frac{x - x_i}{x_{i-1} - x_i} + f(x_i) \frac{x - x_{i-1}}{x_i - x_{i-1}};$$



Рисунок 1.17 - Линейная интерполяция

б) квадратичная (параболическая) интерполяция ($n=2$) (рисунок 1.18).
Начальная точка - $i-1$ - й узел, конечная - $i+1$ - й узел:

$$P_2(x) = f(x_{i-1}) \frac{(x - x_i)(x - x_{i+1})}{(x_{i-1} - x_i)(x_{i-1} - x_{i+1})} + f(x_i) \frac{(x - x_{i-1})(x - x_{i+1})}{(x_i - x_{i-1})(x_i - x_{i+1})} + f(x_{i+1}) \frac{(x - x_{i-1})(x - x_i)}{(x_{i+1} - x_{i-1})(x_{i+1} - x_i)}.$$



Рисунок 1.18 - Квадратичная интерполяция

Приведенные формулы можно использовать как при интерполяции, так и при экстраполяции - достаточно подставить соответствующие значения x (рисунок 17, 18). Алгоритм метода приведен на рисунке 1.19.

Замечания

1. Характер поведения большинства интерполирующих функций таков, что значения старших производных с ростом n резко возрастают. В связи с этим при больших n погрешность интерполяции $R_n(x)$ становится большой. Поэтому на практике используют квадратичные или кубические полиномы ($n=2,3$).

2. Интерполяционный полином Лагранжа удобно использовать при приближении нескольких табличных функций, заданных в одних и тех же узлах сетки. Тогда достаточно один раз рассчитать $\omega(x)$ и $\omega'(x)$ и подставить соответствующие значения табличных функций.

Пример. Изменение температуры в аппарате в зависимости от расхода потока представлено таблицей.

Расход, т/ч	110	120	130	140	150
Температура, °C	147	131	121	115	112

Определить температуру в аппарате при расходе 125 т/ч.

Используем квадратичную интерполяцию. В качестве начальной точки возьмем расход 120 т/ч, конечной - 140 т/ч.

Имеем:

$$P_2(125) = 131 \frac{(125-130)(125-140)}{(120-130)(120-140)} + 121 \frac{(125-120)(125-140)}{(130-120)(130-140)} =$$

$$= 115 \frac{(125-120)(125-130)}{(140-120)(140-130)} = \frac{131 \cdot 3 + 121 \cdot 6 - 115}{8} = 125,5.$$

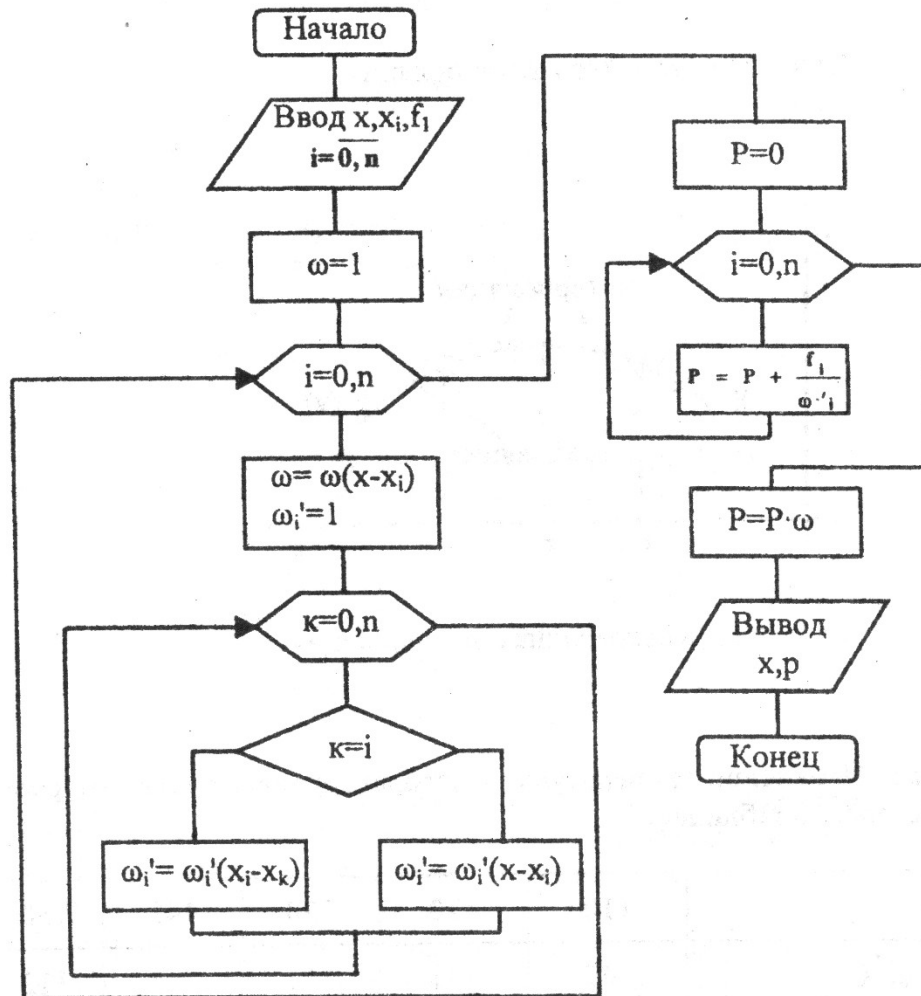


Рисунок 1.19 - Алгоритм интерполяционного полинома Лагранжа

1.4.3.3 Интерполяционная формула Эйткина

Для практического нахождения значений функции интерполяционный полином Лагранжа плохо приспособлен. Он требует не слишком большого числа действий, но мало удобен для записи вычислений. Более удобным является проведение вычислений по схеме Эйткина.

Начнем с рассмотрения линейной интерполяции. Пусть даны два значения аргумента x_0 и x_1 , в которых функция принимает соответственно значе-

ния y_0 и y_1 . Для вычисления значения функции в некоторой точке x с помощью линейной интерполяции необходимо построить прямую, проходящую через эти точки и вычислить величину ординаты этой прямой в требуемой точке.

Это значение может быть вычислено по следующей формуле

$$\mathbb{D}_{0,1}(\delta) = \frac{1}{\delta_1 - \delta_0} \begin{vmatrix} \acute{o}_0 & \delta_0 - \delta \\ \acute{o}_1 & \delta_1 - \delta \end{vmatrix}. \quad (1.16)$$

Действительно, выражение для $P_{0,1}(x)$ представляет собой многочлен первой степени. Легко убедиться, что $P_{0,1}(x)$ принимает в заданных точках нужные значения. Чтобы это сделать, вычислим значения в точках x_0 и x_1 :

$$\mathbb{D}_{0,1}(\delta_0) = \frac{1}{\delta_1 - \delta_0} \begin{vmatrix} \acute{o}_0 & 0 \\ \acute{o}_1 & \delta_1 - \delta_0 \end{vmatrix} = \acute{o}_0,$$

$$\mathbb{D}_{0,1}(\delta_1) = \frac{1}{\delta_1 - \delta_0} \begin{vmatrix} \acute{o}_0 & \delta_0 - \delta_1 \\ \acute{o}_1 & 0 \end{vmatrix} = \acute{o}_1.$$

Пример. По значениям функции $y=e^x$ в двух точках

$$x_0 = 0,40, \quad y_0 = 1,4918,$$

$$x_1 = 0,42, \quad y_1 = 1,5220,$$

найдем с помощью схемы Эйткина значение e^x для $x = 0,411$.

Используя формулу (1.6), находим

$$\mathbb{D}_{0,1} = \frac{1}{0,42 - 0,40} \begin{vmatrix} 1,4918 & -0,011 \\ 1,5220 & 0,009 \end{vmatrix} = \frac{1,4918 \cdot 0,009 + 1,5220 \cdot 0,011}{0,02} = 1,5084$$

Интерполяция более высокой степени производится аналогично, причем для построения каждого многочлена следующей степени используются уже построенные предыдущие.

Если даны три значения аргумента x_0, x_1, x_2 , в которых функция принимает соответственно значения y_0, y_1, y_2 , то предварительно строятся два линейных многочлена

$$\mathbb{D}_{0,1}(\delta) = \frac{1}{\delta_1 - \delta_0} \begin{vmatrix} \acute{o}_0 & \delta_0 - \delta \\ \acute{o}_1 & \delta_1 - \delta \end{vmatrix} \quad (1.17)$$

и

$$\mathbb{D}_{1,2}(\delta) = \frac{1}{\delta_2 - \delta_1} \begin{vmatrix} \acute{o}_1 & \delta_1 - \delta \\ \acute{o}_2 & \delta_2 - \delta \end{vmatrix}. \quad (1.18)$$

Значение функции в точке x , соответствующее интерполяции второй степени по указанным трем точкам (такую интерполяцию называют квадратичной),

получается путем линейной интерполяции между значениями этих линейных выражений, т.е. по формуле

$$\mathfrak{D}_{0,1,2}(\delta) = \frac{1}{\delta_2 - \delta_0} \begin{vmatrix} \mathfrak{D}_{0,1}(\delta) & \delta_0 - \delta \\ \mathfrak{D}_{1,2}(\delta) & \delta_2 - \delta \end{vmatrix}. \quad (1.19)$$

Действительно, так как $P_{0,1}$ и $P_{1,2}$ – многочлены первой степени, то $P_{0,1,2}$ – многочлен второй степени относительно x . Далее, по построению $P_{0,1}(x)$ и $P_{1,2}(x)$ имеем

$$P_{0,1}(x_0) = y_0; \quad P_{0,1}(x_1) = y_1; \quad P_{1,2}(x_1) = y_1; \quad P_{1,2}(x_2) = y_2$$

Поэтому

$$\mathfrak{D}_{0,1,2}(\delta_0) = \frac{1}{\delta_2 - \delta_0} \begin{vmatrix} \mathfrak{D}_{0,1}(\delta_0) & \delta_0 - \delta_0 \\ \mathfrak{D}_{1,2}(\delta_0) & \delta_2 - \delta_0 \end{vmatrix} = \mathfrak{D}_{0,1}(\delta_0) = y_0,$$

$$\mathfrak{D}_{0,1,2}(\delta_2) = \frac{1}{\delta_2 - \delta_0} \begin{vmatrix} \mathfrak{D}_{0,1}(\delta_2) & \delta_0 - \delta_2 \\ \mathfrak{D}_{1,2}(\delta_2) & 0 \end{vmatrix} = \mathfrak{D}_{1,2}(\delta_2) = y_2.$$

Кроме того,

$$\mathfrak{D}_{0,1,2}(\delta_1) = \frac{1}{\delta_2 - \delta_0} \begin{vmatrix} \mathfrak{D}_{0,1}(\delta_1) & \delta_0 - \delta_1 \\ \mathfrak{D}_{1,2}(\delta_1) & \delta_2 - \delta_1 \end{vmatrix} = \frac{1}{\delta_2 - \delta_0} \begin{vmatrix} y_1 & \delta_0 - \delta_1 \\ y_1 & \delta_2 - \delta_1 \end{vmatrix} = y_1.$$

Таким образом, $\mathfrak{D}_{0,1,2}(\delta)$ действительно является многочленом второй степени, принимающим в заданных точках заданные значения, а такой многочлен является единственным.

Пример. Функция $y = f(x)$ задана в таблице. С помощью квадратичной интерполяции по схеме Эйткина найти значение $f(13,13)$.

(1)	(2)	(1)	(2)
x	f(x)	x	f(x)
13,00	1,499 362	13,15	1,504 942
13,05	1,501 061	13,20	1,507 111
13,10	1,502 923	13,25	1,509 425

Применим $x_0 = 13,10$. Тогда последовательно находим

$$\mathfrak{D}_{0,1} = \frac{1}{0,05} \begin{vmatrix} 1,502923 & -0,03 \\ 1,504942 & 0,02 \end{vmatrix} = 1,504134,$$

$$\mathfrak{D}_{1,2} = \frac{1}{0,05} \begin{vmatrix} 1,504942 & 0,02 \\ 1,507111 & 0,07 \end{vmatrix} = 1,504074,$$

$$\mathfrak{D}_{0,1,2} = \frac{1}{10} \begin{vmatrix} 1,504134 & -0,03 \\ 1,504074 & 0,07 \end{vmatrix} = 1,504116.$$

Итак, искомое значение функции, полученное квадратичной интерполяцией, есть 1,504116, тогда как линейная интерполяция на участке (x_0, x_1) дала 1,504134. Точное значение функции, взятое из более подробной таблицы, равно $f(13,13)=1,504116$, так что квадратичная интерполяция сохраняет все знаки точными, в то время как линейная дает погрешность порядка 20 единиц последнего знака.

Рассмотренная схема легко обобщается на более высокие степени. Например, при интерполяции третьей степени по четырём точкам (кубическая интерполяция) имеет место формула

$$\mathbb{D}_{0,1,2,3}(\delta) = \frac{1}{\delta_3 - \delta_0} \begin{vmatrix} \mathbb{D}_{0,1,2}(\delta) & \delta_0 - \delta \\ \mathbb{D}_{1,2,3}(\delta) & \delta_3 - \delta \end{vmatrix}, \quad (1.20)$$

где многочлен $\mathbb{D}_{0,1,2}$ определен формулой (1.19), а $\mathbb{D}_{1,2,3}$ построен аналогично.

Вообще, если даны $n+1$ значений аргумента $x_0, x_1, \dots, x_n$ и функция в них принимает соответственно значения $y_0, y_1, \dots, y_n$, то интерполяционный многочлен n -й степени можно получить по формуле

$$\mathbb{D}_{0,1,\dots,n} = \frac{1}{\delta_n - \delta_0} \begin{vmatrix} \mathbb{D}_{0,1,\dots,n-1}(\delta) & \delta_0 - \delta \\ \mathbb{D}_{1,2,\dots,n}(\delta) & \delta_n - \delta \end{vmatrix}, \quad (1.21)$$

в которую входят предполагаемые построенными многочлены предыдущей степени.

Нетрудно убедиться в том, что формула (1.21), если её выразить последовательно через значения функции $y_0, y_1, \dots, y_n$, совпадает с формулой Лагранжа (1.15).

В самом деле линейный многочлен (1.6) имеет вид

$$\mathbb{D}_{0,1,2}(\delta) = \frac{1}{\delta_1 - \delta_0} [(x_1 - x)y_0 - (x_0 - x)y_1] = \frac{x - x_1}{x_0 - x_1} y_0 + \frac{x - x_0}{x_1 - x_0} y_1,$$

что совпадает с формулой Лагранжа.

Далее, для $\mathbb{D}_{1,2}(\delta)$ аналогично можно написать

$$\mathbb{D}_{1,2}(\delta) = \frac{x - x_2}{x_1 - x_2} y_1 + \frac{x - x_1}{x_2 - x_1} y_2.$$

Поэтому для многочлена второй степени (1.19) получаем

$$\begin{aligned} \mathbb{D}_{0,1,2}(\delta) &= \frac{1}{\delta_2 - \delta_0} \left[\left(\frac{x - x_1}{x_0 - x_1} y_0 + \frac{x - x_0}{x_1 - x_0} y_1 \right) (x_2 - x) - \left(\frac{x - x_2}{x_1 - x_2} y_1 + \frac{x - x_1}{x_2 - x_1} y_2 \right) (x_0 - x) \right] = \\ &= \frac{(x - x_1)(x - x_1)}{(x_0 - x_1)(x_2 - x_1)} y_0 + \frac{(x - x_0)(x - x_2)}{x_2 - x_0} y_1 \left(\frac{1}{x_0 - x_1} + \frac{1}{x_1 - x_2} \right) + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} y_2 = \\ &= \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} y_0 + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} y_1 + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} y_2. \end{aligned}$$

что снова совпадает с формулой Лагранжа для трех узлов интерполяции. Такое совпадение многочлена, полученного по схеме Эйткина с интерполяционной формулой Лагранжа, имеет место для любой степени n .

1.4.3.4 Интерполяционная формула Ньютона

Ранее мы показали, что при любом n функцию $f(x_n)$ можно представить через разностные отношения (1.11):

$$f(x_n) = f(x_0) + (x_n - x_0)f(x_0, x_1) + (x_n - x_0)(x_n - x_1)f(x_0, x_1, x_2) + \dots + (x_n - x_0) \dots (x_n - x_{n-1})f(x_0, \dots, x_n).$$

Если вместо x_n примем x , получим *первую интерполяционную формулу Ньютона*:

$$P_n(x) = f(x_0) + (x - x_0)f(x_0, x_1) + (x - x_0)(x - x_1)f(x_0, x_1, x_2) + \dots + (x - x_0) \dots (x - x_{n-1})f(x_0, \dots, x_n), \quad (1.22a)$$

или

$$P_n(x) = f(x_0) + \sum_{k=1}^n (x - x_0) \dots (x - x_{k-1})f(x_0, \dots, x_k). \quad (1.22б)$$

Здесь $P_n(x)$ - полином степени n .

Покажем, что полином (1.22) удовлетворяет условию (1.13). Действительно, подставив в (1.22) вместо x соответственно $x_0, x_1, \dots, x_n$, убеждаемся, что для каждого x_i $P_n(x_i) = f(x_i)$.

Пусть, например, $x = x_1$:

$$P_n(x_1) = f(x_0) + (x_1 - x_0)f(x_0, x_1) + (x_1 - x_0)(x_1 - x_1)f(x_0, x_1, x_2) + 0 = f(x_0) + (x_1 - x_0)(f(x_1) - f(x_0))/(x_1 - x_0) + 0 = f(x_1).$$

Погрешность интерполяционной формулы Ньютона

$$R_n(x) = f(x) - P_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega(x), \quad \xi \in [a, b]. \quad (1.23)$$

Эта оценка - априорная оценка. Её используют в теоретических исследованиях, до проведения вычислений. На практике обычно используют апостериорную оценку. Для интерполяционной формулы Ньютона этой оценкой является первый отброшенный член полинома:

$$\check{R}_n(x) = |(x - x_0) \dots (x - x_n)f(x_0, \dots, x_{n+1})|. \quad (1.24)$$

Частные случаи интерполяционной формулы Ньютона:

а) линейная интерполяция ($n=1$).

$$P_1(x) = f(x_0) + (x - x_0)f(x_0, x_1); \quad \check{R}_1(x) = |(x - x_0)(x - x_1)f(x_0, x_1, x_2)|;$$

б) квадратичная (параболическая) интерполяция ($n=2$).

$$P_2(x) = f(x_0) + (x-x_0)f(x_0, x_1) + (x-x_0)(x-x_1)f(x_0, x_1, x_2).$$

Пример. Задачу предыдущего примера решить с применением интерполяционной формулы Ньютона.

Имеем:

$$P_2(125) = 131 + (125 - 120) \frac{121 - 131}{130 - 120} + (125 - 120)(125 - 130)$$

$$\left(\frac{115 - 121}{140 - 130} - \frac{121 - 131}{130 - 120} \right) / (140 - 120) = 131 - 5 - 0,5 = 125,5.$$

Первая формула Ньютона используется при интерполировании в начале отрезка $[a, b]$ и при экстраполировании назад ($x < a$). При интерполировании в конце отрезка и при экстраполировании вперед ($b < x$), предпочтительно использовать *вторую интерполяционную формулу Ньютона*:

$$P_n(x) = f(x_n) + (x-x_n)f(x_n, x_{n-1}) + (x-x_n)(x-x_{n-1})f(x_n, x_{n-1}, x_{n-2}) + \dots + (x-x_n) \dots (x-x_1)f(x_n, \dots, x_0). \quad (1.25)$$

Интерполяционную формулу Ньютона-Стирлинга можно использовать при интерполировании в середине отрезка:

$$P_n(x) = f(x_k) + (x-x_k)f(x_k, x_{k+1}) + (x-x_k)(x-x_{k+1})f(x_k, x_{k+1}, x_{k-1}) + \dots + (x-x_k)(x-x_{k+1})(x-x_{k-1})f(x_k, x_{k+1}, x_{k-1}, x_{k+2}) + \dots \quad (1.26)$$

При расчете вручную удобно пользоваться таблицей:

x_0	$f(x_0)$		первая формула Ньютона	
x_1	$f(x_1)$	$f(x_0, x_1)$	$f(x_0, x_1, x_2)$	$f(x_0, x_1, x_2, x_3)$
x_2	$f(x_2)$	$f(x_1, x_2)$	$f(x_1, x_2, x_3)$	
x_3	$f(x_3)$	$f(x_2, x_3)$	вторая формула Ньютона	

Алгоритм первой интерполяционной формулы Ньютона приведен на рисунке 1.19.

Замечания

1. Как интерполяционный полином Лагранжа, так и интерполяционная формула Ньютона имеет одинаковую априорную оценку точности (1.23). Однако на практике более удобно использовать интерполяционную формулу Ньютона. Задавшись некоторым $k < n$, рассчитывают по (1.22) $P_k(x)$. Если значение $\sum_R k(x)$ больше заданной точности, то добавляют к взятым узлам один или несколько узлов. При этом ранее вычисленные слагаемые (1.22) сохраня-

ются и добавляются лишь новые члены $(x-x_0)\dots(x-x_k)f(x_0,\dots,x_{k+1})$, имеющие смысл поправки к уже вычисленному значению $P_k(x)$. В формуле Лагранжа добавление новых узлов требует исправления всех ранее рассчитанных членов.

2. Для расчетов по формуле Ньютона безразличен порядок, в котором пронумерованы узлы интерполяции. Следовательно, при добавлении нового $k+1$ - узла его можно брать и внутри сетки, не расширяя отрезок $[a,b]$.

3. Во многих учебниках для равномерной сетки приводят формулы с использованием конечных разностей. Δ - дань историческим традициям. Расчеты легче проводить в разностных отношениях.

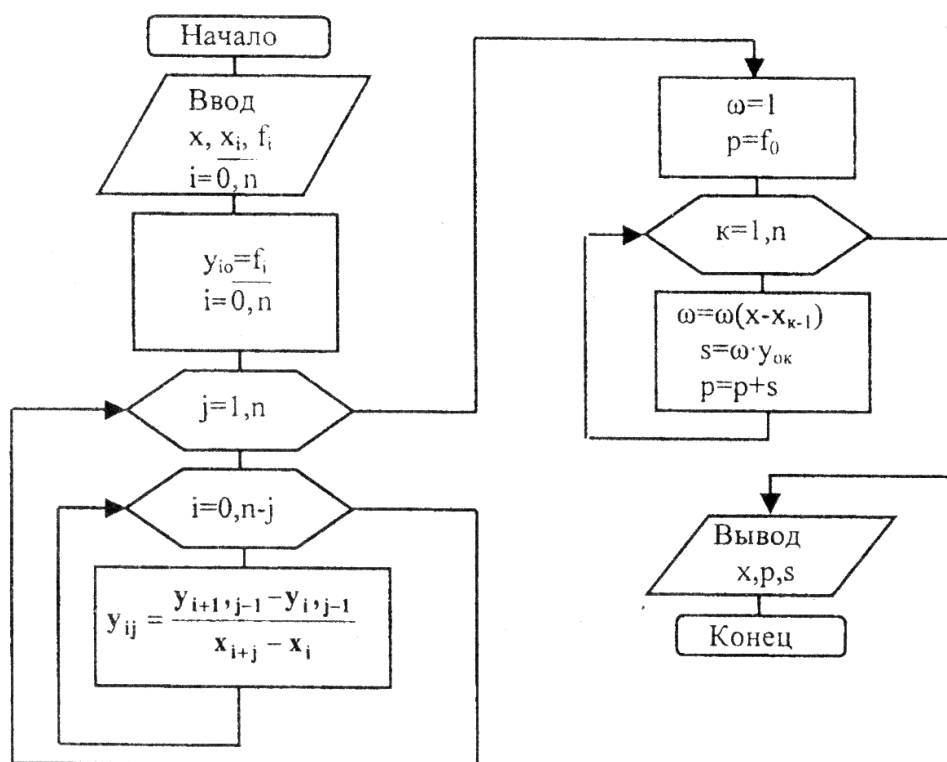


Рисунок 1.19 - Алгоритм интерполяционной формулы Ньютона

1.4.4 Обратное интерполирование

Обратное интерполирование - это нахождение значения x при некотором значении f .

В этом случае таблица читается наоборот: $f(x_i) \rightarrow x(f_i)$, то есть f - аргумент, x - функция.

Частный случай обратного интерполирования - решение уравнения $f(x)=0$. Если функция монотонна, то имеем однозначное соответствие между $f(x_i)$ и $x(f_i)$ (рисунок 1.20).

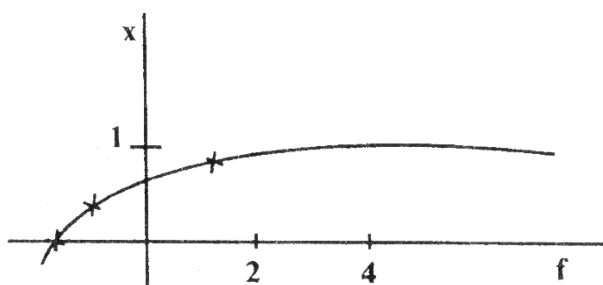


Рисунок 1.20 - График обратного интерполирования

Пример. Решить уравнение $f(x)=(1+x)e^{0,5x}-2,5=0$.

Для этой функции (рисунок 1.20) составлена табличная функция:

i	0	1	2	3
x_i	0	0,5	1,0	2
f_i	-1,5	-0,57	+0,80	5,65

Используем вспомогательную таблицу:

f_i	x_i	$x(f_i, f_{i+1})$	$x(f_0, f_1, f_2)$
-1,5	0	0,54	-0,08
-0,57	0,5	0,36	
0,80	1,0		

Применив 1-ю формулу Ньютона, найдем корень уравнения

$$x(0)=x_0+(0-f_0)x(f_0, f_1)+(0-f_0)(0-f_1)x(f_0, f_1, f_2)=$$

$$0+(0+1,5) \cdot 0,54+(0+1,5)(0+0,57) \cdot (-0,08)=0,74.$$

Точное решение: $x(0)=x^*=0,732$.

Для повышения точности нужно взять новые узлы, близкие к грубо найденному корню, например, $x_4=0,74$.

Существуют более эффективные методы интерполяции, такие, как: приближение Чебышева, сплайн-функции и др., описания которых можно найти в соответствующих учебниках.

1.4.5 Метод наименьших квадратов

Интерполирование позволяет легко приблизить функцию $f(x)$. Однако точность такого приближения гарантирована лишь в интервале нескольких шагов сетки. Для другого интервала приходится заново вычислять коэффициенты интерполяционной формулы. Но хотелось бы иметь формулу, пригодную для большого отрезка $[a, b]$.

Кроме того, если значения табличной функции $f(x_i)$ определены неточно, например, из эксперимента $f_i = f_i^* \pm \epsilon$, то точное приравнивание $f(x_i)$ аппроксимирующей функцией $P(x)$ неразумно. Поэтому целесообразно в этом случае приближать табличную функцию не по точкам, а в среднем. В этом случае говорят об аппроксимации функции.

Конкретизируем понятия интерполяции и аппроксимации функций (рисунок 1.21).

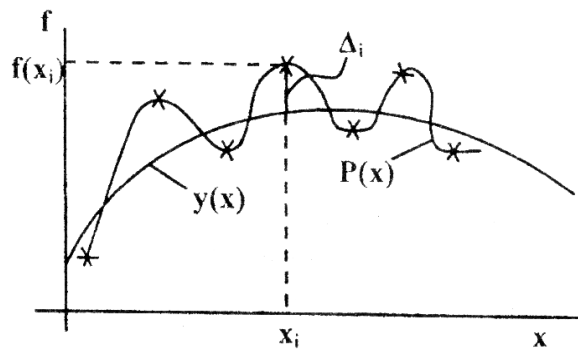


Рисунок 1.21 - Графики интерполирующей и аппроксимирующей функций

Интерполяция - это приближение функции, при котором в узлах сетки интерполирующая функция принимает значения табличной функции

$P(x_i) = f(x_i)$, $i = \overline{1, n}$ - приближение по точкам.

Аппроксимация - это приближение функции, при котором в узлах сетки аппроксимирующая функция может и не принимать значения табличной функции ($y(x)$) - приближение в среднем.

Получим условия для построения одного типа аппроксимирующих функций.

Разность $\Delta_i = y(x_i) - f(x_i)$ называется *невязкой* (рисунок 1.21).

Приближение в среднем означает уменьшение в узлах сетки всех невязок в совокупности. Напрашивается критерий минимизации:

$$\Phi = \sum_{i=1}^n \Delta_i \rightarrow \min$$

- плохой критерий, т.к. знаки невязок будут погашать друг друга.

Несколько лучше $\Phi = \sum_{i=1}^n |\Delta_i| \rightarrow \min$. Однако эту функцию нельзя дифференцировать.

Остановимся на следующем критерии:

$$\Phi = \sum_{i=1}^n (\Delta_i)^2 \rightarrow \min.$$

Задача минимизации этой функции называется *среднеквадратическим приближением*.

Пусть f_i - табличная функция; $y(x)$ - аппроксимирующая функция. Тогда

$$\Phi = \sum_{i=1}^n \rho_i (y(x_i) - f_i)^2 \rightarrow \min. \quad (1.27)$$

Здесь ρ_i - вес i - го эксперимента, соответствующий точности измерения i -го эксперимента. Чем выше точность, тем большее значение веса приписывают точке. Тогда аппроксимирующая кривая будет подходить ближе к точкам с большим весом. Обычно полагают

$$\sum_{i=1}^n \rho_i = 1 \text{ или } \sum_{i=1}^n \rho_i = n.$$

Пусть функция y является функцией с $m+1$ коэффициентами:

$$y=y(x, a_0, \dots, a_m), \quad m < n.$$

Тогда

$$\Phi = \sum_{i=1}^n \rho_i (y(x_i, a_0, \dots, a_m) - f_i)^2 \rightarrow \min.$$

Как видно из уравнения, значение Φ будет определяться численными значениями этих коэффициентов (ρ_i, x_i, f_i - зафиксированные величины). Покажем эту зависимость графически.

Пусть аппроксимирующая функция является линейной функцией $y=a_0+a_1x$. С целью упрощения, полагая $\rho_i=1, \quad i=\overline{1, n}$, имеем

$$\Phi = \sum_{i=1}^n (a_0 + a_1 x_i - f_i)^2 \rightarrow \min.$$

На рисунке 1.22,а показано влияние углового коэффициента a , на величину невязок. Прямая, которая наилучшим образом аппроксимирует табличную функцию, обозначена символом (*). Ордината пересечения этой прямой с осью Оу соответствует свободному члену a_0^* . Через эту точку проведем две прямые с разными углами наклона. Первая прямая $y = a_0^* + a_1^I x$ на рисунке расположена ниже прямой (*), угловой ее коэффициент $a_1^I < a_1^*$. Вторая прямая $y = a_0^* + a_1^{II} x$ имеет угловой коэффициент $a_1^{II} > a_1^*$.

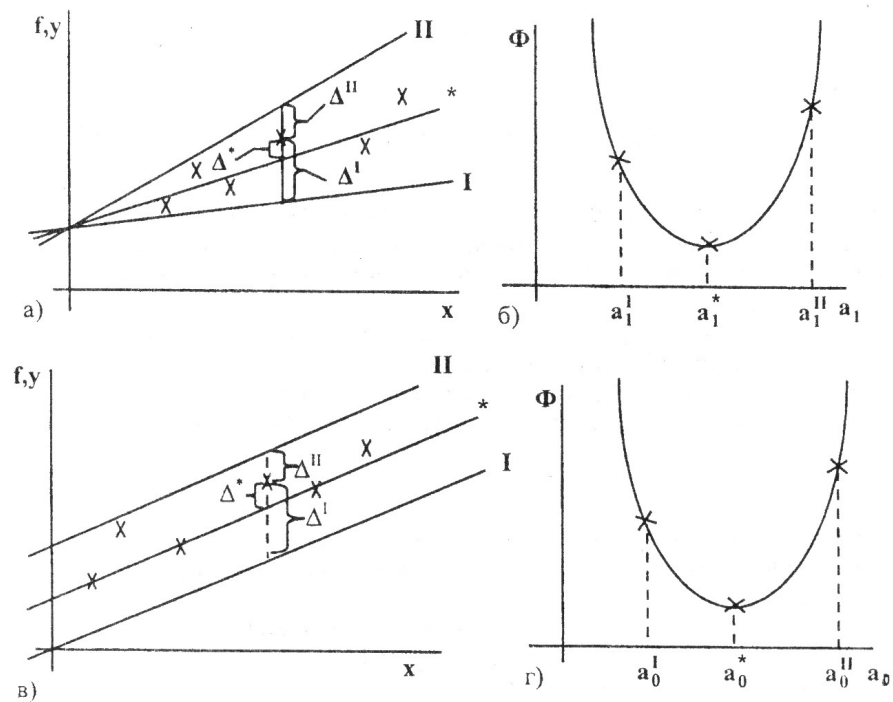


Рисунок 1.22 - Влияние коэффициентов аппроксимирующей функции на величину невязок и на величину критерия Φ

Ввиду того, что прямая (*) расположена наилучшим образом среди значений табличной функции, совокупность абсолютных значений невязок Δ_i^* , $i=\overline{1, n}$ минимальна. Из рисунка видно, что совокупность абсолютных значений невязок прямых (I) и (II) больше, чем прямой (*). Следовательно, значение функций Φ для этих прямых будет больше, чем для прямой (*) (рисунок 1.22,б). Наименьшее значение Φ соответствует прямой (*).

Аналогичным образом при варьировании свободного члена a_0 при постоянном угловом коэффициенте a_1^* (рисунок 1.22,в) совокупность абсолютных значений невязок будет минимальна для прямой (*), для которой значение функции Φ соответственно будет минимальным (рисунок 1.22,г).

Из курса математического анализа известно, что в точке экстремума частные производные функции по переменным (в нашем случае - по коэффициентам) равны нулю.

Тогда в общем случае имеем

$$\frac{\partial \Phi}{\partial a_j} = 2 \sum_{i=1}^n (y(x_i, a_0, \dots, a_m) - f_i) \frac{\partial y}{\partial a_j} = 0, \quad j = \overline{0, m}$$

или

$$\sum_{i=1}^n (y(x_i, a_0, \dots, a_m) - f_i) \frac{\partial y}{\partial a_j} = 0, \quad j = \overline{0, m}. \quad (1.28)$$

Получили систему из $m+1$ уравнений. Решая эту систему относительно неизвестных a_j , можно определить их значения. Описанный способ находже-

ния коэффициентов аппроксимирующей функции называют **методом наименьших квадратов** (МНК).

Получили две разные формулировки решения одной и той же задачи - (1.27) и (1.28). Метод МНК (1.28) используется чаще.

Отметим, что если табличная функция f_i , определена точно, то достаточно было бы взять любые $m+1$ уравнений из переопределенной системы

$$y(x_i) - f(x_i) = 0, \quad i = \overline{1, n}, \quad m < n$$

и рассчитать из новой системы коэффициенты a_j . В этом случае функция y , была бы интерполирующей функцией, т.е. в узлах сетки приняла бы значения табличной функции.

Таким образом, для обеспечения решения (1.28) необходимо выполнение неравенства $m < n$. При этом, если величина m будет слишком малой величиной, то аппроксимирующая функция будет грубой.

Оптимальное число коэффициентов определяют следующим образом. При некотором m находят соответствующие коэффициенты a_j и рассчитывают погрешность аппроксимации δ . Если $\delta \gg \varepsilon$, где ε - погрешность эксперимента, то m надо увеличить. Если $\delta \ll \varepsilon$, то m надо уменьшить. Оптимальное m такое, что $\delta \approx \varepsilon$.

В качестве аппроксимирующей функции y , наиболее часто используют следующие функции.

1. Полиномиальную: $y(x) = a_0 + a_1x + \dots + a_mx^m$.
2. Дробно-линейную: $y(x) = P_k(x)/Q_m(x)$.
3. Экспоненциальную: $y(x) = \sum_k a_k \cdot \exp(e_k x)$.
4. Логарифмическую, тригонометрическую и др.

Рассмотрим более подробно **полиномиальную аппроксимацию**.

Имеем:

$$\Phi = \sum_{i=1}^n (a_0 + a_1x_i + \dots + a_mx_i^m - f_i)^2 \rightarrow \min.$$

Система (1.28) представится в виде

$$\begin{cases} \frac{\partial \Phi}{\partial a_0} = 2 \sum_i (a_0 + a_1x_i + \dots + a_mx_i^m - f_i) \cdot 1 = 0 \\ \frac{\partial \Phi}{\partial a_1} = 2 \sum_i (a_0 + a_1x_i + \dots + a_mx_i^m - f_i) \cdot x_i = 0 \\ \dots \dots \dots \\ \frac{\partial \Phi}{\partial a_m} = 2 \sum_i (a_0 + a_1x_i + \dots + a_mx_i^m - f_i) \cdot x_i^m = 0. \end{cases}$$

Раскрыв скобки, получим систему нормальных уравнений:

$$\begin{cases} na_0 + a_1 \sum_i x_i + \dots + a_m \sum_i x_i^m = \sum_i f_i \\ a_0 \sum_i x_i + a_1 \sum_i x_i^2 + \dots + a_m \sum_i x_i^{m+1} = \sum_i x_i f_i \\ \dots \\ a_0 \sum_i x_i^m + a_1 \sum_i x_i^{m+1} + \dots + a_m \sum_i x_i^{2m} = \sum_i x_i^m f_i \end{cases}$$

Имеем систему линейных уравнений $B \cdot A = F$, $A^T = \{a_0, \dots, a_m\}$, решение которой не вызывает затруднений.

Нередко приходится аппроксимировать экспериментальные данные нелинейной функцией. В этом случае при решении системы (1.28) могут возникнуть трудности. Вычисления можно облегчить линеаризацией исходных уравнений, путем выпрямления искривленной гиперповерхности.

Пример. Экспериментальные данные представлены в виде таблицы:

x_i	100	152	193	228	261	304
z_i	3,23	7,80	15,4	24,2	37,9	60,1
$f_i = \ln z_i$	1,17	2,05	2,73	3,19	3,63	4,10

Графическим анализом можно убедиться, что эти данные удовлетворительно приблизятся экспоненциальной кривой вида $z = v_0 \exp(v_1 x)$.

Для облегчения расчета коэффициентов v_0, v_1 линеаризуем аппроксимирующую функцию путём $\ln z = \ln v_0 + v_1 x$, которую, в свою очередь, обозначим в виде $y(x) = a_0 + a_1 x$. При этом табличная функция будет определяться $f_i = \ln z_i$.

Искомые коэффициенты будут находиться через новые по уравнениям $v_0 = \exp(a_0)$, $v_1 = a_1$.

Минимизируемый функционал примет вид

$$\Phi = \sum_{i=1}^n (a_0 + a_1 x_i - f_i)^2 \rightarrow \min.$$

Система (1.28) и система нормальных уравнений имеет вид

$$\begin{cases} \sum_i (a_0 + a_1 x_i - f_i) = 0 \\ \sum_i (a_0 + a_1 x_i - f_i) \cdot x_i = 0 \end{cases} \Rightarrow \begin{cases} na_0 + a_1 \sum_i x_i = \sum_i f_i \\ a_0 \sum_i x_i + a_1 \sum_i x_i^2 = \sum_i x_i f_i \end{cases}$$

Решая данную систему линейных уравнений, получим

$$a_0 = (\sum f_i - a_1 \sum x_i) / n$$

$$a_1 = (\sum x_i f_i - (\sum x_i)(\sum f_i) / n) / (\sum x_i^2 - (\sum x_i)^2 / n).$$

Для рассматриваемого примера

$$\sum x_i = 1238; \quad \sum x_i^2 = 282874; \quad \sum f_i = 16,87; \quad \sum x_i f_i = 3877$$

$$a_0 = -0,168; \quad a_1 = 0,0144.$$

По рассчитанным значениям a_0, a_1 находятся коэффициенты b_0, b_1 искомой аппроксимирующей экспоненциальной функции:

$$b_0 = \exp(-0,168) = 0,8454; \quad b_1 = 0,0144.$$

Метод наименьших квадратов находит применение также при *сглаживании экспериментальных данных* (сглаживание - получение плавного вида табличной функции путем уменьшения случайных ошибок).

Рассмотрим *линейное сглаживание*.

Возьмем несколько соседних точек и в этом узком интервале построим среднеквадратическую аппроксимацию с одним параметром:

$$\Phi = \sum_i (\tilde{f} - f_i)^2 \rightarrow \min \Rightarrow \sum_i (\tilde{f} - f_i) = 0;$$

$$\tilde{f} \cdot n = \sum f_i; \quad \tilde{f} = 1/n \cdot \sum f_i.$$

Здесь под $\tilde{f}$ понимается точка табличной функции, значение которой исправляется. Так, при $n=3$ имеем

$$f_i = 1/3 \cdot (f_{i-1} + f_i + f_{i+1}).$$

В радиотехнике этот способ сглаживания называется *фильтром*. Пример *кубического сглаживания*:

$$f_i = \frac{1}{35} (-3f_{i-2} + 12f_{i-1} + 17f_i + 12f_{i+1} - 3f_{i+2}).$$

Метод наименьших квадратов успешно используется при решении *переопределенных систем уравнений*.

Рассмотрим такую систему:

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m = f_1 \\ \hline a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m = f_n, \quad m < n. \end{array}$$

Минимизируемый функционал имеет вид

$$\Phi = \sum_{i=1}^n (a_{i1}x_1 + \dots + a_{im}x_m - f_i)^2 \rightarrow \min.$$

Ввиду того, что неизвестными являются переменные $x_j, j = \overline{1, m}$, то в отличие от (1.28), частные производные берутся по этим переменным:

$$\partial \Phi / \partial x_j = 0, \quad j = \overline{1, m}$$

$$\frac{\partial \Phi}{\partial x_1} = 2 \sum_{i=1}^n (a_{i1}x_1 + \dots + a_{im}x_m - f_i) a_{i1} = 0$$

$$\frac{\partial \Phi}{\partial x_m} = 2 \sum_{i=1}^n (a_{i1}x_1 + \dots + a_{im}x_m - f_i) a_{im} = 0.$$

В итоге, после раскрытия скобок, получим определенную систему линейных уравнений с квадратной матрицей:

$$\tilde{A} X = \tilde{F}, \quad \tilde{A} = \{a_{jk}\}_{m,m}$$

Замечания

1. Метод наименьших квадратов не может сам раскрыть истинного характера зависимости. Он только позволяет рассчитать значения коэффициентов конкретной аппроксимирующей функции, которую решил проверить исследователь. Функцию выбирает исследователь, и он же оценивает ее пригодность.

2. Если нет никаких теоретических указаний о виде зависимости, то следует искать наиболее простую формулу, содержащую как можно меньшее количество параметров a_j .

3. Полное совпадение с данными эксперимента не желательно, т.к. аппроксимирующая функция будет повторять ошибки эксперимента.

Вопросы и упражнения для самоконтроля.

1. Что понимают под шагом сетки?
2. Какие типы сеток вы знаете?
3. Как находятся конечные разности различных порядков через значения функции в узловых точках?
4. Как могут быть получены разностные отношения?
5. Что понимают под интерполированием функции?
6. В каких случаях используют термины интерполяция и экстраполяция.
7. В какой форме строится интерполяционный многочлен Лагранжа?
8. Какому условию должен удовлетворять интерполяционный многочлен Лагранжа?
9. В чем заключается особенность получения многочлена по схеме Эйткина?
10. Что означает априорная и апостериорная оценка?
11. Почему первую интерполяционную формулу Ньютона нецелесообразно применять для интерполирования в конце отрезка интерполяции?
12. Что понимают под обратным интерполированием?
13. Дайте определение понятиям интерполяция и аппроксимация.
14. В чем суть приближения таблично заданной функции по методу наименьших квадратов?
15. Чем метод наименьших квадратов отличается от метода интерполяции.
16. Функция задана таблицей значений $\{x_i, y_i\}$ найти ее приближенное значение для x'_1 и x'_2 с помощью полинома Лагранжа, составить формулу полинома, с помощью интерполяционной формулы Ньютона. Априорным путем оценить погрешности в указанных точках. Дать анализ погрешностей.

№ вар	x_0	x_1	x_2	y_0	y_1	y_2	x'_1	x'_2
1	-6	-5	-4	4	-4	5	-5.6	-4.5
2	-3	-1	3	7	1	11	0	2
3	2	3	5	4	1	7	2.5	4
4	0	2	3	-1	-4	2	1	2.5
5	7	9	13	2	-2	3	8	11
6	3.1	3.6	4.1	1.85	1.82	1.74	3.3	3.9
7	-2	-1	2	4	9	1	0.5	1.5
8	3.6	4.2	4.6	1.82	1.74	1.63	3.8	4.4
9	-3	-1	1	-11	-1	6	-2	0.5
10	0	1	3	5	-3	2	0.5	2
11	1	2	3	-3	-7	2	1.5	2.5
12	2	4	6	-1	-6	-3	3	5
13	0	1.5	3	4	-7	1	1	2.5
14	-1	0	3	-3	5	2	2	2.7
15	0.5	0.16	0.17	0.15	0.16	0.17	0.155	0.165
16	2.1	2.6	3.1	1.65	1.8	1.85	2.3	2.9
17	0	3	6	1	4	6	2	5
18	1.1	1.6	2.1	1.03	1.39	1.65	1.4	1.9
19	-4	-2	0	2	8	5	-3	-1
20	7	8	10	6	-2	7	7.5	8.6

17. Найти корень уравнения $\delta^2 \sin \frac{1}{x} - 5 = 0$, находящийся на отрезке $[3;8]$, методом обратной интерполяции с точностью $1 \cdot 10^{-5}$.

18. По данным таблицы построить точечный график и определить подходящий вид приближающей функции:

x	0,11	0,13	0,14	0,17	0,20	0,25	0,33	0,55
y	1,22	1,24	1,30	1,32	1,41	1,54	1,66	2,02

1.5 Численное дифференцирование

С дифференцированием встречаются на практике, когда необходимо найти производные, имея в распоряжении только табличные значения функции (например, экспериментальные данные).

Простейшие формулы численного дифференцирования получаются в результате дифференцирования интерполяционных формул. Если интерполяционный многочлен на рассматриваемом участке с достаточной точностью совпадает с заданной функцией, а сама функция достаточно гладкая и плавно изменяется на рассматриваемом участке, то можно считать, что производная интерполяционного многочлена также мало отличается от требуемой производной.

При этом следует предположить, что расстояние между узлами интерполяции достаточно мало, чтобы функция не имела между ними большого числа экстремумов. В противном случае может оказаться (рисунок 1.23), что разность между значениями функции и интерполяционного многочлена мала, тогда как их производные не имеют между собой ничего общего.

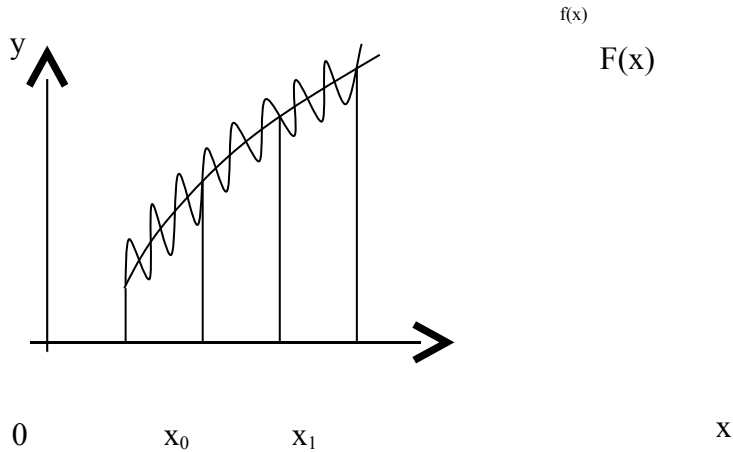


Рисунок 1.23

Если рассматриваемая функция плавно изменяется между узлами интерполяции, то для отыскания производной функцию, заданную таблично, следует заменять интерполяционным многочленом.

Пусть функция $f(x)$ определена на отрезке $[a, b]$ в виде табличной функции $f(x_i)$. Пусть узлы x_i сетки Ω_h различны. Интерполируем f_i посредством полинома $P_n(x)$. Тогда

$$f(x) = P_n(x) + R_n(x).$$

Если на отрезке $[a, b]$ функция $f(x)$ имеет непрерывную k -ю производную, то

$$f^{(k)}(x) = P_n^{(k)}(x) + R_n^{(k)}(x), \quad k \leq n.$$

Полагая $R_n^{(k)} \approx 0$, имеем

$$f^{(k)}(x) \approx P_n^{(k)}(x). \quad (1.29)$$

Можно показать, что погрешность будет

$$R_n^{(k)} = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega^{(k)}(x) \quad \text{при} \quad \omega^{(k)}(x) \neq 0. \quad (1.30)$$

В качестве интерполирующего полинома возьмем формулу Ньютона:

$$P_n(x) = f(x_0) + (x - x_0)f(x_0, x_1) + (x - x_0)(x - x_1)f(x_0, x_1, x_2) + \dots$$

Последовательное дифференцирование этого полинома дает приближенное выражение производных:

$$f'(x) \approx P'_n(x) = f(x_0, x_1) + ((x-x_0) + (x-x_1))f(x_0, x_1, x_2) + \dots + ((x-x_0)(x-x_1) + (x-x_0)(x-x_2) + (x-x_1)(x-x_2))f(x_0, x_1, x_2, x_3) + \dots$$

$$f^{(k)}(x) \approx P_n^{(k)}(x) = k! \left[f(x_0, \dots, x_k) + \left(\sum_{i=0}^k (x-x_i) \right) f(x_0, \dots, x_{k+1}) + \dots \right]. \quad (1.31)$$

Частный случай

$$f'(x) \approx f(x_0, x_1) = \frac{f_1 - f_0}{x_1 - x_0} = \frac{\Delta f}{\Delta x}.$$

Если $\Delta x \rightarrow 0$, то приближенное равенство заменяется точным.

Замечания

1. Оценка (1.30) - априорная оценка. На практике используют апостериорную оценку — по величине первого отброшенного члена.

2. С ростом k сходимость ряда (1.31) ухудшается. В связи с этим численным дифференцированием можно хорошо определить первую и вторую производные. Третью и четвертую - удовлетворительно, более высшего порядка - плохо.

3. Если сетка равномерная, то предпочтительно использовать следующие уравнения:

$$f'(x) = \frac{f_1 - f_0}{h}; \quad f'(x) = \frac{f_2 - f_0}{2h}; \quad f''(x) = \frac{f_2 - 2f_1 + f_0}{h^2} \quad \text{и т.д.}$$

Вопросы и упражнения для самоконтроля.

1. Что понимают под дифференцированием?
2. Как решается задача численного дифференцирования для таблично заданной функции с помощью интерполирования?

3. Найти значения производных в точках:

а) $f(x)=2x^2+1$, $x=-12$; б) $f(x)=\sin(x)$, $x=\frac{\pi}{2}$; в) $f(x)=10x^3+1$, $x=3$;

г) $f(x)=e^{2x}$, $x=5$; д) $f(x)=\sqrt{1-2\delta}$, $x=-8$; е) $f(x)=\sin(x) \cdot \ln(x)$, $x=2$;

ж) $f(x)=\frac{\cos(x)}{e^x}$, $x=0,005$; з) $f(x)=\cos \frac{\delta}{2} - \frac{1}{\delta}$, $x=0,2$;

и) $f(x)=x^3 \cos(x) - \sin(2x)$, $x=0,5$; к) $f(x)=4x^3 - 2x^2 + 3$, $x=-7$;

л) $f(x)=3x - x(x^3+9)$, $x=-2$; м) $f(x)=6x - 12x^3$, $x=4$;

н) $f(x)=-5x^2+2x$, $x=-10$; о) $f(x)=\frac{1}{\delta} \cos(4x)$, $x=1$;

п) $f(x)=e^{2x}\sin(5x)$, $x=0,6$; р) $f(x)=x+4x^3-5$, $x=13$:

с) $f(x)=e^{-x}+4x^2$, $x=3$; т) $f(x)=\ln(x^2+4x+4)$, $x=5$.

4. Пусть функция $f(x)=e^x$ задана четырехзначной таблицей с шагом $h=0,1$. Вычислить производную $f'(1,06)$.

x	f(x)	Δy	$\Delta^2 y$	$\Delta^3 y$
0,6	1,8221	1917		
0,7	2,0138	2117	200	
0,8	2,2255	2341	224	24
0,9	2,4596	2587	246	22
1,0	2,7183	2859	272	26
1,1	3,0042	3159	300	28
1,2	3,3201			

1.6 Численное интегрирование

1.6.1 Понятие квадратурной формулы

Согласно определению интеграла

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

Геометрической интерпретацией определенного интеграла является площадь, ограниченная кривой $f(x)$, осью Ox и прямыми $x=a$ и $x=b$.

Далеко не каждый интеграл имеет аналитическое решение. Чаще решение определяют приближенно, с некоторой погрешностью.

Приближенные формулы вычисления интегралов называют *квадратурными формулами*. Для расчета интеграла функции многих переменных разработаны m -мерные кубатуры. Это название связано с тем, что в m -мерном пространстве единицей объема является m -мерный куб. Мы рассмотрим только квадратурные формулы, т.к. кратный интеграл можно представить в виде

$$\int_{\Omega_m} f(x_1, \dots, x_m) dx_1 \dots dx_m = \int_{\Omega_{x_1}} dx_1 \int_{\Omega_{x_2}} \dots \int_{\Omega_{x_m}} f(x_1, \dots, x_m) dx_m$$

и каждый интеграл вычислить с помощью квадратурной формулы. Хотя отметим - это приведет к увеличению объема вычислений (в сравнении с расчетом исходного интеграла).

Большинство квадратурных формул можно представить в виде

$$\int_a^b f(x) dx \approx \sum_{i=0}^n A_i f(x_i), \quad (1.32)$$

где x_i -узлы квадратурной формулы, $x_i \in [a, b]$; A_i - веса квадратурной формулы.

Формула (1.32) содержит $2(n+1)$ параметра: $A_i, x_i, i=\overline{0, n}$. Их нужно выбрать так, чтобы формула давала, возможно лучший результат.

1.6.2 Интерполяционные квадратуры

Пусть $f(x)$ достаточно гладкая и многократно дифференцируемая функция. Из теоремы Вейерштрасса следует, что любую непрерывную функцию можно сколь угодно точно приблизить полиномами (обычными или тригонометрическими). В частности, используем интерполяционный полином Лагранжа

$$f(x) = \sum_{i=0}^n f(x_i) \omega_i(x) + R_n(x).$$

Проинтегрируем это уравнение:

$$\int_a^b f(x) dx = \sum_{i=0}^n f(x_i) \int_a^b \omega_i(x) dx + \int_a^b R_n(x) dx.$$

Обозначим

$$A_i = \int_a^b \omega_i(x) dx; \quad r(x) = \int_a^b R_n(x) dx, \quad (1.33)$$

где $r(x)$ - погрешность квадратурной формулы. Полагая, что она является малой величиной, имеем

$$\int_a^b f(x) dx \approx \sum_{i=0}^n f(x_i) A_i.$$

Отличительной особенностью $\omega_i(x)$ является то, что этот полином не зависит от значения функции $f(x)$ в узлах сетки. Следовательно, для фиксированных узлов $x_0, \dots, x_n$, независимо от $f(x)$, величины $A_i = \text{const}$. Таким образом, вычисление интеграла будет заключаться в расчете $f(x_i)$ и некоторых постоянных весов A_i . В справочной литературе приводятся численные значения A_i . При разных n они различны.

Оценку погрешности получим, исходя из оценки погрешности интерполяционной формулы Лагранжа. Погрешность интерполяционного полинома Лагранжа

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega(x), \quad \xi \in [a, b].$$

Следовательно,

$$r(x) = \int_a^b R(x) dx = \int_a^b \frac{f^{(n+1)}(x)}{(n+1)!} \omega(x) dx.$$

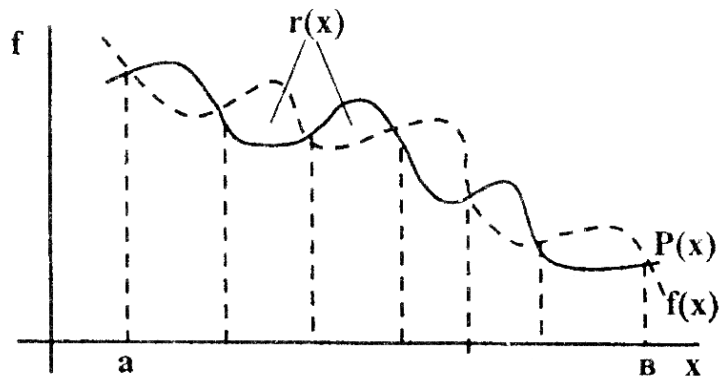


Рисунок 1.24 - Геометрическая интерпретация погрешности

Пусть $|f^{(n+1)}(x)| < M$. Тогда

$$r(x) = \frac{M}{(n+1)!} \int_a^b \omega(x) dx. \quad (1.34)$$

Здесь имеются те же недостатки, которые сопутствуют интерполяционному полиному Лагранжа:

- 1) сложность и неустойчивость решений при больших n ;
- 2) с ростом n величины A_i принимают разные знаки и неограниченно возрастают, что приводит к существенному возрастанию ошибок округления.

Эти трудности удастся исключить, используя *кусочно-квадратурные формы*. Идея способа заключается в следующем. Исходный отрезок делят на N малых отрезков точками X_k .

Тогда

$$\int_a^b f(x) dx = \sum_{k=1}^N \int_{x_{k-1}}^{x_k} f(x) dx,$$

где $a = X_0 < X_1 < \dots < X_N = b$, $X_k = a + kH$, $H = (b-a)/N$, $k = \overline{0, n}$.

Далее внутри отрезков $[X_{k-1}, X_k]$ располагают соответствующие узлы x_i $X_{k-1} = x_0 < x_1 < \dots < x_n = X_k$ (рисунок 1.25) и интегрирование проводят внутри этих

отрезков с малыми n . Искомый результат получается суммированием N рассчитанных значений интеграла.

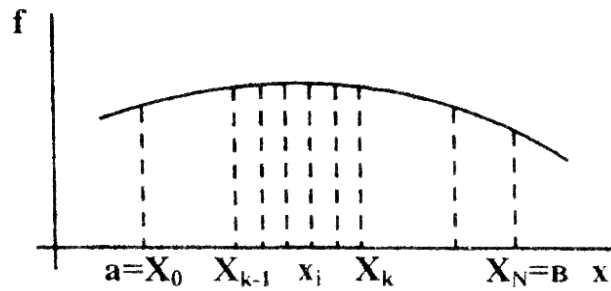


Рисунок 1.25 - Построение кусочно-квадратурной формы

При анализе точности расчетов обычно n оставляют неизменным, а варьируют N .

Исходя из квадратурной формулы (1.34) выведем часто используемые на практике формулы.

Метод трапеций (рисунок 1.26,а)

Пусть $n=1$. Тогда имеем $h=b-a=x_1-x_0$

$$\int_a^b f(x) dx \approx \sum_{i=0}^1 A_i f(x_i) = A_0 f(x_0) + A_1 f(x_1).$$

Значения весов равны:

$$A_0 = \int_a^b \frac{x-x_1}{x_0-x_1} dx = \int_a^b \frac{x-b}{a-b} dx = -\frac{1}{2h} (x-b)^2 \Big|_a^b = -\frac{1}{2h} ((b-b)^2 - (a-b)^2) = \frac{h}{2},$$

$$A_1 = \int_a^b \frac{x-x_0}{x_1-x_0} dx = \frac{1}{2h} (x-a)^2 \Big|_a^b = \frac{1}{2h} ((b-a)^2 - (a-a)^2) = \frac{h}{2}.$$

Таким образом, значение интеграла

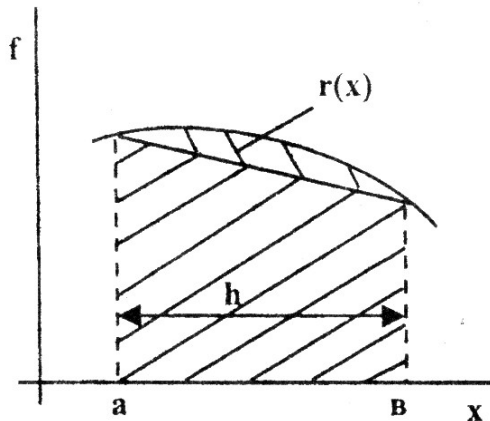
$$\int_a^b f(x) dx \approx \frac{h}{2} f(a) + \frac{h}{2} f(x_1) = \frac{h}{2} (f(a) + f(b)). \quad (1.35a)$$

Погрешность метода

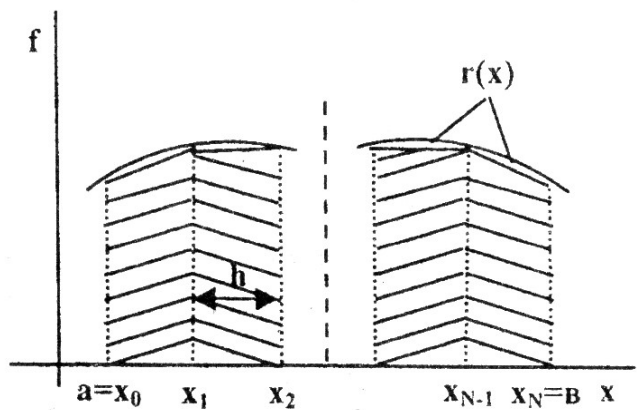
$$r(x) = \frac{M}{2!} \int_a^B (x - x_0)(x - x_1) dx = \left| \begin{matrix} x_0 = a = 0 \\ x_1 = B = h \end{matrix} \right| = \frac{M}{2} \int_0^h x(x - h) dx = -\frac{M}{12} h^3.$$

Погрешность метода пропорциональна длине шага в третьей степени. Для большого интервала погрешность может быть очень большой. Так, если $h=10$, то $r(x) \approx 10^3$. В таких случаях используют кусочно-квадратурные формы. Для метода трапеций они строятся следующим образом (рисунок 1.26,б):

$$\begin{aligned} \int_a^B f(x) dx &= \sum_{k=1}^N \frac{h}{2} (f(x_{k-1}) + f(x_k)) = h \left(\frac{1}{2} (f_0 + f_1) + \frac{1}{2} (f_1 + f_2) + \dots + \frac{1}{2} (f_{N-1} + f_N) \right) = \\ &= h \left(\frac{1}{2} (f_0 + f_N) + \sum_{k=1}^{N-1} f_k \right). \end{aligned} \quad (1.356)$$



а)



б)

Рисунок 1.26 - Иллюстрация метода трапеций

Метод прямоугольников

Если в качестве узла взять середину отрезка $[a, b]$, то имеем

$$\int_a^B f(x) dx \approx h f(\bar{x}), \quad h = B - a, \quad \bar{x} = (B + a)/2. \quad (1.36a)$$

Для
ков имеем

N малых отрез-

$$\int_a^B f(x) dx \approx h \sum_{k=1}^N f(\bar{x}_k); \quad h = H = (B - a)/N. \quad (1.36б)$$

Можно показать, что $r(x) = Mh^3/24$, т.е. погрешность одного порядка, что и у метода трапеций, но в 2 раза точнее.

Метод Симпсона (парабол) (рисунок 1.27)

Пусть $n=2$: $a=x_0 < x_1 < x_2=b$. Квадратурная формула имеет вид

$$\int_a^b f(x) dx \approx \sum_{i=0}^2 A_i f(x_i) = A_0 f_0 + A_1 f_1 + A_2 f_2.$$

Определим значения весов квадратурной формулы:

$$\begin{aligned} A_0 &= \int_a^b \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} dx = \left. \begin{array}{l} x_0 = a = 0 \\ x_1 = h \\ x_2 = b = 2h \end{array} \right| = \int_0^{2h} \frac{(x-h)(x-2h)}{(0-h)(0-2h)} dx = \\ &= \frac{1}{2h^2} \left(\frac{x^3}{3} - \frac{3h}{2} x^2 + 2h^2 x \right) \Big|_0^{2h} = \frac{1}{2h^2} \left(\frac{8h^3}{3} - \frac{12h^3}{2} + 4h^3 \right) = \frac{h}{3}. \end{aligned}$$

$$\begin{aligned} A_1 &= \int_a^b \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} dx = \int_0^{2h} \frac{(x-0)(x-2h)}{(h-0)(h-2h)} dx = -\frac{1}{h^2} \int_0^{2h} (x^2 - 2hx) dx = \\ &= -\frac{1}{h^2} \left(\frac{x^3}{3} - \frac{2h}{2} x^2 \right) \Big|_0^{2h} = -\frac{1}{h^2} \left(\frac{8h^3}{3} - 4h^3 \right) = -h \frac{(8-12)}{3} = \frac{4}{3} h. \end{aligned}$$

$$\begin{aligned} A_2 &= \int_a^b \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} dx = \int_0^{2h} \frac{(x-0)(x-h)}{2h(2h-h)} dx = \frac{1}{2h} \int_0^{2h} (x^2 - hx) dx = \\ &= \frac{1}{2h^2} \left(\frac{x^3}{3} - \frac{hx^2}{2} \right) \Big|_0^{2h} = \frac{1}{2h^2} \left(\frac{8h^3}{3} - \frac{4h^3}{2} \right) = \frac{h^3}{2h^2} \left(\frac{16-12}{6} \right) = \frac{h}{3}. \end{aligned}$$

Таким образом,

$$\int_a^B f(x)dx \approx \frac{h}{3} (f(x_0) + 4f(x_1) + f(x_2)). \quad (1.37a)$$

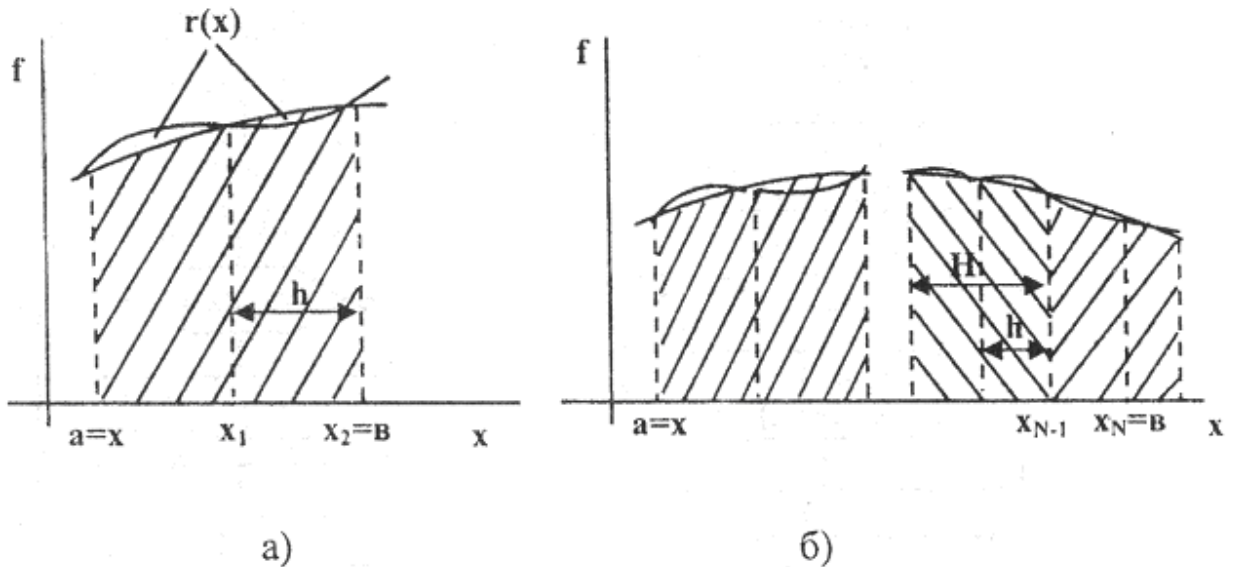


Рисунок 1.27 - Иллюстрация метода Симпсона

Можно показать, что погрешность метода $r(x) \approx h^4$.

Кусочно-квадратурные формы для этого метода строятся следующим образом:

$$\int_a^B f(x)dx \approx \frac{h}{3} (f_0^1 + 4f_1^1 + f_2^1) + \frac{h}{3} (f_0^2 + 4f_1^2 + f_2^2) + \dots + \frac{h}{3} (f_0^N + 4f_1^N + f_2^N).$$

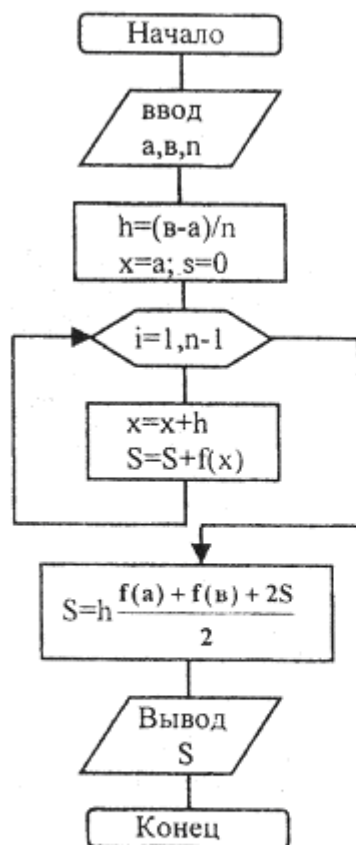
Здесь верхний индекс соответствует номеру малого отрезка $[x_{k-1}, x_k]$.

Возьмем сквозную нумерацию узлов, обозначив $\bar{N} = 2N$. Тогда

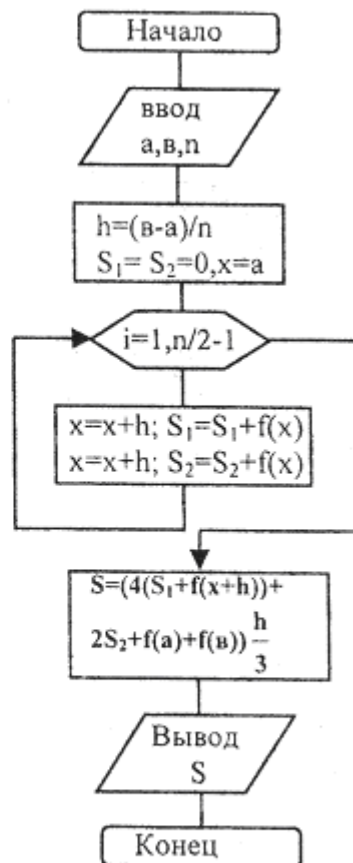
$$\int_a^B f(x)dx \approx \frac{h}{3} (f_0 + 4f_1 + 2f_2 + 4f_3 + 2f_4 + \dots + 2f_{\bar{N}-2} + 4f_{\bar{N}-1} + f_{\bar{N}});$$

$$\int_a^B f(x)dx \approx \frac{h}{3} [f(a) + f(b) + 4 \sum_{k=1}^{\bar{N}/2} f(x_{2k-1}) + 2 \sum_{k=1}^{\bar{N}/2-1} f(x_{2k})]. \quad (1.37б)$$

Здесь $h = (b-a) / \bar{N}$, где $\bar{N}$ - четное число.



а)



б)

Рисунок 1.28 - Алгоритмы метода трапеций (а) и Симпсона (б)

1.6.3 Вычисление несобственных интегралов

$$\int_a^b f(x) dx$$

Интеграл называется *собственным*, если:

- 1) промежуток интегрирования $[a, b]$ конечен;
- 2) функция $f(x)$ непрерывна на $[a, b]$,

в противном случае называется *несобственным* интегралом.

1. Интеграл с бесконечным промежутком интегрирования.

Такой интеграл называется *сходящимся*, если

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

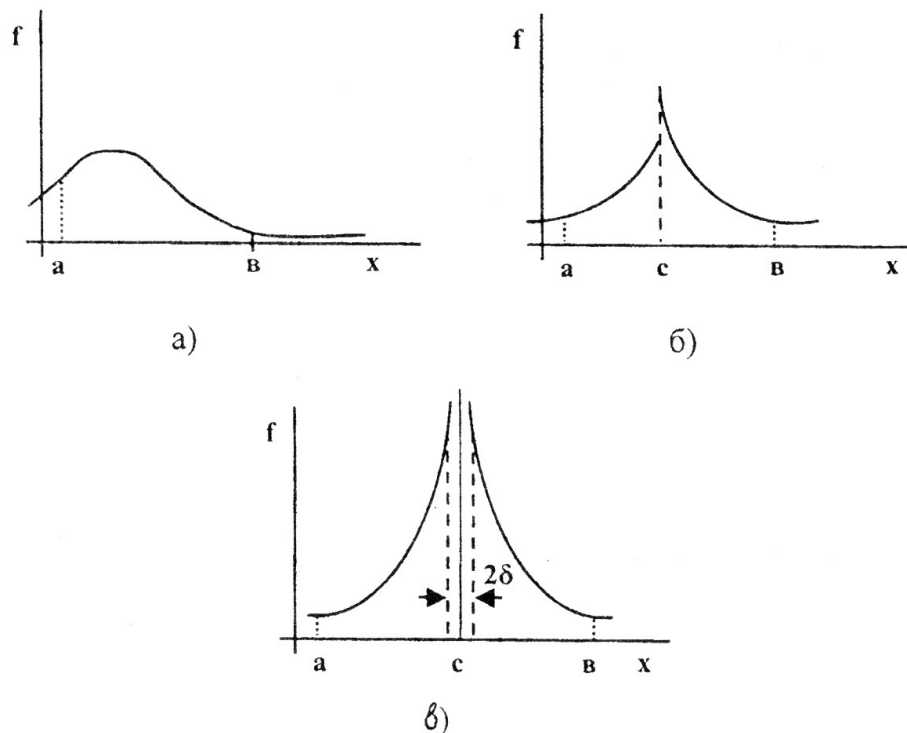


Рис. 1.29 - Графики подынтегральных функций

Если интеграл не сходится, то нет смысла его вычислять. Чтобы вычислить сходящийся несобственный интеграл с заданной точностью ε , представим его в виде (рисунок 1.29,а)

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

Число b выберем столь большим, чтобы имело место неравенство

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

Тогда последний интеграл отбрасывается и вычисляется лишь интеграл

$$\int_a^b f(x) dx$$

с использованием одной из квадратурных формул.

2. Интервал $[a, b]$ конечен, но функция $f(x)$ имеет конечное число *точек разрыва* на отрезке $[a, b]$. Пусть для определенности имеется одна точка разрыва.

а) Разрыв первого рода.

Исходный интеграл расчлняют на 2 интеграла (рисунок 1.29,б):

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

и рассчитывают каждый интеграл в отдельности, используя одну из квадратурных формул.

б) Разрыв второго рода.

Исходный интеграл расчлняют в этом случае следующим образом (рисунок 1.28,в):

$$\int_a^B f(x) dx = \int_a^{c-\delta} f(x) dx + \int_{c+\delta}^B f(x) dx,$$

где δ - малая положительная величина. Ее выбирают так, чтобы функция в точке $c \pm \delta$ была ограничена.

Вопросы и упражнения для самоконтроля.

- 1) Как называются приближенные формулы вычисления интегралов?
- 2) Как связаны задачи численного интегрирования и интерполирования?
- 3) Чем объясняется название формулы трапеции?
- 4) В чем выражается преимущество формулы Симпсона перед формулой трапеций?
- 5) Дайте определение собственному интегралу.

6) Вычислить значение интеграла $I = \int_0^4 f(x) dx$, методами трапеций, прямоугольников, Симпсона при $h=1$. В качестве под интегральных функций взять:

а) $f(x) = x^2 + 1$; б) $f(x) = 0,5 + x \lg x$; в) $f(x) = 1 + x^4$ г) $f(x) = \sqrt{\frac{2-x}{x-8}}$;

д) $f(x) = x^2 + 4x + 4$; е) $f(x) = 5x + \ln x$; ж) $f(x) = x^2 \sqrt{2-x^2}$;

з) $f(x) = \frac{\sin x}{x}$; и) $f(x) = (x+1)(x-5)$; к) $f(x) = 5xe^{\cos x}$;

л) $f(x) = \frac{\exp\left(\sqrt{\frac{2-x}{x+2}}\right)}{(2+x)\sqrt{4-x^2}}$; м) $f(x) = x^4 \operatorname{tg} \frac{x}{4}$; н) $f(x) = \frac{1}{x^2 \sqrt{x^2+1}}$;

о) $f(x) = (x-0,4) \ln \frac{x}{2}$; п) $f(x) = \frac{\sin^2 x}{x^3}$; р) $f(x) = \frac{1}{\pi \sqrt{16-x^2}}$;

с) $f(x) = \sqrt{9-x^2}$; т) $f(x) = \frac{\sin(0,1x-3)}{x^2-1}$.

7) Вычислить несобственный интеграл по неограниченному промежутку непосредственно и через предел. Постройте график функции, заданной интегралом с переменным пределом.

$$\begin{aligned}
& \text{а) } \int_1^{\infty} \frac{dx}{x^2 + 4}; \quad \text{б) } \int_1^{\infty} \ln^2 x dx; \quad \text{в) } \int_1^{\infty} \frac{dx}{1 + x^2}; \quad \text{г) } \int_1^{\infty} x e^{-x} dx; \quad \text{д) } \int_1^{\infty} \frac{\sqrt{x} dx}{(1 + x)^2}; \\
& \text{е) } \int_1^{\infty} \frac{1}{4} e^{-x^2} dx; \quad \text{ж) } \int_1^{\infty} \frac{dx}{1 - x + x^2}; \quad \text{з) } \int_1^{\infty} \frac{dx}{x \ln^2 x}; \quad \text{и) } \int_1^{\infty} \frac{dx}{x^2 + 9x + 9}; \\
& \text{к) } \int_1^{\infty} x^2 e^{-x} dx; \quad \text{л) } \int_1^{\infty} \frac{dx}{x^4 + 2}; \quad \text{м) } \int_1^{\infty} 6x \ln x dx; \quad \text{н) } \int_1^{\infty} \frac{e^{-x}}{\sqrt{x}} dx; \\
& \text{о) } \int_1^{\infty} \frac{x + 1}{x^2 + 1} dx; \quad \text{п) } \int_1^{\infty} \frac{x dx}{1 + x^4}; \quad \text{р) } \int_1^{\infty} \sqrt{x} e^{-x} dx; \quad \text{с) } \int_1^{\infty} \frac{x dx}{(1 + x)^2}; \quad \text{т) } \int_1^{\infty} \frac{dx}{x \sqrt{\ln x}}.
\end{aligned}$$

2 Решение систем алгебраических уравнений

2.1 Решение систем линейных уравнений

2.1.1 Основные понятия

Система линейных уравнений

$$AX=B, \quad A_{m,n}, \quad X_{1,n}^T, \quad B_{m,1} \quad (2.1)$$

называется определенной, если число уравнений системы равно числу неизвестных $m=n$. Здесь m - число уравнений; n - число неизвестных.

Если число уравнений меньше числа неизвестных ($m < n$), - не доопределенная система, и, если число уравнений больше числа неизвестных ($m > n$), - переопределенная система. Мы рассмотрим лишь вопросы решения определенной системы.

Важной характеристикой определенных систем линейных уравнений является величина определителя матрицы A - $\det A$.

Пример. Величина определителя следующей системы второго порядка равна:

$$\begin{aligned}
x_1 + 2x_2 &= 3 \\
2x_1 + 5x_2 &= 7
\end{aligned}
\quad \det A = 1 \cdot 5 - 2 \cdot 2 = 1.$$

Каждое уравнение системы представляет собой $n-1$ - мерную плоскость. Решением системы является точка пересечения этих плоскостей X^* . Однако плоскости могут и не пересекаться. Судить о взаимном расположении плоскостей можно по величине определителя. На рисунке 2.1 приведены возможные расположения прямых, соответствующих уравнениям системы второго порядка ($n=2$). Здесь представлены следующие случаи.

1) $\det A \neq 0$ - система имеет единственное решение X^* . В этом случае говорят, что система обусловлена (не вырождена).

2) $\det A=0$ - система не имеет решения или имеет бесконечное число решений (прямые на рисунке параллельны друг другу или совпадают). Говорят, что система не обусловлена (вырождена).

3) $\det A \approx 0$ - система плохо обусловлена (почти вырождена). Незначительные ошибки округления могут привести к тому, что величина определителя может стать равной нулю.

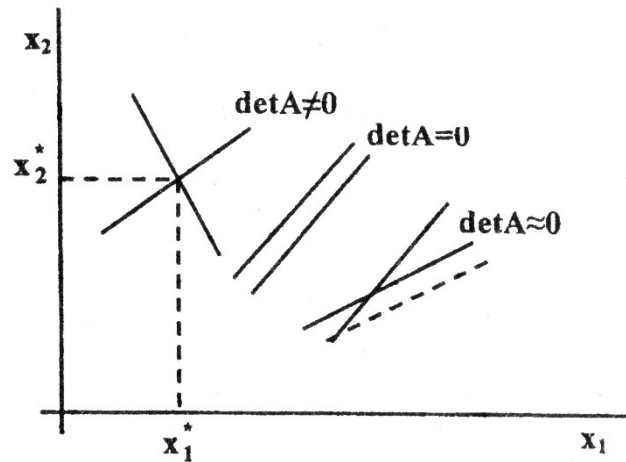


Рисунок 2.1 - Возможное взаимное расположение плоскостей

Матрицы вида

$$A = \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ a_{ij} & & a_{nn} \end{bmatrix} \quad \tilde{A} = \begin{bmatrix} \tilde{a}_{11} & \tilde{a}_{ij} \\ & \ddots & \\ 0 & & \tilde{a}_{nn} \end{bmatrix} \quad C = \begin{bmatrix} c_{11} & c_{12} & 0 \\ & \ddots & \\ 0 & c_{n,n-1} & c_{n,n} \end{bmatrix} \quad D = \begin{bmatrix} d_{11} & 0 \\ & \ddots & \\ 0 & & d_{nn} \end{bmatrix}$$

называются соответственно нижней треугольной, верхней треугольной, ленточной и диагональной матрицами. Определители матриц $A, \tilde{A}, D$ рассчитываются просто:

$$\det A = a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn}; \quad \det \tilde{A} = \tilde{a}_{11} \cdot \tilde{a}_{22} \cdot \dots \cdot \tilde{a}_{nn}; \quad \det D = d_{11} \cdot d_{22} \cdot \dots \cdot d_{nn}$$

Решение систем вида $DX = B$, где D - диагональная матрица, находится также просто:

$$x_i^* = b_i / d_{ii}, i = \overline{1, n}.$$

Ниже будет показано, что также не вызывает затруднений решение систем с треугольной матрицей.

Различают точные методы решения систем уравнений, итерационные и вероятностные. К точным методам относятся метод Крамера и метод исключения Гаусса.

2.1.2 Метод Крамера

В этом методе решением является отношение

$$x_i^* = \det A_i / \det A, \quad (2.2)$$

где A_i матрица, образованная из матрицы A заменой i -го столбца на вектор свободных членов B .

Пример.

$$2x_1 + 4x_2 = 8$$

$$x_1 + 5x_2 = 7$$

$$\det A = 2 \cdot 5 - 4 \cdot 1 = 6$$

$$\det A_1 = 8 \cdot 5 - 4 \cdot 7 = 12$$

$$\det A_2 = 2 \cdot 7 - 8 \cdot 1 = 6$$

$$x_1^* = \frac{12}{6} = 2; \quad x_2^* = \frac{6}{6} = 1$$

Для нахождения решения требуется около $n^2 \cdot n!$ арифметических действий. Уже при $n=20$ число вычислений достигает порядка 10^{21} , что практически исключает возможность применения метода даже с использованием современных ЭВМ.

2.1.3 Метод исключения Гаусса

Идея метода заключается в приведении квадратной матрицы системы (2.1) к треугольному виду. Метод состоит из двух этапов:

- 1) прямой ход - приведение матрицы к треугольному виду;
- 2) обратный ход - решение системы с треугольной матрицей.

1. Прямой ход. Предполагается, что $\det A \neq 0$.

Шаг 1. Обнуление поддиагональных элементов 1-го столбца. Умножим первую строку на коэффициент $m_{i1} = -a_{i1}/a_{11}$ и прибавим к i -ой строке:

$$a_{ij}^{(1)} = a_{ij}^{(0)} + m_{i1} \cdot a_{1j}^{(0)}; \quad \hat{a}_i^{(1)} = \hat{a}_i^{(0)} + m_{i1} \cdot \hat{a}_1^{(0)}; \quad i = \overline{2, n}, \quad j = \overline{1, n}.$$

После завершения шага 1 получим систему, эквивалентную исходной (2.1) и с элементами a_{i1} , $i = \overline{2, n}$, равным нулю:

$$A^{(1)} = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \dots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{bmatrix}$$

В этой системе элементы первой строки остались неизменными, а элементы нижележащих строк путем элементарного преобразования видоизменились. Однако решение системы осталось неизменным.

В вышеприведенных обозначениях элементов верхние индексы соответствуют номеру шага.

Шаг 2. Умножим вторую строку на коэффициент $m_{i2} = -a_{i2}^{(1)} / a_{22}^{(1)}$ и прибавим к i — ой строке:

$$a_{ij}^{(2)} = a_{ij}^{(1)} + m_{i2} \cdot a_{2j}^{(1)}; \quad \hat{a}_i^{(2)} = \hat{a}_i^{(1)} + m_{i2} \cdot \hat{a}_2^{(1)}; \quad i = \overline{3, n}, \quad j = \overline{2, n}$$

В результате получим систему, также эквивалентную исходной и с матрицей, у которой поддиагональные элементы как первого, так и второго столбца равны нулю.

$$A^{(2)} = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \dots & a_{1n}^{(0)} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ 0 & 0 & \dots & a_{nn}^{(2)} \end{bmatrix}$$

Шаг К. В общем случае уравнения, по которым осуществляются преобразования строк, имеют вид

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} + m_{ik} \cdot a_{kj}^{(k-1)}; \quad \hat{a}_i^{(k)} = \hat{a}_i^{(k-1)} + m_{ik} \cdot \hat{a}_k^{(k-1)}; \quad (2.3)$$

$$m_{ik} = -a_{ik}^{(k-1)} / a_{kk}^{(k-1)}; \quad k = \overline{1, n-1} \quad i = \overline{k+1, n} \quad j = \overline{k, n}$$

Здесь на каждом шаге $a_{kk} \neq 0$. После $n-1$ - го шага получим систему с верхней треугольной матрицей.

$$A^{(n-1)} = \begin{bmatrix} a_{11}^{(0)} & & a_{1j}^{(k)} \\ & \ddots & \\ 0 & & a_{nn}^{(n-1)} \end{bmatrix}$$

2. Обратный ход. Последняя строка полученной системы позволяет определить значение x_n^* :

$$x_n^* = \hat{a}_n^{(n-1)} / a_{nn}^{(n-1)}.$$

Значение x_n^* можно использовать далее для расчета x_{n-1}^* подставив в $n-1$ - ю строку:

$$x_{n-1}^* = (\hat{a}_{n-1}^{(n-2)} - a_{n-1,n}^{(n-2)} \cdot x_n^*) / a_{n-1,n-1}^{(n-2)}.$$

Затем последовательно вычисляются $x_{n-2}^*, \dots, x_1^*$, по формуле

$$x_n^* = (\hat{a}_k^{(k-1)} - \sum_{j=k+1}^n a_{kj}^{(k-1)} x_j^*) / a_{kk}^{(k-1)}, \quad k = \overline{n-1, 1} \quad (2.4)$$

Замечания

1. Исключение Гаусса нельзя проводить, если в ходе расчета на главной диагонали оказался нулевой элемент $a_{kk}=0$. В этом случае необходимо использовать модификацию метода с выбором главного элемента.

2. Для реализации метода требуется около $2n^3/3$ арифметических вычислений. Так, при $n = 20$ требуется около 5300 вычислений.

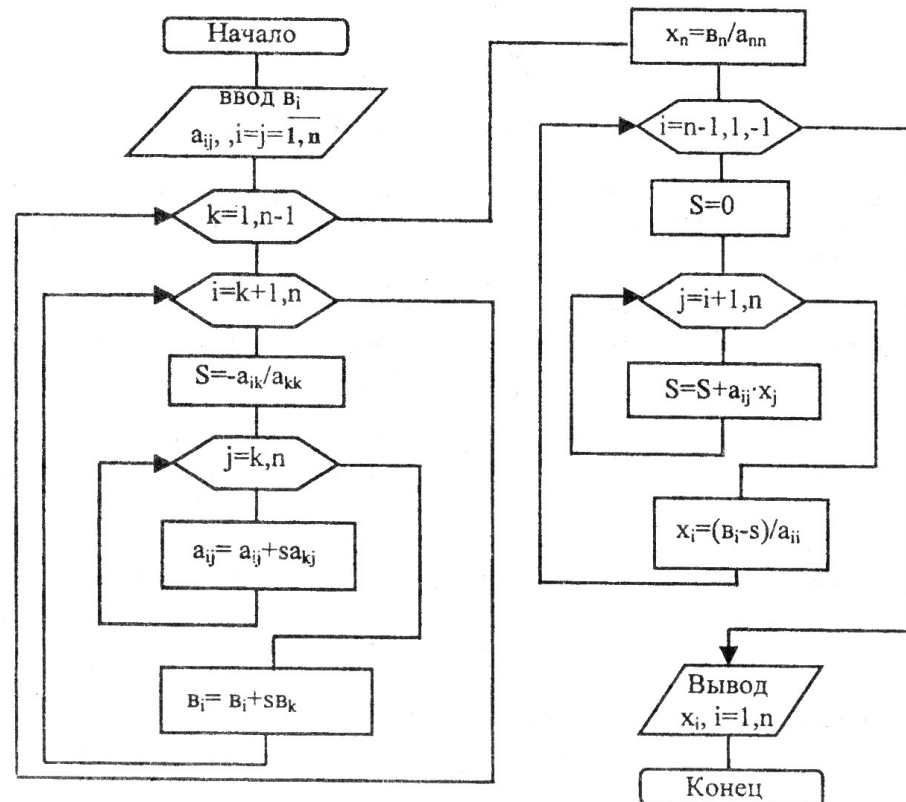


Рисунок 2.2 - Алгоритм метода исключения Гаусса

Пример. Решить систему линейных уравнений

$$\begin{cases} x_1 + 2x_2 + x_3 = 4 \\ 2x_1 + 3x_2 + 3x_3 = 8 \\ 3x_1 + 9x_2 + x_3 = 13 \end{cases}$$

Прямой ход

Шаг I. Исключим x_1 из второго и третьего уравнений. Для этого сначала умножим элементы первой строки на коэффициент $m_{21} = -a_{21}/a_{11} = -2/1 = -2$ и прибавим к соответствующим элементам второй строки:

$$(2-2) \cdot x_1 + (3-4) \cdot x_2 + (3-2) \cdot x_3 = 8-8 \Rightarrow -x_2 + x_3 = 0.$$

Затем умножим элементы первой строки на коэффициент $m_{31} = -a_{31}/a_{11} = -3/1$ и прибавим к элементам третьей строки;

$$(3-3) \cdot x_1 + (9-6) \cdot x_2 + (1-3) \cdot x_3 = 13-12 \Rightarrow 3 \cdot x_2 - 2x_3 = 1.$$

После завершения шага I получим систему, эквивалентную исходной:

$$\begin{cases} x_1 + 2x_2 + x_3 = 4 \\ 0 \cdot x_1 - x_2 + x_3 = 0 \\ 0 + 3x_2 - 2x_3 = 1 \end{cases}$$

Шаг 2. Исключим x_2 из третьего уравнения. Умножим элементы второй строки последней системы на коэффициент $m_{32} = -3/(-1) = 3$ и прибавим к третьей строке:

$$(3-3) \cdot x_2 + (3-2) \cdot x_3 = 1+0 \Rightarrow x_3 = 1.$$

В итоге получим систему с верхней треугольной матрицей:

$$\begin{cases} x_1 + 2x_2 + x_3 = 4 \\ -x_2 + x_3 = 0 \\ x_3 = 1 \end{cases}$$

Обратный ход

Из последнего уравнения следует $x_3^* = 1$. Предпоследнее уравнение имеет вид

$$-x_2 + 1 = 0 \Rightarrow x_2^* = 1.$$

И, наконец, из первого уравнения имеем

$$x_1 + 2 \cdot 1 + 1 = 4 \Rightarrow x_1^* = 1.$$

2.1.4 Метод исключения Гаусса с выбором главного элемента

Если элемент главной диагонали a_{ii} мал ($|a_{ii}| \ll |a_{ij}|$), то при осуществлении исключения Гаусса возникают большие ошибки округления, которые, накопляясь, сильно искажают результат. Чтобы избежать этого, рекомендуется каждый шаг исключения начинать с выбора главного элемента.

Пусть $\det A \neq 0$ и $a_{11}^{(0)} \approx 0$. Выберем в первом столбце максимальный по модулю элемент. Если такого элемента не окажется, то все элементы первого столбца близки к нулю и, следовательно, практически $\det A = 0$, что противоречит начальному предположению. Итак, существует элемент столбца $a_{k1} \neq 0$, который по модулю больше остальных.

Переставим местами 1 и К строки:

$$\vec{a}_{1j}^{(0)} \leftrightarrow \vec{a}_{kj}^{(0)}, \quad j = \overline{1, n}; \quad \vec{v}_1^{(0)} \leftrightarrow \vec{v}_k^{(0)}.$$

Для полученной системы проведем шаг 1 исключения Гаусса, после чего осуществим поиск главного элемента по второму столбцу. Выбираем максимальный по модулю элемент, начиная со 2-ой строки, и переставим строки:

$$\vec{a}_{2j}^{(1)} \leftrightarrow \vec{a}_{kj}^{(1)}, \quad j = \overline{2, n}; \quad \vec{v}_2^{(1)} \leftrightarrow \vec{v}_k^{(1)} \quad \text{и т.д.}$$

Замечания

1. Для плохо обусловленных систем устойчивость метода может оказаться недостаточной.
2. Погрешность округления может быть еще несколько уменьшена выбором главного элемента по всей матрице. Однако точность при этом возрастает незначительно, а расчет заметно усложняется.

2.1.5 Матричная формулировка метода Гаусса

Как известно, если систему $AX=B$ умножить на некоторую ненулевую матрицу C , получим систему $CAx=CB$, эквивалентную исходной. При выполнении первого шага исключения Гаусса первая строка системы последовательно умножалась на коэффициенты $m_{21}, \dots, m_{n1}$ и прибавлялась к соответствующей строке, в результате чего поддиагональные элементы первого столбца обращались в нуль. В матричном виде эта операция выглядит так:

$$M_1 A^{(0)} X = M_1 B^{(0)} \Rightarrow A^{(1)} X = B^{(1)}, \quad \text{где } M_1 = \begin{bmatrix} 1 & 0 & 0 \\ m_{21} & \ddots & \\ m_{n1} & 0 & 1 \end{bmatrix}$$

Второй шаг исключения

$$M_2 M_1 A^{(0)} X = M_2 M_1 B^{(0)} \Rightarrow A^{(2)} X = B^{(2)}, \text{ где } M_2 = \begin{bmatrix} 1 & 0 & & \\ 0 & 1 & & \\ \vdots & \vdots & \ddots & \\ 0 & m_{n2} & 0 & 1 \end{bmatrix}$$

Последний шаг исключения соответствует операции

$$M A^{(0)} X = M B^{(0)} \Rightarrow A^{(n-1)} X = B^{(n-1)}, \quad M = \begin{bmatrix} 1 & 0 & 0 & \\ m_{21} & 1 & & \\ m_{31} & m_{32} & \ddots & \\ \vdots & \vdots & \vdots & \ddots \\ m_{n1} & m_{n2} & & 1 \end{bmatrix},$$

где матрица M является результатом произведения $n-1$ матриц:

$$M = M_{n-1} \dots M_2 M_1.$$

Особенность матрицы M такова, что обратная ей матрица M^{-1} является также нижней треугольной с единичными элементами на главной диагонали, а поддиагональные элементы равны соответствующим элементам матрицы M , взятым с противоположным знаком.

Примем следующие обозначения: $L = M^{-1}$ и $U = A^{(n-1)}$.

Умножим полученную систему с верхней треугольной матрицей

$A^{(n-1)} X = B^{(n-1)}$ или по-другому $MA^{(0)} X = MB^{(0)}$ слева на матрицу L :

$$LMA^{(0)} X = LMB^{(0)} \Rightarrow LUX = B. \quad (2.5)$$

Данное представление матрицы A называется LU разложением. В частности, мы подтвердили следующую теорему.

Теорема. Всякую квадратную матрицу A , у которой отличны от нуля главные миноры, можно представить в виде произведения нижней треугольной матрицы на верхнюю треугольную матрицу.

2.1.6 Вычисление обратной матрицы и значения определителя

Непосредственное вычисление значения определителя с использованием миноров квадратной матрицы требует большого объема вычислений. С другой стороны, значение определителя треугольной матрицы вычисляется просто: он равен произведению ее диагональных элементов.

В процессе исключения поддиагональных элементов матрицы методом Гаусса величина определителя не меняется. Таким образом, значение определителя равно

$$\det A = \pm \prod_{i=1}^n U_{ii},$$

где U - верхняя треугольная матрица исключения Гаусса.

Знак зависит от суммарного числа перестановок строк или столбцов (при использовании метода Гаусса с выбором главного элемента).

Нахождение матрицы Z , обратной матрице A , эквивалентно решению матричного уравнения

$$AZ = E.$$

Данное уравнение можно записать в виде системы n^2 уравнений

$$\sum_k a_{ik} \cdot Z_{kj} = \delta_{ij}, \quad i, j = 1, \bar{n},$$

где δ - символ Кронекера.

Полученная система распадается на n независимых систем уравнений с одной и той же матрицей A , но с различными правыми частями. Тогда достаточно один раз совершить прямой ход метода исключения Гаусса, т.е. получить разложение $A=LU$ и запомнить нижнюю треугольную матрицу L . Верхняя треугольная матрица U совпадает с матрицей $A^{(n-1)}$. Обратный ход осуществляется путем решения систем уравнений

$$Ly^{(j)} = \delta^{(j)}$$

$$A^{(n-1)} Z^{(j)} = y^{(j)}, \quad j = 1, \bar{n},$$

где $y^{(j)}$ – промежуточный вектор; $\delta^{(j)}$ – вектор, у которого j -я компонента равна единице, остальные - нулю.

На рисунке 2.3 приведен фрагмент алгоритма обращения матрицы, которым заменяется обратный ход метода исключения Гаусса (рисунок 2.2).

В алгоритме прямого хода метода Гаусса (рисунок 2.2) после расчета $s = -a_{ik}/a_{kk}$ необходимо вставить блок $L_{ik} = s$. После выполнения прямого хода метода исключения Гаусса полученная верхняя треугольная матрица будет являться матрицей U .

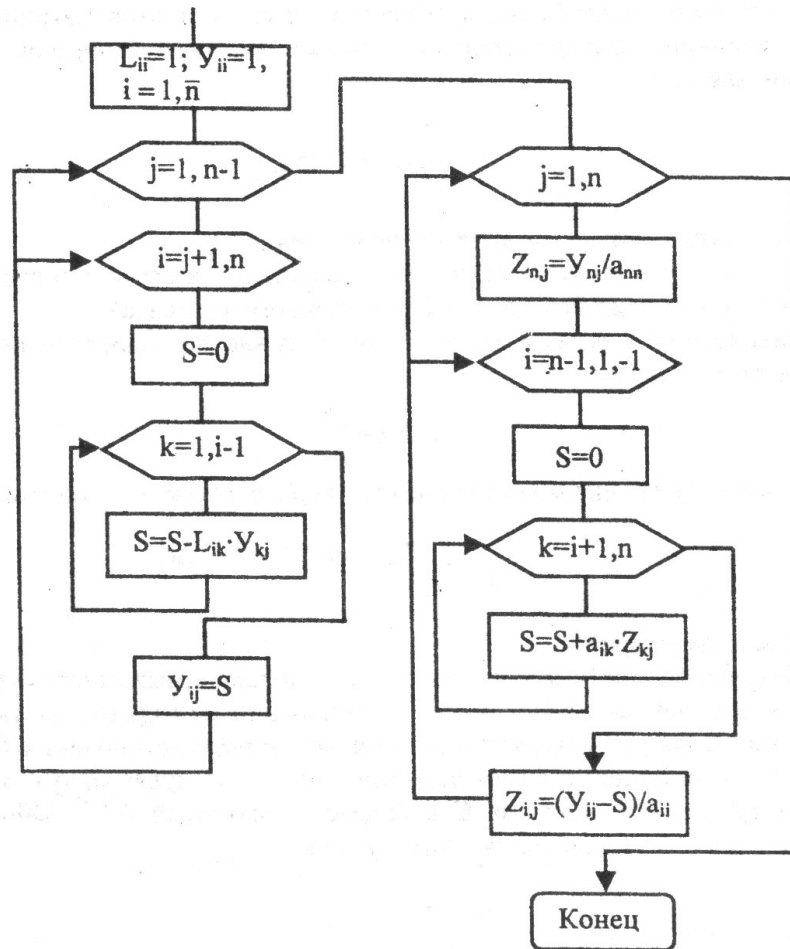


Рисунок 2.3 - Алгоритм обратного хода обращения матрицы

2.1.7 Метод исключения для ленточных матриц. Прогонка

Матрица A называется $2\alpha+1$ диагональной (ленточной), если $a_{ij}=0$ для всех $|i-j|>\alpha$, и $a_{ij}\neq 0$ – в противном случае.

Например, при $\alpha=1$ имеем $A = \begin{bmatrix} & & 0 \\ & \diagdown & \\ & 0 & \diagup \\ 0 & & \end{bmatrix}$.

Такие матрицы образуются при решении краевых задач для дифференциальных уравнений второго порядка.

Для ленточных матриц метод исключения Гаусса может быть усовершенствован путем игнорирования многочисленных нулевых элементов матрицы, что позволит уменьшить объем вычислений и необходимой памяти ЭВМ.

Рассмотрим систему с трехдиагональной матрицей ($\alpha=1$):

$$\begin{aligned} a_{i,i-1}x_{i-1} + a_{ii}x_i + a_{i,i+1}x_{i+1} &= b_i, & i = \overline{1, n} \\ a_{i0} &= 0; & a_{n,n+1} &= 0. \end{aligned} \quad (2.6)$$

Соответствующие матрицы L и U будут иметь вид

$$A = \begin{bmatrix} 1 & & 0 \\ -m_{21} & & \\ 0 & -m_{n,n-1} & 1 \end{bmatrix}; \quad U = \begin{bmatrix} a_{11}^{(0)} & a_{12}^{(0)} & 0 \\ & & a_{n-1,n}^{(n-2)} \\ 0 & & a_{nn}^{(n-1)} \end{bmatrix}.$$

Из матрицы U следует, что в каждом уравнении содержатся только по два неизвестных:

$$a_{ii}^{(i-1)} \cdot x_i + a_{i,i+1}^{(i-1)} \cdot x_{i+1} = \hat{a}_i^{(i-1)}, \quad i = \overline{1, n}$$

Обозначим

$$x_i = \zeta_{i+1} \cdot x_{i+1} + \eta_{i+1}, \quad i = \overline{n, 1}; \quad x_{n+1} = 0. \quad (2.7)$$

Уравнения (2.7) являются *формулами обратного хода*. Подставив их в (2.6), получим

$$x_i = \frac{a_{i,i+1}}{-a_{ii} - a_{i,i-1} \cdot \zeta_i} \cdot x_{i+1} + \frac{a_{i,i-1} \cdot \eta_i - \hat{a}_i}{-a_{ii} - a_{i,i-1} \cdot \zeta_i}. \quad (2.8)$$

Из (2.7) и (2.8) следует

$$\begin{cases} \zeta_{i+1} = -a_{i,i+1} / (a_{ii} + a_{i,i-1} \cdot \zeta_i) \\ \eta_{i+1} = + (a_{i,i-1} \eta_i - \hat{a}_i) / (-a_{ii} - a_{i,i-1} \cdot \zeta_i), i = \overline{1, n} \\ \zeta_1 = 0; \eta_1 = 0 \end{cases} \quad (2.9)$$

Уравнения (2.9) называется *формулами прямого хода*.

Согласно методу по уравнениям (2.9) рассчитываются значения промежуточных переменных ζ_i и η_i с помощью которых затем по уравнениям (2.7) определяются значения x_i^* .

Замечания

1. При диагональном преобладании элементов матрицы A метод абсолютно устойчив.
2. Объем вычислений – $9n$ арифметических действий.

2.1.8 Метод простых итераций

Исходную систему (2.1) приведем к эквивалентному виду $X = \overline{A} X + \overline{B}$ и построим итерационный процесс:

$$X^{(k+1)} = \overline{A} X^{(k)} + \overline{B}, \quad k=0,1,2,\dots \quad (2.10)$$

Утверждение. Если последовательность приближений $X^{(0)}, X^{(1)}$ имеет предел $X^* = \lim_{k \rightarrow \infty} X^{(k)}$, то этот предел является решением (2.1):

$$\exists \lim_{k \rightarrow \infty} X^{(k+1)} = \bar{A} \lim_{k \rightarrow \infty} X^{(k)} + \bar{B} \Leftrightarrow X^* = \bar{A}X^* + \bar{B}.$$

Возможны два вида системы эквивалентных уравнений. Приведем процесс (2.10) для обоих видов.

Вид 1. $(a_{ii} \neq 0)$

$$x_i^{(k+1)} = \sum_{j=1}^n \bar{a}_{ij} \cdot x_j^{(k)} + \bar{b}_i, \quad i = \overline{1, n}, \quad k = 0, 1, 2, \dots,$$

где $\bar{a}_{ij} = -a_{ij}/a_{ii}$; $\bar{b}_i = b_i/a_{ii}$, $j \neq i$, $\bar{a}_{ii} = 0$, $i = \overline{1, n}$.

Вид 2. Пусть $a_{ii} = \alpha_{i1} + \alpha_{i2}$ и $\alpha_{i1} \neq 0$.

$$x_i^{(k+1)} = + \sum_{j=1}^n \bar{a}_{ij} \cdot x_j^{(k)} + \bar{b}_i, \quad i = \overline{1, n}, \quad k = 0, 1, 2, \dots,$$

где $\bar{a}_{ij} = -a_{ij}/\alpha_{i1}$; $\bar{b}_i = b_i/\alpha_{i1}$, $j \neq i$, $\bar{a}_{ii} = -\alpha_{i2}/\alpha_{i1}$, $i = \overline{1, n}$.

Вид системы эквивалентных уравнений следует выбирать в соответствии со следующей теоремой.

Теорема 1. (достаточное условие сходимости). Итерационный процесс (2.10) сходится, если для матрицы $\bar{A}$ выполняется одно из неравенств:

$$a) \sum_{j=1}^n |\bar{a}_{ij}| < 1, \quad i = \overline{1, n}; \quad б) \sum_{i=1}^n |\bar{a}_{ij}| < 1, \quad j = \overline{1, n}.$$

(2.11a)

Следствием теоремы 1 является теорема 2.

Теорема 2. Итерационный процесс (2.10) сходится, если у исходной матрицы A выполняется диагональное преобладание элементов:

$$a) |a_{ii}| > \sum_{j=1}^n |a_{ij}|, \quad i = \overline{1, n}, \quad i \neq j; \quad б) |a_{jj}| > \sum_{i=1}^n |a_{ij}|, \quad j = \overline{1, n}, \quad j \neq i.$$

(2.11б)

Если условия теорем выполняются, то итерационный процесс (2.10) сходится при любом начальном приближении $X^{(0)}$, т.е. метод является самоисправляющимся.

Обычно в качестве начального приближения берут $X^{(0)} = B$ или $X^{(0)} = \vec{0}$, если не известно приближение, близкое к решению $X^{(0)} \approx X^*$.

Критерием завершения итерационного расчета обычно служит неравенство

$$\sum_{i=1}^n |x_i^{(k+1)} - x_i^{(k)}| < \varepsilon.$$

Пример. Привести систему к эквивалентному виду:

$$\begin{cases} 4x_1 - x_2 + x_3 = 4 \\ -x_1 - 2x_2 + 5x_3 = 2 \\ x_1 + 6x_2 + 2x_3 = 9 \end{cases} \xrightarrow[\text{(2.11б)}]{\text{условие}} \begin{cases} 4x_1 - x_2 + x_3 = 4 \\ x_1 + 6x_2 + 2x_3 = 9 \\ -x_1 - 2x_2 + 5x_3 = 2 \end{cases} \xrightarrow{\text{Вид1}} \begin{cases} x_1 = \frac{1}{4}(0 + x_2 - x_3 + 4) \\ x_2 = \frac{1}{6}(-x_1 + 0 - 2x_3 + 9) \\ x_3 = \frac{1}{5}(x_1 + 2x_2 + 0 + 2) \end{cases}$$

Для последней системы выполняется условие (2.11 а).

Алгоритм преобразования системы в соответствии с диагональным преобладанием элементов приведен на рисунок 2.4.

2.1.9 Метод Зейделя

В методе простых итераций на k -ой итерации рассчитываются все элементы вектора X , которые затем используются на $k+1$ -ой итерации в качестве нового приближения. Сходимость метода можно ускорить, если рассчитанные значения неизвестных используются на этой же итерации для вычисления значений последующих неизвестных. Например, при расчете $x_2^{(k+1)}$ использовать не $x_1^{(k)}$, а уже рассчитанное значение $x_1^{(k+1)}$, при расчете $x_3^{(k+1)}$ использовать $x_1^{(k+1)}$, $x_2^{(k+1)}$ и т.д.

$$x_i^{(k+1)} = \sum_{j=1}^{i-1} \tilde{a}_{ij} x_j^{(k+1)} + \sum_{j=i+1}^n \tilde{a}_{ij} x_j^{(k)} + \tilde{a}_i, \quad i = \overline{1, n}, \quad k = 0, 1, 2, \dots \quad (2.12)$$

Здесь достаточное условие сходимости определяется также теоремами 1 и 2, рассмотренными выше.

Пример. Методом Зейделя решить систему

$$\begin{cases} x_1 - 2x_2 = -2 \\ 2x_1 + x_2 = 2 \end{cases} \quad \begin{aligned} x_1^{(k+1)} &= (-2 + 2x_2^{(k)})/1 \\ x_2^{(k+1)} &= (2 - 2x_1^{(k+1)})/1 \end{aligned}$$

k	x_1	x_2
0	0	0
1	-2	6
2	10	-18
3	-33	78

$$\begin{aligned} x_1^{(1)} &= -2 + 2 \cdot 0 = -2 & x_1^{(2)} &= -2 + 2 \cdot 6 = 10 \\ x_2^{(1)} &= 2 - 2(-2) = 6 & x_2^{(2)} &= 2 - 2 \cdot 10 = -18 \end{aligned}$$

и т.д.

Данный итерационный процесс расходится, т.к. не выполняется условие (2.11). Переставим уравнения - обеспечим диагональное преобладание элементов.

$$\begin{cases} 2x_1 + x_2 = 2 \\ x_1 - 2x_2 = -2 \end{cases} \quad \begin{aligned} x_1^{(k+1)} &= (2 - x_2^{(k)}) / 2 \\ x_2^{(k+1)} &= (2 + x_1^{(k+1)}) / 2 \end{aligned}$$

k	x_1	x_2
0	0	0
1	1	1,5
2	0,25	1,125
3	0,44	1,22

$$x_1^{(1)} = (2 - 0) / 2 = 1 \quad x_2^{(2)} = (2 - 1,5) / 2 = 0,25$$

$$x_2^{(1)} = (2 + 1) / 2 = 1,5 \quad x_2^{(2)} = (2 + 0,25) / 2 = 1,125$$

и т.д.

Итерационный процесс сходится.

Пределом (решением) является $x_1^* = 0,40$; $x_2^* = 1,20$.

Алгоритм метода приведен на рисунке 2.5.

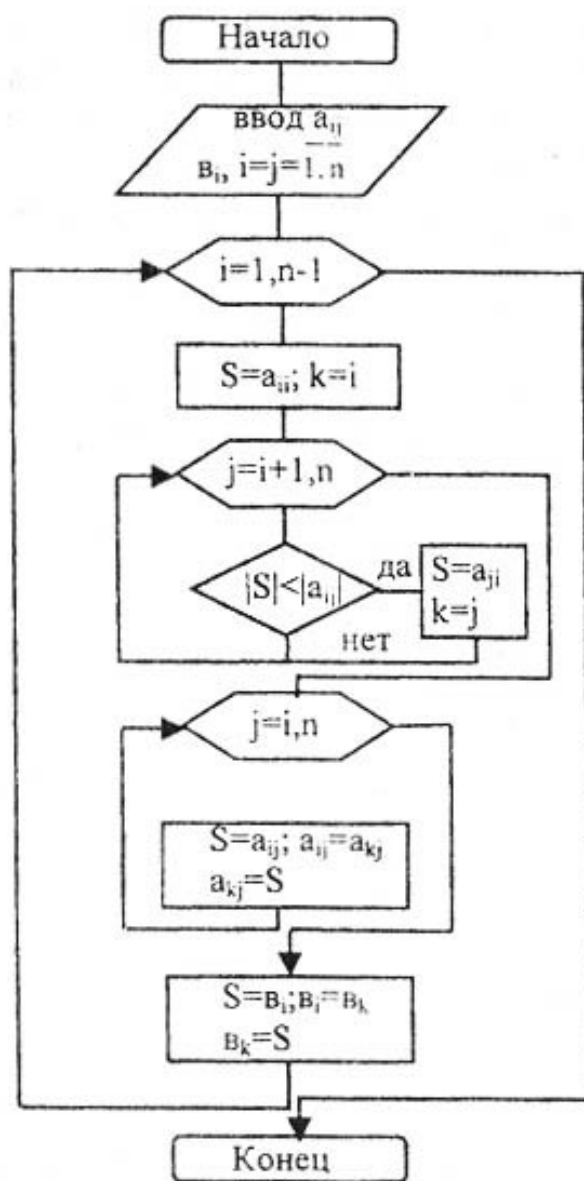


Рисунок 2.4 - Алгоритм приведения к диагональному преобладанию

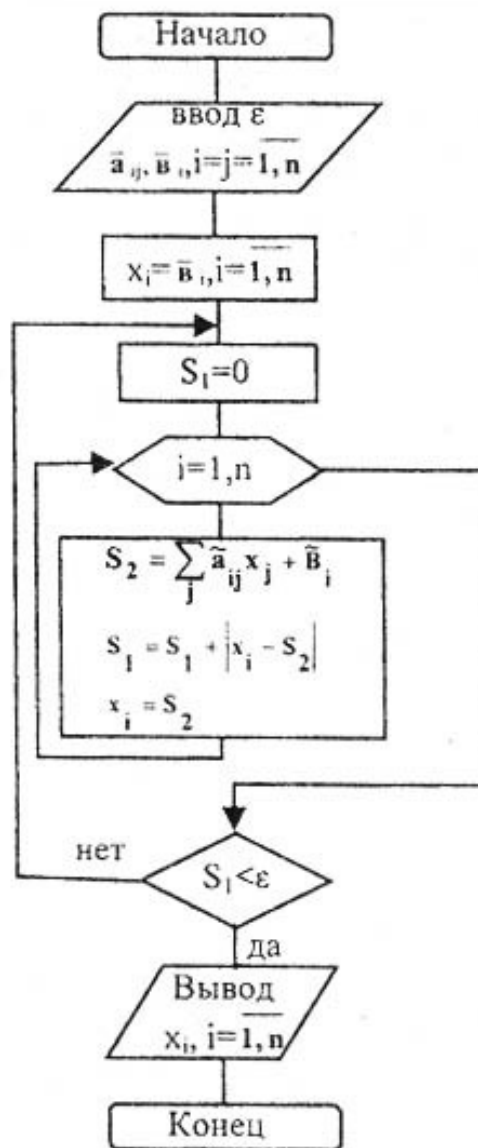


Рисунок 2.5 - Алгоритм метода Зейделя

Вопросы и упражнения для самоконтроля.

- 1) Какая система уравнений называется определенной?
- 2) К какой категории методов вычислительной математики относится метод Гаусса?
- 3) В чем заключается прямой и обратный ход метода Гаусса?
- 4) В каком случае нельзя проводить исключение Гаусса?
- 5) Почему метод Гаусса с выбором главного элемента дает более точный результат, по сравнению с простым методом Гаусса?
- 6) В чем сущность метода LU разложения?
- 7) Для каких систем линейных уравнений применяется метод прогонки?
- 8) В чем заключаются этапы прямой и обратной прогонки?
- 9) Как преобразуется система линейных уравнений к итерационному виду?

- 10) При каких условиях итерационный процесс сходится?
- 11) Когда завершается итерационный процесс?
- 12) В чем отличие метода Зейделя от метода простых итераций?
- 13) Решить систему линейных уравнений методом исключения Гаусса:

$$\text{а) } \begin{cases} 2,34\delta_1 - 4,21\delta_2 - 11,61\delta_3 = 14,41 \\ 8,04\delta_1 + 5,22\delta_2 + 0,27\delta_3 = -6,44 \\ 3,92\delta_1 - 7,99\delta_2 + 8,37\delta_3 = 55,56 \end{cases} \quad \text{б) } \begin{cases} 2\delta_1 - \delta_3 = 1 \\ \delta_1 - 3\delta_2 + \delta_3 = 2 \\ \delta_1 + \delta_2 + 3\delta_3 = 4 \end{cases}$$

$$\text{в) } \begin{cases} 300x_1 - 12x_2 + 11x_3 = -3504 \\ 1,5x_1 + 600x_2 + 8,5x_3 = 1884 \\ -7x_1 + 7x_2 + 900x_3 = 1091 \end{cases} \quad \text{г) } \begin{cases} 3\delta_1 + \delta_2 - \delta_3 = 6 \\ 2\delta_1 + 4\delta_2 + \delta_3 = 4 \\ \delta_1 - \delta_2 + 3\delta_3 = 4 \end{cases}$$

$$\text{д) } \begin{cases} \delta_1 + 10\delta_2 - 5\delta_3 = 1 \\ 2\delta_1 + 4\delta_2 - \delta_3 = 2 \\ \delta_1 + 3\delta_2 + \delta_3 = 3 \end{cases} \quad \text{е) } \begin{cases} -4,21\delta_2 - 11,61\delta_3 = 14,41 \\ 8,04\delta_1 + 5,22\delta_2 + 0,27\delta_3 = -6,44 \\ 3,92\delta_1 - 7,99\delta_2 + 8,37\delta_3 = 55,56 \end{cases}$$

$$\text{ж) } \begin{cases} 200x_1 - 13x_2 + 12x_3 = -2470 \\ x_1 + 400x_2 + 9x_3 = 904 \\ -8x_1 + 8x_2 + 600x_3 = 7920 \end{cases} \quad \text{з) } \begin{cases} 2\delta_1 - \delta_3 = 1 \\ \delta_1 - 4\delta_2 + 2\delta_3 = -5 \\ \delta_1 + \delta_2 + \delta_3 = 6 \end{cases}$$

$$\text{и) } \begin{cases} 611\delta_1 + 196\delta_2 - 192\delta_3 = 407 \\ 196\delta_1 + 899\delta_2 + 113\delta_3 = -192 \\ -192\delta_1 + 113\delta_2 + 899\delta_3 = 196 \end{cases} \quad \text{к) } \begin{cases} 1,00\delta_1 + 10,00\delta_2 - 5,00\delta_3 = 1 \\ 1,01\delta_1 + 9,99\delta_2 - 5,01\delta_3 = 3 \\ 0,99\delta_1 + 1,01\delta_2 - 4,99\delta_3 = 3 \end{cases}$$

$$\begin{array}{ll}
\text{л)} \left\{ \begin{array}{l} 4\delta_1 + \delta_2 + \delta_3 = 9 \\ \delta_1 + 5\delta_2 + \delta_3 = 14 \\ 2\delta_1 + 2\delta_2 + 5\delta_3 = 21 \end{array} \right. ; & \text{м)} \left\{ \begin{array}{l} 600x_1 - 9x_2 + 8x_3 = -5382 \\ 3x_1 + 1200x_2 + 7x_3 = 72360 \\ -4x_1 + 4x_2 + 1800x_3 = 16260 \end{array} \right. ; \\
\text{н)} \left\{ \begin{array}{l} \delta_1 + \delta_2 + \delta_3 = 1 \\ 8\delta_1 + 4\delta_2 + 2\delta_3 = 1 \\ 27\delta_1 + 9\delta_2 + 3\delta_3 = 1 \end{array} \right. ; & \text{о)} \left\{ \begin{array}{l} 3\delta_1 + \delta_2 - \delta_3 = -2 \\ 3\delta_1 + 5\delta_2 - \delta_3 = 8 \\ 2\delta_1 + 3\delta_3 = 1 \end{array} \right. ; \\
\text{п)} \left\{ \begin{array}{l} 3x_1 + x_2 - x_3 = -1 \\ -2x_1 + 4x_2 + x_3 = 5 \\ x_1 + x_2 + 3x_3 = -3 \end{array} \right. ; & \text{р)} \left\{ \begin{array}{l} 900x_1 - 6x_2 + 5x_3 = -5424 \\ 4.5x_1 + 1800x_2 + 5.5x_3 = 16210 \\ -x_1 + x_2 + 2700x_3 = 16220 \end{array} \right. ; \\
\text{с)} \left\{ \begin{array}{l} 3\delta_1 + \delta_2 + 0.5\delta_3 = 7.5 \\ \delta_1 + 3\delta_2 + \delta_3 = 6 \\ \delta_1 + \delta_2 + 4\delta_3 = 7 \end{array} \right. ; & \text{т)} \left\{ \begin{array}{l} 1900x_1 + 4x_2 - 5x_3 = 76960 \\ 9.5x_1 + 3800x_2 + 0.5x_3 = 72240 \\ 9x_1 - 9x_2 + 5700x_3 = -22940 \end{array} \right. .
\end{array}$$

14) Решить систему методом прогонки:

$$\begin{aligned}
4x_1 + 2x_2 &= \frac{3}{2} + \frac{1}{5}N, \\
2x_1 + 6x_2 + 2x_3 &= \frac{13}{6} + \frac{1}{4}N, \\
-4x_2 + 6x_3 + 2x_4 &= 1, \\
2x_3 - 4x_4 &= \frac{2}{3} + \frac{1}{3}N,
\end{aligned}$$

где N – параметр, соответствующий номеру по списку группы.

15) Методами простых итераций и Зейделя решить систему линейных алгебраических уравнений $AX=B$, где

$$\hat{A} = \begin{pmatrix} 10 & 2 & 2 \\ 4 & 12 & -2 \\ 4 & 6 & 20 \end{pmatrix}, \quad \hat{A} = \begin{pmatrix} 22 + 0.4 \cdot N \\ 26 + 0.5 \cdot N \\ 36 - 0.3 \cdot N \end{pmatrix},$$

где N – параметр, соответствующий номеру по списку группы, с точностью $\varepsilon = 0.001$.

2.2 Решение систем нелинейных уравнений

Рассмотрим систему нелинейных уравнений

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0 \\ f_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ f_n(x_1, x_2, \dots, x_n) = 0 \end{cases}$$

или в операторном виде

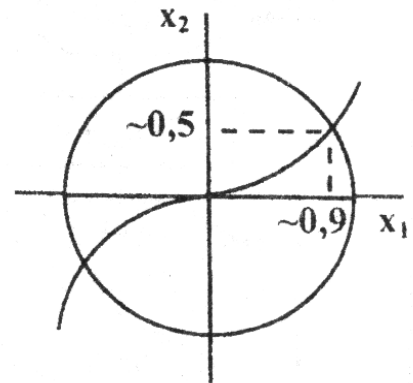
$$F(x) = 0. \quad (2.13)$$

Решением системы являются такие $x_1^*, \dots, x_n^*$ которые удовлетворяют (2.13). При их поиске наиболее часто используются итерационные методы. Начальные приближения, необходимые при применении итерационного метода, определяются априори или графически.

Пример. Определить графически начальное приближение для решения системы

$$\begin{cases} x_1^2 + x_2^2 - 1 = 0 \\ x_1^3 - 1,5x_2 = 0 \end{cases}$$

$$X^{(0)} = \{0,9; 0,5\}.$$



2.2.1 Метод простых итераций

Приведем исходную систему (2.13) к виду

$$\begin{cases} x_1 = \varphi_1(x_1, \dots, x_n) \\ \dots \\ x_n = \varphi_n(x_1, \dots, x_n) \end{cases}$$

или в операторной форме $X = \varphi(X)$,

(2.14)

и построим итерационный процесс

$$x_1^{(k+1)} = \varphi_1(x_1^{(k)}, \dots, x_n^{(k)})$$

...

$$x_n^{(k+1)} = \varphi_n(x_1^{(k)}, \dots, x_n^{(k)})$$

(2.15)

Если итерации сходятся, то они сходятся к решению X^* :

$$X^* = \lim_{k \rightarrow \infty} \varphi(X^{(k)}) = \varphi(\lim_{k \rightarrow \infty} X^{(k)}).$$

Условием сходимости (2.15) являются неравенства

$$\sum_{i=1}^n \left| \frac{\partial \varphi_i}{\partial x_1} \right| < 1, \dots, \sum_{i=1}^n \left| \frac{\partial \varphi_i}{\partial x_n} \right| < 1. \quad (2.16)$$

Из исходной системы (2.13) можно получить несколько систем вида (2.14). Из этих систем нужно выбрать одну, удовлетворяющую в точке $X^{(0)}$ условию (2.16).

Пример. Решить систему

$$\begin{cases} x_1 + 3 \lg x_1 - x_2^2 = 0 \\ 2x_1^2 - x_1 x_2 - 5x_1 + 1 = 0 \end{cases}$$

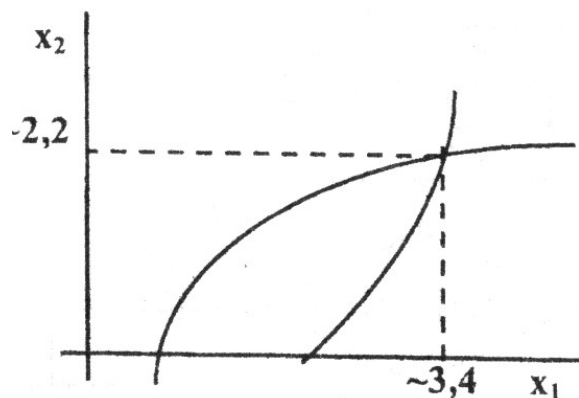


Рисунок 2.6 - Определение начального приближения

Приведем эту систему к виду (2.14) следующим образом:

$$\begin{aligned} x_1 &= x_2^2 - 3 \lg x_1 \\ x_2 &= 2x_1 - 5 + 1/x_1 \end{aligned} \quad \begin{aligned} \frac{\partial \varphi_1}{\partial x_1} &= -\frac{3 \lg e}{x_1}; & \frac{\partial \varphi_1}{\partial x_2} &= 2x_2 \\ \frac{\partial \varphi_2}{\partial x_1} &= 2 - \frac{1}{x_1^2}; & \frac{\partial \varphi_2}{\partial x_2} &= 0 \end{aligned}$$

Примем (рисунок 2.6) $X^{(0)} = \{3,4; 2,2\}$.

В окрестностях точки $X^{(0)}$ имеем

$$\left| -3 \lg e / 3,4 + \left| 2 - 1/11,6 \right| \right| > 1; \quad \left| 2 \cdot 2,2 \right| + 0 > 1.$$

Условие (2.16) не выполняется. Приведем систему к виду (2.14) иначе:

$$\begin{aligned} x_1 &= \sqrt{(x_1(5 + x_2) - 1)} \\ x_2 &= \sqrt{x_1 + 3 \lg x_1}. \end{aligned}$$

Проверкой можно убедиться, что в прямоугольнике $3 < x_1 < 4$, $2 < x_2 < 2,5$ для данного вида системы справедливо

$$\left| \frac{\partial \varphi_1}{\partial x_1} \right| < 0,6; \quad \left| \frac{\partial \varphi_1}{\partial x_2} \right| < 0,32; \quad \left| \frac{\partial \varphi_2}{\partial x_1} \right| < 0,34; \quad \left| \frac{\partial \varphi_2}{\partial x_2} \right| = 0,$$

в связи с чем данный вид системы обеспечит сходимость решения.

Ответ: $X^* = \{3,483; 2,260\}$.

Алгоритм метода приведен на рисунок 2.7.

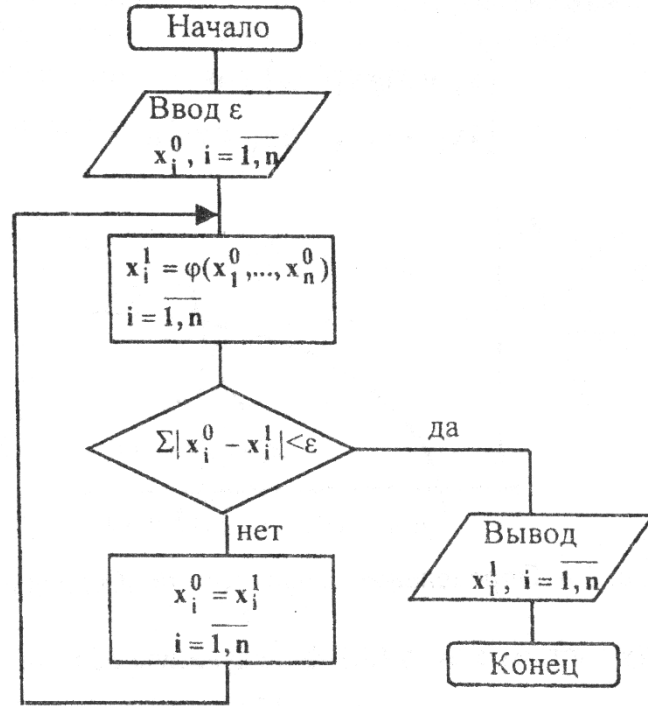


Рисунок 2.7 - Алгоритм метода простых итераций

2.2.2 Метод Зейделя

Этот метод предназначен для решения систем, записанных в форме (2.14). Метод Зейделя является модификацией метода простых итераций, в котором после задания начального приближения $X^{(0)}$ вместо параллельного итерирования производится последовательное итерирование, причем на каждой итерации в каждое последующее уравнение подставляются значения неизвестных, полученных из предыдущих уравнений.

$$\begin{aligned} x_1^{(k+1)} &= \varphi_1(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}), \\ x_2^{(k+1)} &= \varphi_2(x_1^{(k+1)}, x_2^{(k)}, \dots, x_n^{(k)}), \\ &\dots \\ x_n^{(k+1)} &= \varphi_n(x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{n-1}^{(k+1)}, x_n^{(k)}). \end{aligned} \quad (2.17)$$

Критерием завершения итерационного расчета обычно служит неравенство

$$\sum_{i=1}^n |x_i^{(k+1)} - x_i^{(k)}| < \varepsilon.$$

Пример. Найти корни системы уравнений, расположенные в первом квадранте, методом Зейделя с точностью $\varepsilon = 0,001$.

$$f_1(x_1, x_2) = 2x_1^2 - x_1x_2 - 5x_1 + 1 = 0,$$

$$f_2(x_1, x_2) = x_1 + 3\lg x_1 - x_2^2 = 0.$$

Преобразуем систему к виду (2.14) так, чтобы выполнялось условие сходимости:

$$x_1 = \sqrt{\frac{x_1(x_2 + 5) - 1}{2}} = \varphi_1(x),$$

$$x_2 = \sqrt{x_1 + 3\lg x_1} = \varphi_2(x).$$

Зададим начальное приближение $X^{(0)} = (3,5; 2,2)$.

Выполним расчеты по формуле (2.17):

$$\delta_1^{(k+1)} = \sqrt{\frac{\delta_1^{(k)}(x_2^{(k)} + 5) - 1}{2}}, k = 0, 1, \dots$$

$$x_2^{(k+1)} = \sqrt{x_1^{(k+1)} + 3\lg x_1^{(k+1)}}.$$

Результаты расчета представлены в таблице:

k	0	1	2	3	4	5
$x_1^{(k)}$	3,5	3,4785	3,4821	3,48425	3,485537	3,486305
$x_2^{(k)}$	2,2	2,588	2,26	2,260658	2,261049	2,26128
Δ		0,0588	0,0036	0,00215	0,00128	0,0007

2.2.3 Метод Ньютона-Рафсона

Недостаток метода простой итерации заключается в необходимости прибегать к искусственным приемам при приведении системы к виду, пригодному для итераций. При решении нелинейных уравнений часто используют метод Ньютона.

Пусть найдено некоторое приближение $X^{(k)}$. Разложим уравнения $F(X)$ системы (2.13) в ряды и ограничимся первыми дифференциалами:

$$F(X) = F(X^{(k)}) + F'(X^{(k)}) \cdot \Delta X = 0.$$

Здесь $\Delta X = X^{(n+1)} - X^{(k)}$; $F'(X)$ - матрица Якоби.

Имеем

$$X^{(k+1)} = X^{(k)} - [F'(X^{(k)})]^{-1} \cdot F(X^{(k)}), \quad k = 0, 1, \dots \quad (2.18)$$

Итерационный процесс (2.18) называется *методом Ньютона*. В этом методе на каждой итерации приходится обращать матрицу Якоби, что само по себе является трудоемкой задачей. Иногда, с целью упрощения, один раз рассчитывают Якобиан и используют его на всех остальных итерациях. Од-

нако скорость сходимости резко уменьшается, т.к. матрица производных на начальных итерациях может существенно отличаться от окончательной.

Метод можно упростить, особенно при малых n , если вместо решения задачи обращения матрицы решать систему линейных уравнений относительно приращений Δx_i :

$$F'(X^{(k+1)}) \cdot \Delta X^{(k)} = -F(X^{(k)}).$$

Тогда новое приближение будет определяться так:

$$X^{(k+1)} = X^{(k)} + \Delta X^{(k)}, \quad k=0,1,2,\dots \quad (2.19)$$

Процесс (2.19) называется методом *Ньютона-Рафсона*, алгоритм которого представлен на рисунке 2.8.

Если $\det [F'(X)] \neq 0$, то в окрестностях корня метод сходится квадратично.

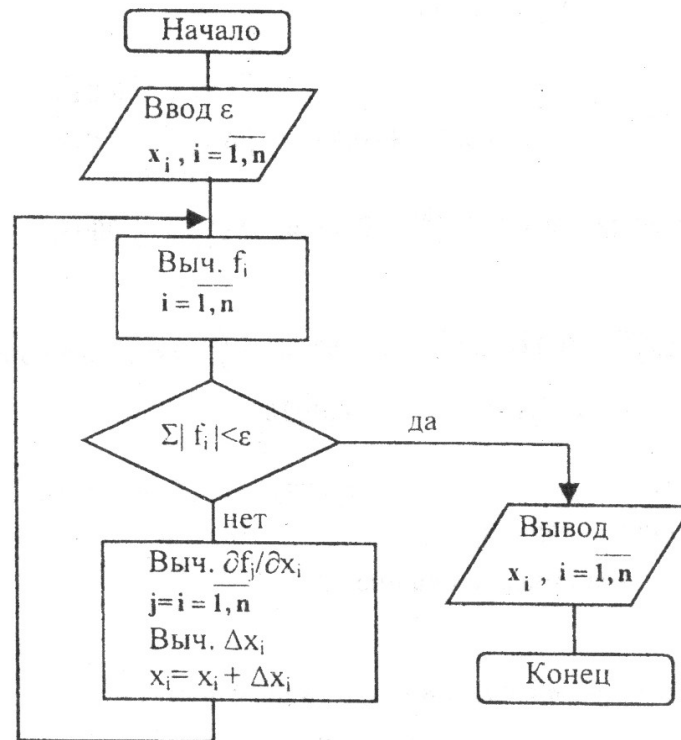


Рисунок 2.8 - Алгоритм метода Ньютона-Рафсона

Пример. Решить систему

$$\begin{cases} 3x_1 + 6x_2 + 2 = 0 \\ x_1^3 + x_2^3 - 1 = 0 \end{cases}$$

с точностью $\varepsilon=0,001$.

В качестве начального приближения примем $X^{(0)}=\{0,83; 0,75\}$.

Тогда

$$\begin{aligned}\frac{\partial f_1}{\partial x_1} &= 3; & \frac{\partial f_1}{\partial x_2} &= 6; \\ \frac{\partial f_2}{\partial x_1} &= 3x_1^2; & \frac{\partial f_2}{\partial x_2} &= 3x_2^2;\end{aligned}$$

$$f_1(x_1^{(0)}, x_2^{(0)}) = -0,01;$$

$$f_2(x_1^{(0)}, x_2^{(0)}) = -0,01;$$

$$|f_1| + |f_2| > \varepsilon.$$

Имеем систему

$$\begin{bmatrix} 3 & 6 \\ 3 \cdot 0,83^2 & 3 \cdot 0,75^2 \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} = \begin{bmatrix} 0,01 \\ 0,01 \end{bmatrix}.$$

Решение этой системы: $\Delta x_1^{(0)} = 0,0044$; $\Delta x_2^{(0)} = 0,0005$.

Следовательно,

$$x_1^{(1)} = x_1^{(0)} + \Delta x_1^{(0)} = 0,83 + 0,0044 = 0,8344; \quad x_2^{(1)} = 0,75 + 0,0005 = 0,7505;$$

$$f_1(x_1^{(1)}, x_2^{(1)}) = 0,002; \quad f_2(x_1^{(1)}, x_2^{(1)}) = 0,0036.$$

Заданная точность еще не достигнута ($|f_1| + |f_2| > \varepsilon$). Необходимо вернуться к уточнению корня и т.д.

2.2.4 Метод минимизации квадратичной функции

Рассмотрим функцию $\varphi(X) = \sum_{i=1}^n (f_i(X))^2$. Эта функция всегда положительна и обращается в нуль в том и только том случае, если $F(X)=0$, т.е. если $X=X^*$. Таким образом, решение системы (2.13) будет одновременно и точкой минимума функции $\varphi(X)$.

Иногда бывает проще искать точку минимума функции $\varphi(X)$, чем решать систему нелинейных уравнений. Методы поиска точки минимума функции рассмотрены в следующей главе. Здесь следует отметить лишь, что $\varphi(X)$ может иметь не нулевые локальные точки минимума, которые не являются решением системы (2.13).

Вопросы и упражнения для самоконтроля.

- 1) Почему задача решения системы нелинейных уравнений сложнее задачи решения системы линейных уравнений?
- 2) В чем заключается метод простых итераций для решения системы нелинейных уравнений?
- 3) Чем отличается метод Зейделя от метода простых итераций?

4) В чем метод Ньютона-Рафсона для решения систем нелинейных уравнений превосходит метод простых итераций?

5) Методами простых итераций, Зейделя, Ньютона-Рафсона решить системы с точностью $\varepsilon = 0,001$:

$$\text{a)} \begin{cases} \delta^3 - \delta^2 - 1 = 0 \\ \delta\delta^3 - \delta - 4 = 0 \end{cases}; \quad \text{б)} \begin{cases} 2\delta^3 - \delta^2 - 1 = 0 \\ \delta\delta^3 - \delta - 4 = 0 \end{cases}; \quad \text{в)} \begin{cases} \sin(x+y) - 1.4x = 0 \\ x^2 + y^2 = 1 \end{cases};$$

$$\text{г)} \begin{cases} \delta + 3\lg x - y^2 = 0 \\ 2x^2 - xy - 5x + 1 = 0 \end{cases}; \quad \text{д)} \begin{cases} \operatorname{tg}(xy + 0.1) = x^2 \\ x^2 + y^2 = 1 \end{cases}; \quad \text{е)} \begin{cases} \sin(x+1) - y = 1 \\ 2x + \cos y = 2 \end{cases};$$

$$\text{ж)} \begin{cases} \delta^3 - \delta^2 - 1 = 0 \\ \delta\delta^3 - \delta - 4 = 0 \end{cases}; \quad \text{з)} \begin{cases} \delta - \sin x = 0 \\ x^2 + y^2 = 1 (x > 0) \end{cases}; \quad \text{и)} \begin{cases} x \cos x - y = 0 \\ x^2 + y^2 = 0 (x > 0) \end{cases};$$

$$\text{к)} \begin{cases} \operatorname{tg}(xy) = x^2 \\ 0.8x^2 + 2y^2 = 1 \end{cases}; \quad \text{л)} \begin{cases} x - 2y^2 + 1 = 0 \\ -x^2 + 2\delta - 1 = 0 \end{cases}; \quad \text{м)} \begin{cases} e^x + 2y = 4 \\ x^2 + 3y^2 = 0 (x > 0) \end{cases};$$

$$\text{н)} \begin{cases} \operatorname{tg}(xy + 0.2) = x^2 \\ x^2 + 2y^2 = 1 \end{cases}; \quad \text{о)} \begin{cases} x - \sin y - \cos y = 0.8 \\ y - 0.02 \sin x^2 - 0.4x = 0 \end{cases}; \quad \text{п)} \begin{cases} \operatorname{tg}(xy + 0.4) = x^2 \\ 0.8x^2 + 2y^2 = 1 \end{cases};$$

$$\text{р)} \begin{cases} \cos(x + 0.5) - y = 2 \\ \sin y - 2x = 1 \end{cases}; \quad \text{с)} \begin{cases} x - 2 \sin x + y = 1 \\ x^2 + y^2 = 1 (x > 0) \end{cases}; \quad \text{т)} \begin{cases} y(xy + 0.1) = x^2 \\ 0.9x^2 + 2y^2 = 1 \end{cases}.$$

3 Минимизация функций

3.1 Одномерная минимизация

Под *минимизацией функции* понимают поиск точек минимума функции в заданной области изменения аргумента:

$$\Phi = \min_{x \in [a, b]} \Phi(x) \quad (3.1)$$

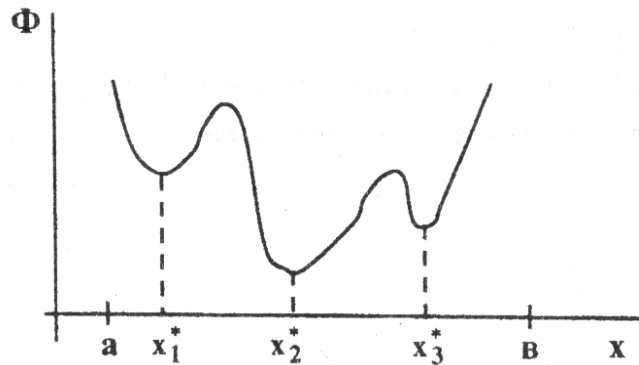


Рисунок 3.1 - Точки минимума функции

На рисунке 3.1 отмечены точками x_1^*, x_2^*, x_3^* - точки локальных минимумов, из них x_2^* - точка глобального минимума.

При поиске точки глобального минимума обычно определяют все точки локальных минимумов и из них выделяют точку, соответствующую наименьшему значению функции.

Замечание. Поиск точек максимума функции можно свести к задаче поиска точек минимума, обратив знак функции на обратный:

$$\Phi = \min (-\Phi(x)) = \max (\Phi(x)).$$

В точках экстремума выполняется равенство $d\Phi/dx=0$. Можно решать это уравнение относительно неизвестных x . Иногда так и поступают, но для сложных функций данный путь не годится. Кроме того, при таком подходе будут определяться как точки минимума, так и точки максимума функции. Мы рассмотрим другой подход - приближенное, численное вычисление точки экстремума.

3.1.1 Метод сканирования

Разделим $[a_0, b_0]$ на m малых интервалов длиной $h_0=(b_0-a_0)/m$ точками x_j , то есть построим равномерную сетку $x_j=a_0+h_0 \cdot j$, $j = \overline{0, m}$ (рисунок 3.2). В точках x_j рассчитываем значения Φ и выберем наименьшее значение:

$$\Phi^{(1)} = \min_{x_j \in [a, b]} \Phi(x_j) \Rightarrow x^{*(1)}$$

Сузим отрезок $[a_0, b_0]$ до отрезка $[a_1, b_1]$:

$$a_1 = x^{*(1)} - p \cdot h_0; \quad b_1 = x^{*(1)} + p \cdot h_0,$$

где $2p < m$; $p=1$, или $p=2$, или $p=3, \dots$

Аналогичным образом разделим отрезок $[a_1, b_1]$ на m более мелких интервалов длиной $h_1 = (b_1 - a_1)/m$ и построим сетку $x_j = a_1 + h_1 \cdot j$, $j = \overline{0, m}$.

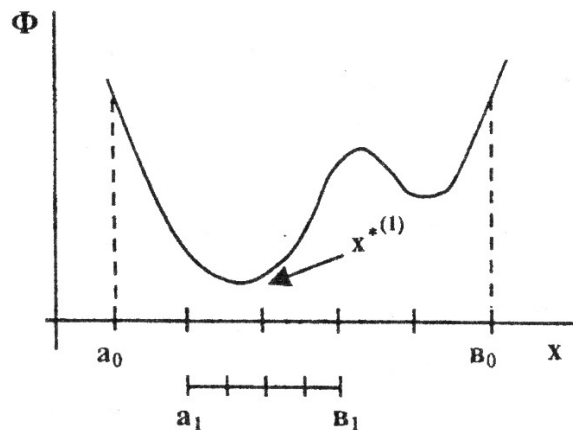


Рисунок 3.2 - Иллюстрация метода сканирования

Вновь рассчитаем значения функции в узлах сетки и выберем из них наименьшее значение.

$$\Phi^{(2)} = \min_{x_j \in [a_1, b_1]} \Phi(x_j) \Rightarrow x^{*(2)}$$

Сузим отрезок $[a_1, b_1]$ до $[a_2, b_2]$ и т.д.

Критерий останова итерационного процесса

$$|\Phi^{(k-1)} - \Phi^{(k)}| < \varepsilon \quad \text{или} \quad h_k < \varepsilon. \quad (3.2)$$

Алгоритм метода приведен на рисунке 3.3.

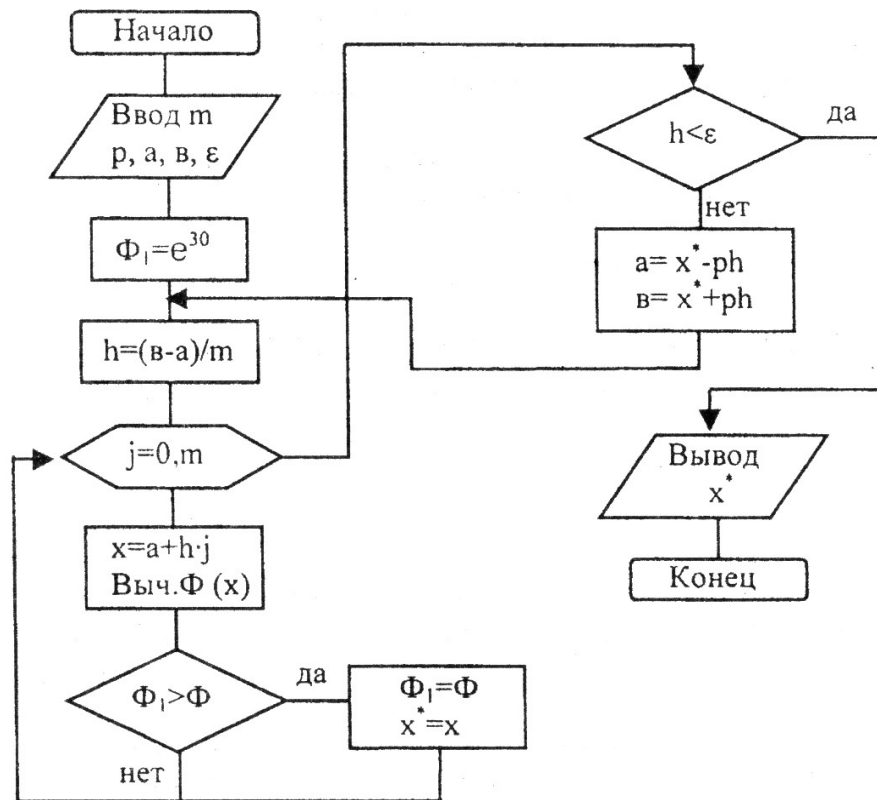


Рисунок 3.3 - Алгоритм метода сканирования

Достоинство метода - при достаточно «густых» точках можно определить точку глобального минимума любой функции, в том числе и разрывной.

Недостаток метода - требуется большой объем вычислений, метод медленно сходится.

Замечание. Для унимодальной функции, имеющей на данном отрезке только одну точку минимума, следует принять $p=1$.

3.1.2 Метод золотого сечения

В начале рассмотрим геометрическую иллюстрацию метода.

Пусть $x^* \in [a_0, b_0]$. Разделим отрезок на три неравных отрезка точками x_1 и x_2 . Рассчитаем в этих точках значения функции и сопоставим их друг с другом:

$$\Phi^{(1)} = \min(\Phi(x_1), \Phi(x_2)).$$

Удалив один из трех отрезков, наиболее удаленных от минимального значения функции, сузим отрезок до $[a_1, b_1]$. На рисунке 3.4 значение функции в точке x_1 меньше, чем в точке x_2 . Следовательно, отрезок $[a_1, b_1]$ образуется из $[a_0, b_0]$ удалением $[x_2, b_0]$.

В свою очередь, разделим $[a_1, b_1]$ на три неравных отрезка. Одна точка уже есть - $\bar{x}_2 \in [a_1, b_1]$ - оставшаяся после сужения точка x_1 или x_2 . На рисунке ей соответствует точка x_1 (точка x_2 в результате сужения исходного отрезка превратилась в точку b_1). Остается нанести вторую точку - точку x_3 . Затем рассчитаем значение функции в точке x_3 (в точке $\bar{x}_2$ мы уже рассчитали), вновь выберем наименьшее значение

$$\Phi^{(2)} = \min(\Phi(\bar{x}_2), \Phi(x_3))$$

и сузим отрезок до $[a_2, b_2]$ и т.д.

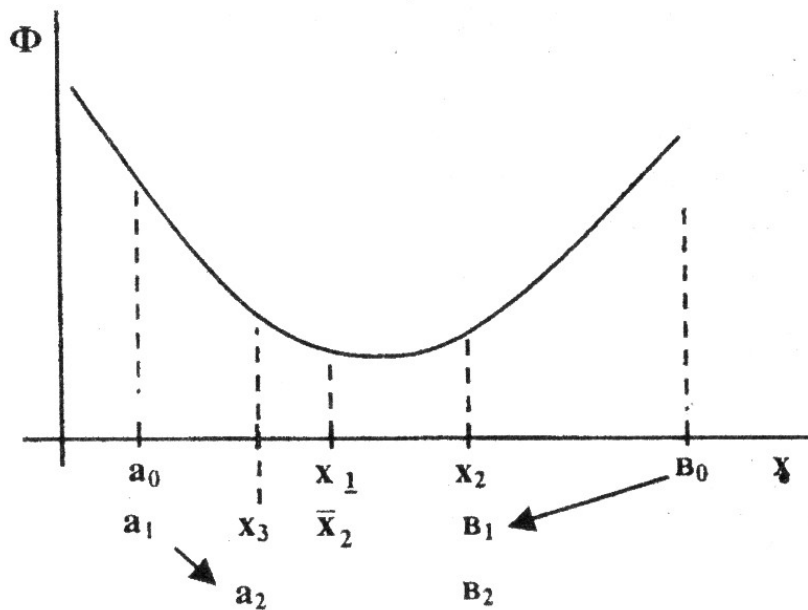


Рисунок 3.4 - Иллюстрация метода золотого сечения

В каком соотношении делить отрезки?

Золотым сечением называется деление отрезка на две части так, чтобы отношение большей части ко всему отрезку равнялось отношению меньшей части к большей части:

$$\frac{b_0 - x_1}{b_0 - a_0} = \frac{x_1 - a_0}{b_0 - x_1} = C \Rightarrow x_1 = b_0 - C(b_0 - a_0);$$

$$\frac{x_2 - a_0}{b_0 - a_0} = \frac{b_0 - x_2}{x_2 - a_0} = C \Rightarrow x_2 = a_0 + C(b_0 - a_0).$$

Решая эти уравнения, можно получить

$$\tilde{N} = (\sqrt{5} - 1)/2 \approx 0,618. \quad (3.3)$$

Такой выбор x_2 замечателен еще и тем, что отрезок $[a_0, x_2]$ делится точкой x_1 по золотому сечению. Аналогично и отрезок $[x_1, b_0]$ делится точкой x_2 также по золотому сечению. Остается лишь точку x_3 выбрать так, чтобы осуществлялось такое же золотое сечение.

Имеем

Уведе-
декс на едини-
обозначим $\bar{x}_2$, где под $\bar{x}_2$ понимается оставшаяся после сужения $[a_0, b_0]$ внутренняя точка x_1 или x_2 : $x_3 = a_1 + b_1 - \bar{x}_2$. На следующем шаге имеем $x_4 = a_2 + b_2 - \bar{x}_3$, где $\bar{x}_3$ – внутренняя точка $\bar{x}_2$ или $\bar{x}_3$, и т.д.

Таким образом, получим рекуррентную формулу

$$x_{k+1} = a_{k-1} + \hat{a}_{k-1} - \bar{x}_k, \quad k = 2, 3, 4, \dots \quad (3.4)$$

Причем на первом шаге

$$x_1 = b_0 - C(b_0 - a_0);$$

$$x_2 = a_0 + C(b_0 - a_0).$$

Алгоритм метода приведен на рисунке 3.5.

Пример. Определить точку минимума функции $\Phi = (x - 1,55)^2 + 2$. Очевидно, что $x^* = 1,55$. Для численного определения координаты этой точки применим метод золотого сечения. Примем начальный отрезок $x^* \in [1; 2]$.

Шаг 1. $x_1 = 2 - 0,618(2 - 1) = 1,382$; $x_2 = 1 + 0,618(2 - 1) = 1,618$

$\Phi(x_1) = (1,382 - 1,55)^2 + 2 = 2,028$; $\Phi(x_2) = (1,618 - 1,55)^2 + 2 = 2,005$.

Так как $\Phi(x_1) > \Phi(x_2)$, то в качестве $\bar{x}_2$ примем точку x_2 .

Тогда $a_1 = 1,382$, $b_1 = 2$.

Шаг 2. $x_3 = 1,382 + 2 - 1,618 = 1,764$; $\Phi(x_3) = 2,046$.

Так как $\Phi(x_2) < \Phi(x_3)$, то $a_2 = 1,382$ и $b_2 = 1,768$, и т.д.

Замечание. На исследуемом отрезке должна располагаться только одна точка минимума - точка глобального минимума.

3.1.3 Метод парабол

Ранее рассмотренные методы надежны, но медленны. Если на отрезке $[a, b]$ функция $\Phi(x)$ дифференцируема и унимодальна, то можно построить более быстрые методы, основанные на численном решении уравнения $\Phi'(x) = 0$.

Линеаризуем $\Phi'(x)$ в точке x_n :

$$\Phi'(x) \approx \Phi'(x_n) + \Phi''(x_n)(x - x_n) \approx 0$$

и построим итерационный процесс Ньютона, полагая $x = x_{n+1}$:

$$x_{n+1} = x_n - \Phi'(x_n) / \Phi''(x_n), \quad n = 0, 1, 2, \dots$$

Обычно расчет производных затруднен. Поэтому выгодно их заменить конечно-разностными аппроксимациями:

$$\Phi'(x_n) \approx (\Phi(x_n + \Delta x) - \Phi(x_n - \Delta x)) / 2\Delta x$$

$$\Phi''(x_n) \approx (\Phi(x_n + \Delta x) - 2\Phi(x_n) + \Phi(x_n - \Delta x)) / \Delta x^2.$$

Тогда имеем

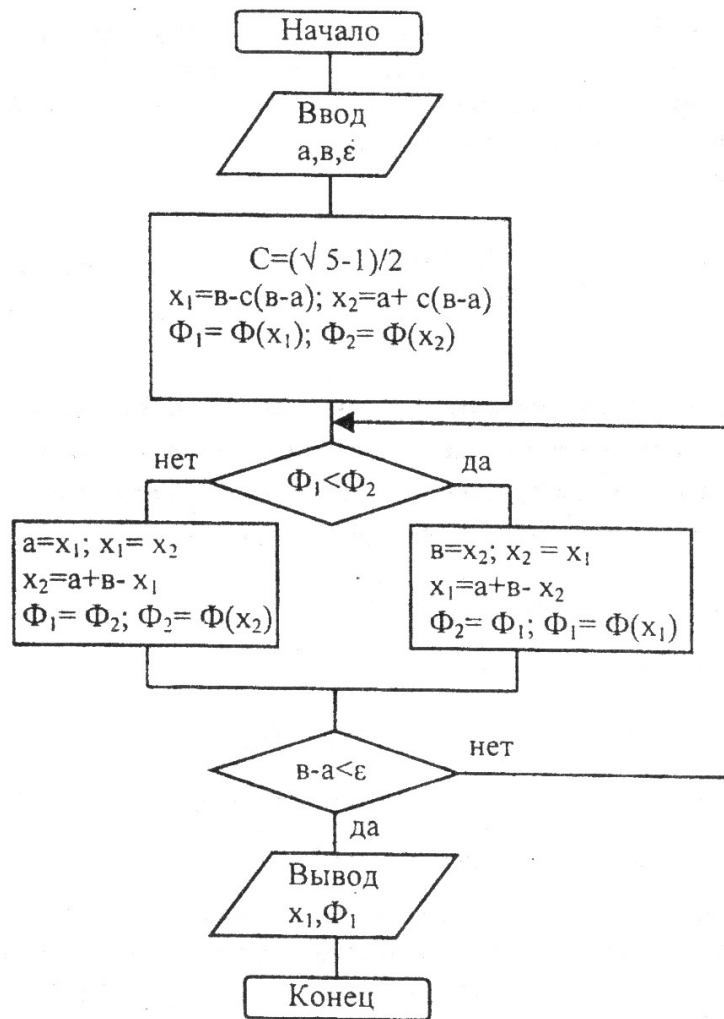


Рисунок 3.5 - Алгоритм метода золотого сечения

$$x_{n+1} = x_n - \frac{\Delta x}{2} \frac{\Phi(x_n + \Delta x) - \Phi(x_n - \Delta x)}{\Phi(x_n + \Delta x) - 2\Phi(x_n) + \Phi(x_n - \Delta x)}, \quad n = 0, 1, 2, \dots \quad (3.5)$$

Такой подход эквивалентен замене кривой на интерполяционную параболу, построенную по трем точкам.

Скорость сходимости вблизи точки экстремума близка к квадратичной.

Вопросы и упражнения для самоконтроля.

- 1) Что понимают под минимизацией функции?
- 2) В чем состоит метод сканирования?
- 3) Назовите достоинства и недостатки метода сканирования.
- 4) В чем отличие метода золотого сечения от метода сканирования?
- 5) На чем основывается метод парабол?
- 6) Найти минимум функций на указанном интервале методами сканирования, золотого сечения и парабол. Сравните эти методы:
 - а) $f(x) = 4x^3 - 8x^2 - 11x + 5$, $[0, 2]$; б) $f(x) = -\sin(x) - \sin 3x/3$, $[0, 2]$;

- в) $f(x) = e^x / x^2$, $[0.5, 3]$; г) $f(x) = (x - 1)^4$, $[-2, 3]$;
 д) $f(x) = \cos(x) + 2\sin(2x)$, $[0, 6]$; е) $f(x) = \sin(x) + 2,3\sin(3x)$, $[0.5, 5.5]$;
 ж) $f(x) = x^3 - 3x + 5$, $[-2, 4]$; з) $f(x) = \text{arctg}x - 1/2 \ln(1 + x)$, $[-6, 6]$;
 и) $f(x) = (x - 1)^2 \sin x$, $[-2, 3]$; к) $f(x) = -3 + 2x^4 - x^2 + 1.2x$, $[0.1, 3.1]$;
 л) $f(x) = \sin(2x) + 1.1\cos(4x)$, $[0, 5]$; м) $f(x) = x \sin(1/x)$, $[0.2, 1]$;
 н) $f(x) = x^2 + 2 - 1/x$, $[0.4, 3.4]$; о) $f(x) = \cos(2.5) - 0.12\text{tg}(x)$, $[0, 4]$;
 п) $f(x) = 6x - 14 + e^x$, $[2, 4]$; р) $f(x) = 0.9\cos(2x) + 1.4\cos(3.7x)$, $[0, 5]$;
 с) $f(x) = x^5 - x - 0.2$, $[-2, 2]$; т) $f(x) = (\sin(2.8))^3 + 1.6\cos(0.8)$, $[0, 6.5]$.

3.2 Многомерная минимизация

Задача

$$\Phi = \min_{X \in \Omega} \Phi(X) \quad (3.6)$$

означает поиск координат точки минимума $n+1$ -мерной криволинейной поверхности в $n+1$ -мерном пространстве. С целью упрощения геометрическую иллюстрацию методов будем рассматривать при $n=2$.

Как в топографии изобразим рельеф этой поверхности линиями уровня. Для этого проведем линии их пересечения с исходной криволинейной поверхностью $\Phi = \Phi(x_1, x_2)$ (рисунке 3.6,а). Проекции этих линий на плоскость x_1 - x_2 называются *линиями уровня*.

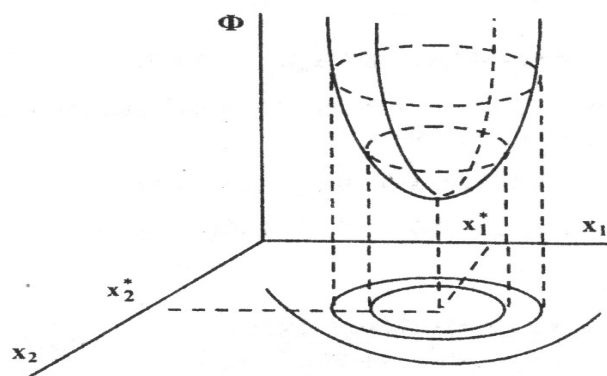
Различают три типа рельефа: овальный (рисунок 3.6,а), овражный (3.6,б) и сложный (3.6,в). В последнем случае поверхность может содержать точки минимума, максимума и седловые точки. На рисунке 3.6,в точками x_1^* и x_3^* отмечены точки локальных минимумов. Из них точкой глобального минимума является x_3^* . Точка x_2^* называется седловой точкой.

В каждой из точек экстремума, а также в седловой точке выполняются условия

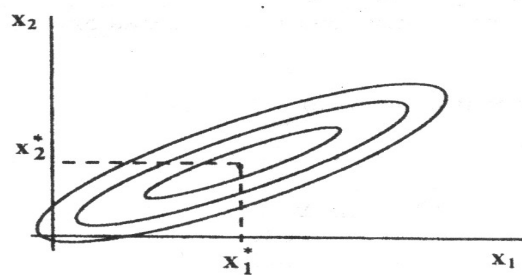
$$\frac{\partial \Phi}{\partial x_i} = 0, \quad i = \overline{1, n}. \quad (3.7)$$

Все методы поиска минимума функции сводятся к построению траектории, вдоль которой $\Phi \rightarrow \min$, а $\frac{\partial \Phi}{\partial \Phi} = 0$. Казалось бы, проще решить систему (3.7) относительно неизвестных x_i . В некоторых случаях так и поступают.

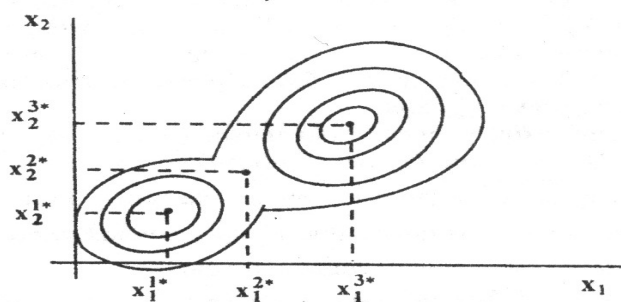
Однако большинство задач настолько сложны, что процесс решения нелинейной системы оказывается более трудоемкой задачей. Кроме того, может случиться, что решение системы (3.7) не будет являться решением задачи (3.6), например, соответствовать координате седловой точки. Поэтому чаще используют численные методы минимизации.



а)



б)



в)

Рисунок 3.6 - Построение линий уровня (а); линии уровня овражного (б) и сложного (в) рельефов

3.2.1 Метод сканирования

Пусть $X^* \in \Omega_0$, где Ω_0 представляет собой прямоугольный n -мерный параллелепипед со сторонами $[a_i^{(0)}, \hat{a}_i^{(0)}]$ (рисунок 3.7). В области Ω_0 построим n -мерную равномерную сетку. Для этого каждый отрезок $[a_i^{(0)}, \hat{a}_i^{(0)}]$ разделим на m малых интервалов длиной $h_i^{(0)} = \frac{(\hat{a}_i^{(0)} - a_i^{(0)})}{m}$ и координаты узлов сетки определим: $x_i^{(k)} = a_i^{(0)} + h_i^{(0)} \cdot k$, $k = \overline{0, m}$.

В каждом узле рассчитаем значения функции и выберем наименьшее значение:

$$\Phi^{(1)} = \min_{x \in \Omega_0} \Phi(x_1^j, x_2^i, \dots, x_n^k) \Rightarrow \{x_1^{(1)*}, x_2^{(1)*}, \dots, x_n^{(1)*}\}$$

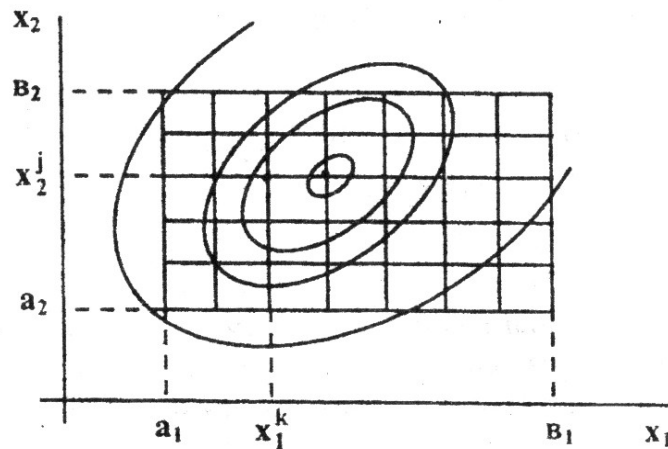


Рисунок 3.7 - Построение сетки

Затем, как в случае одномерного поиска, сузим область Ω_0 до Ω_1 :

$$a_i^{(1)} = x_i^{(1)*} - p \cdot h_i^{(0)}; \quad b_i^{(1)} = x_i^{(1)*} + p \cdot h_i^{(0)}, \quad i = 1, n.$$

Каждый отрезок $[a_i^{(1)}, \hat{a}_i^{(1)}]$ разделим на m более мелких интервалов длиной

$$h_i^{(k)} = \frac{(\hat{a}_i^{(1)} - a_i^{(1)})}{m}.$$

В каждом узле рассчитаем значения функции и выберем наименьшее значение и т.д.

Критерием останова является одно из условий:

$$\hat{O}^{k+1} - \hat{O}^{(k)} < \varepsilon \quad \text{или} \quad \max_i h_i < \varepsilon. \quad (3.8)$$

Замечание. Ввиду резкого увеличения объема вычислений с увеличением размерности решаемой задачи метод сканирования при $n \geq 3$ используется лишь для предварительной оценки рельефа поверхности функции.

3.2.2 Спуск по координатам

Зафиксируем все переменные $x_i^{(0)}$ за исключением x_1 . Варьируя значение x_1 . Используя любой из методов поиска минимума функции одной переменной, определим минимум:

$$\hat{O}_1^{(1)} = \min_{x_1 \in [a_1, \hat{a}_1]} \hat{O}(\tilde{o}_1, \tilde{o}_2^{(0)}, \dots, \tilde{o}_n^{(0)}).$$

Пусть эта точка - $x_1^{(1)}$. Далее фиксируем $x_1^{(1)}$ и, варьируя x_2 найдем минимум:

$$\hat{O}_2^{(1)} = \min_{x_2 \in [a_2, \hat{a}_2]} \hat{O}(\tilde{o}_1^{(1)}, x_2, \tilde{o}_2^{(0)}, \dots, \tilde{o}_n^{(0)}),$$

и т.д. по всем x_i , $i = \overline{1, n}$.

Затем на втором шаге вновь начинаем варьировать x_1 , x_2 и т.д. по x_n включительно, что позволит определить $X^{(2)}$. Итерационный процесс завершается по достижении минимума функции с заданной точностью (рисунок 3.8,а).

При овальных линиях уровня метод сходится хорошо. Однако при сильной овражности минимум может быть не достигнут. Эта трудность в особенности возникает при расположении оврага под углом $\pm 45^\circ$ к осям координат (рисунок 3.8,б) - возникают частые колебания с малыми шагами, что может привести к преждевременному завершению итерационного процесса. При наличии нескольких локальных минимумов эффективность поиска точки глобального минимума существенно зависит от выбора точки начального приближения (рисунок 3.8,в).

Алгоритм метода приведен на рисунке 3.9.

3.2.3 Метод градиента

В отличие от спуска по координатам в методе градиента шаги совершаются в направлении наибыстрейшего уменьшения целевой функции, что ускоряет процесс поиска точки экстремума.

Градиент скалярной функции $\Phi(X)$ - это n-мерный вектор:

$$\text{grad } \Phi(X) = \nabla \Phi(X) = \left\{ \frac{\partial \Phi}{\partial x_1}, \dots, \frac{\partial \Phi}{\partial x_n} \right\}.$$

Направление градиента совпадает с направлением нормали (нормаль-перпендикуляр ко всем касательным в рассматриваемой точке поверхности) и указывает направление наискорейшего возрастания функции.

Поиск точки минимума производится в два этапа:

1) рассчитываются все

$$\frac{\partial \Phi}{\partial x_1}, \dots, \frac{\partial \Phi}{\partial x_n}$$

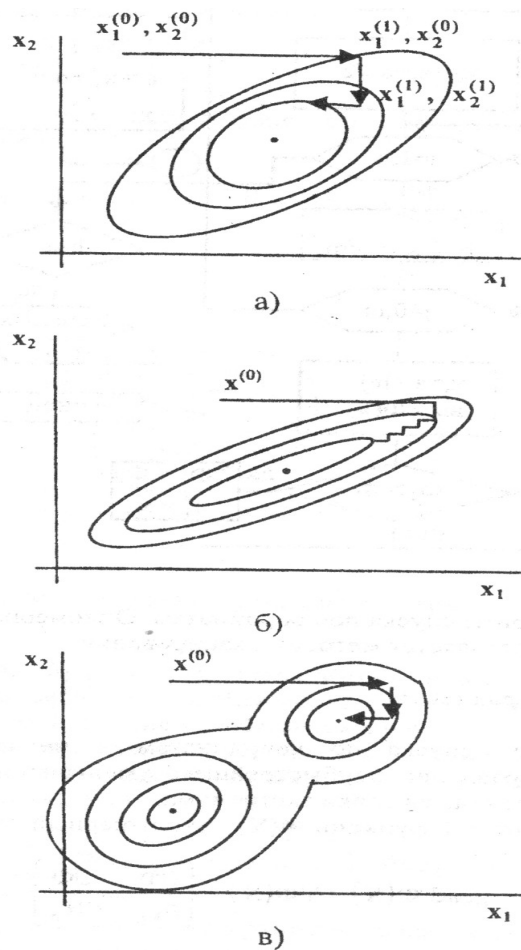


Рисунок 3.8 - Иллюстрация спуска по координатам

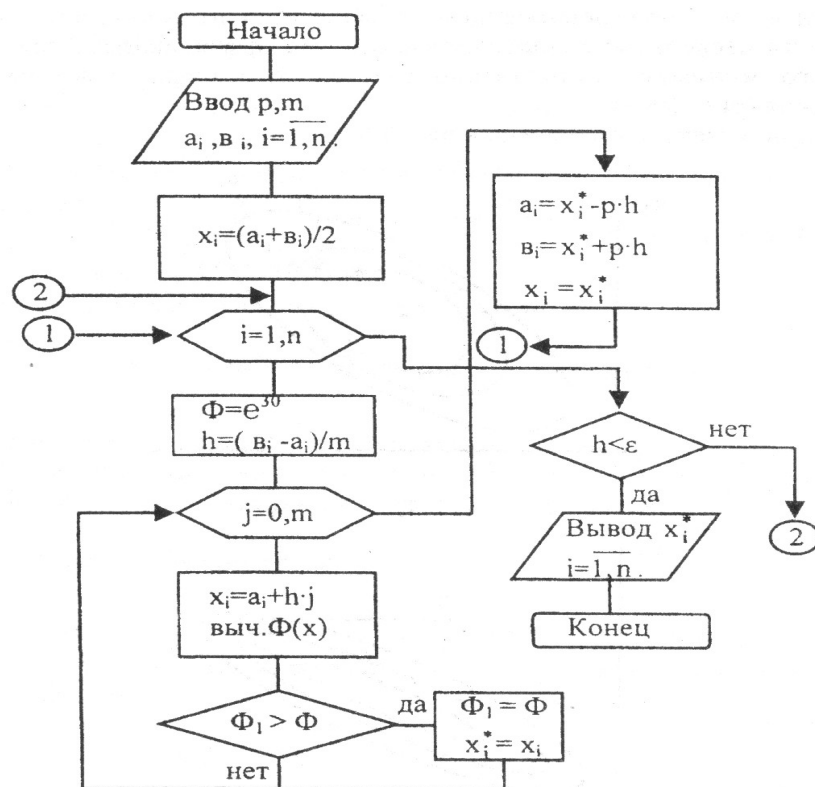


Рисунок 3.9 - Алгоритм спуска по координатам. Одномерная минимизация осуществляется методом сканирования

2) осуществляется шаг в направлении убывания функции, т.е. в обратном направлении градиента:

$$x_1^{(k+1)} = x_1^{(k)} - h \frac{\partial \Phi}{\partial x_1}, \dots, x_n^{(k+1)} = x_n^{(k)} - h \frac{\partial \Phi}{\partial x_n}, \quad k = 0, 1, 2, \dots \quad (3.9)$$

Величина шага будет определяться абсолютной величиной градиента, а также величиной коэффициента h . Чем меньше h , тем дольше поиск. С другой стороны, при больших h дольше колебания около точки экстремума.

При минимизации сложной функции, для которой невозможно выписать в явном виде выражение вектор-градиента, обычно используют линейную аппроксимацию производной:

$$\frac{\partial \Phi}{\partial x_i} \approx \frac{1}{\Delta x_i} (\Phi(x_1, \dots, x_i + \Delta x_i, \dots, x_n) - \Phi(x_1, \dots, x_i, \dots, x_n)).$$

При сильной овражности поверхности точка экстремума из-за частых колебаний может быть не достигнута. Иллюстрация метода приведена на рисунке 3.10, алгоритм - на рисунке 3.11.

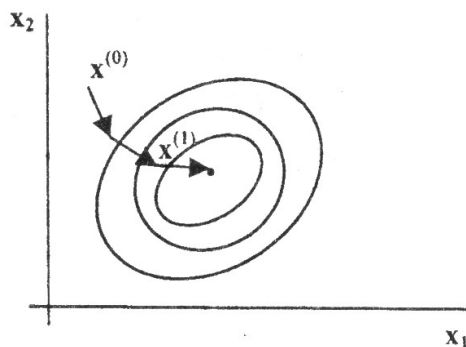


Рисунок 3.10 - Иллюстрация метода градиента

Пример. Найти точку минимума функции

$$\Phi = x_1^2 - x_1 x_2 + 2x_2^2.$$

Определим частные производные:

$$\frac{\partial \Phi}{\partial x_1} = 2x_1 - x_2; \quad \frac{\partial \Phi}{\partial x_2} = -x_1 + 4x_2.$$

Приравняв их нулю, можем найти аналитическое решение:

$$x_2 = 2x_1; \quad -x_1 + 8x_1 = 0 \quad \Rightarrow \quad x_1^* = 0; \quad x_2^* = 0.$$

Рассмотрим далее численное решение.

Пусть $x^{(0)} = \{1, 1\}$ и $h=0,2$

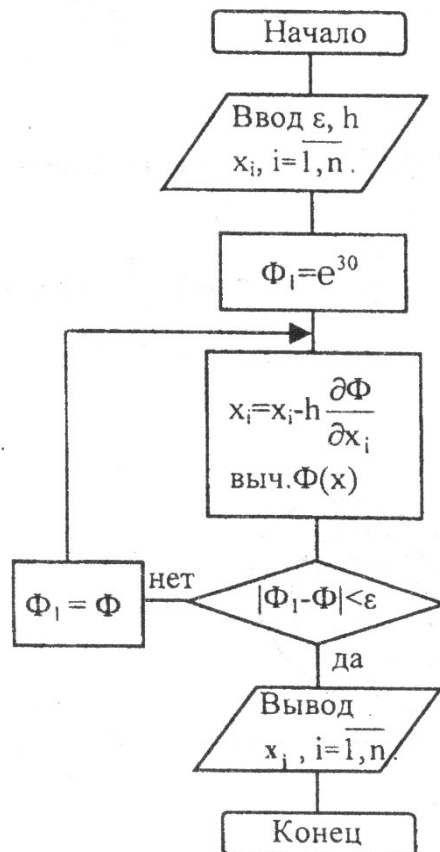


Рисунок 3.11 - Алгоритм метода градиента

$$x_1^{(1)} = 1 - 0,2 \cdot (2 \cdot 1 - 1) = 0,8$$

$$x_2^{(1)} = 1 - 0,2 \cdot (-1 + 4 \cdot 1) = 0,4$$

$$x_1^{(2)} = 0,8 - 0,2 \cdot (2 \cdot 0,8 - 0,4) = 0,56$$

$$x_2^{(2)} = 0,4 - 0,2 \cdot (-0,8 + 4 \cdot 0,4) = 0,24$$

$$x_1^{(3)} = 0,56 - 0,2 \cdot (2 \cdot 0,56 - 0,24) = \dots$$

$$x_2^{(3)} = 0,24 - 0,2 \cdot (-0,56 + 4 \cdot 0,24) = \dots$$

и т.д. до достижения заданной точности.

В точке $x^{(0)} = \{1,1\}$ $\text{grad } \phi = \{1,3\}$, следовательно, по оси x_2 функция возрастает круче, чем по оси x_1 (рисунок 3.12).

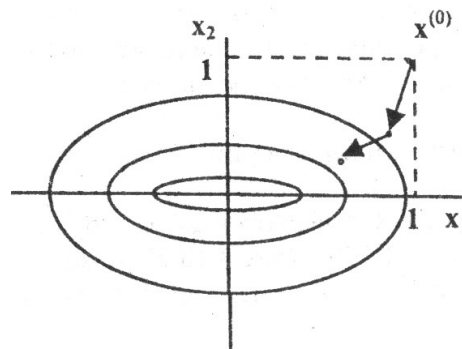


Рисунок 3.12 - Графическая иллюстрация примера

3.2.4 Слепой поиск

Методы случайного поиска представляются рекуррентной формулой

$$X^{(k+1)} = X^{(k)} + h\alpha^{(k)}, \quad k=0,1,2,\dots,$$

где α - вектор случайных чисел.

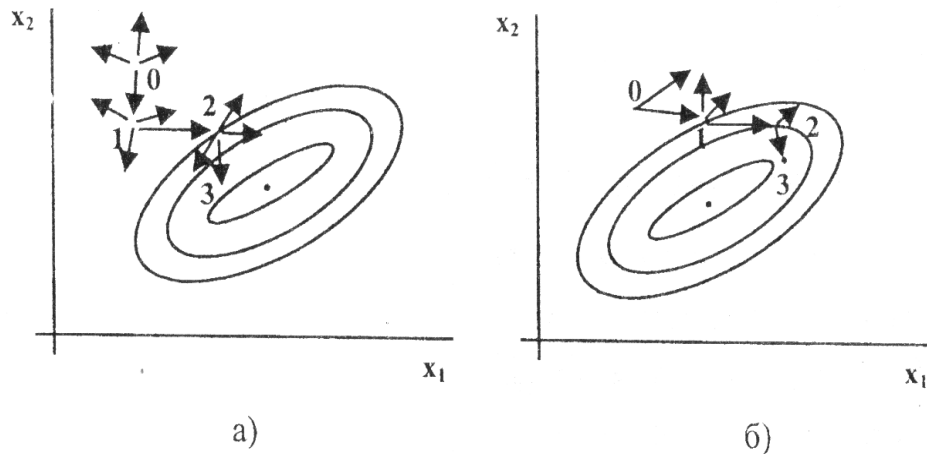


Рисунок 3.13 - Иллюстрации методов слепого (а) и случайного (б) поисков

В слепом поиске выбираются N случайных точек, в которых насчитываются значения функции. На следующем этапе осуществляется шаг в точку с наименьшим значением функции и из этой точки вновь осуществляется слепой поиск (рисунок 3.13,а).

Теоретически можно достигнуть сколь угодно большой точности. Однако необходимо много раз рассчитывать значение функции. Если функция достаточно сложная, то поиск затягивается надолго.

3.2.5 Метод случайных направлений

Пробные шаги из точки $X^{(k)}$ делаются до тех пор, пока не найдется точка, в которой $\Phi(X^{(k+1)}) < \Phi(X^{(k)})$ (рисунок 3.13,б). Тогда следующий поиск осуществляют из точки $X^{(k+1)}$ и т.д. При приближении к точке экстремума величину h уменьшают.

По сравнению со слепым поиском длительность поиска точки экстремума существенно уменьшается.

Методы случайного поиска эффективно используются при больших n .

3.2.6 Условная минимизация. Метод штрафных функций

При наличии ограничений на переменные имеем задачу на условный минимум:

$$\begin{aligned}\hat{O} &= \min_{X \in \Omega} \hat{O}(\delta) \\ \varphi_j(X) &= 0, j = \overline{1, m} \\ \psi_k(X) &\geq 0, k = \overline{1, p}.\end{aligned}\quad (3.11)$$

Функция Φ должна достигать минимума на кривой φ и не должна входить в запретную область ψ (рисунок 3.14).

Для решения этой задачи вводят следующую функцию:

$$J = \left[\Phi(X) + \mu \sum_{j=1}^m \varphi_j^2(X) + \eta \sum_{k=1}^p \psi_k^2(X) (1 - \text{sign } \psi_k(X)) \right] \quad (3.12)$$

$$\mu, \eta > 0$$

Последние два слагаемых взяты таким образом, что они обращаются в нуль при выполнении условий исходной задачи. Если условия нарушены, то они увеличивают J , причем тем больше, чем сильнее нарушены условия -штраф за нарушения.

Если коэффициенты μ и η взяты достаточно большими, то малейшее отклонение от условий приводит к резкому возрастанию J . Однако при этом вблизи границ появляется сильная овражность, что затрудняет поиск. Обычно поиск точки экстремума начинают с малых μ и η и последовательно увеличивают их на последующих шагах. Наиболее часто при его реализации используют методы случайного поиска.

Алгоритм метода приведен на рисунке 3.15.

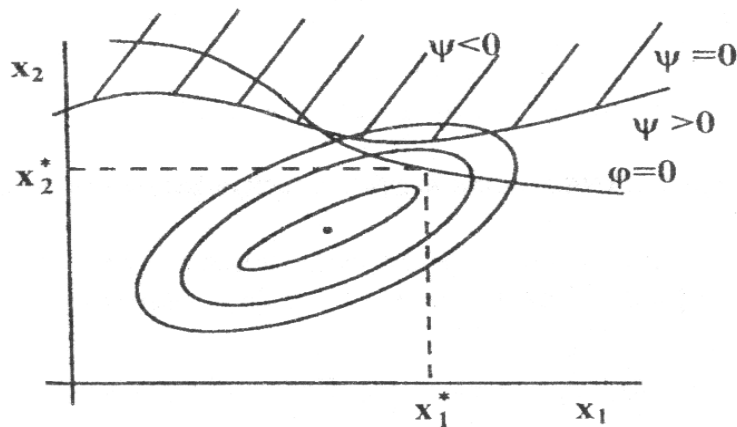


Рисунок 3.14 - Иллюстрация задачи с ограничениями

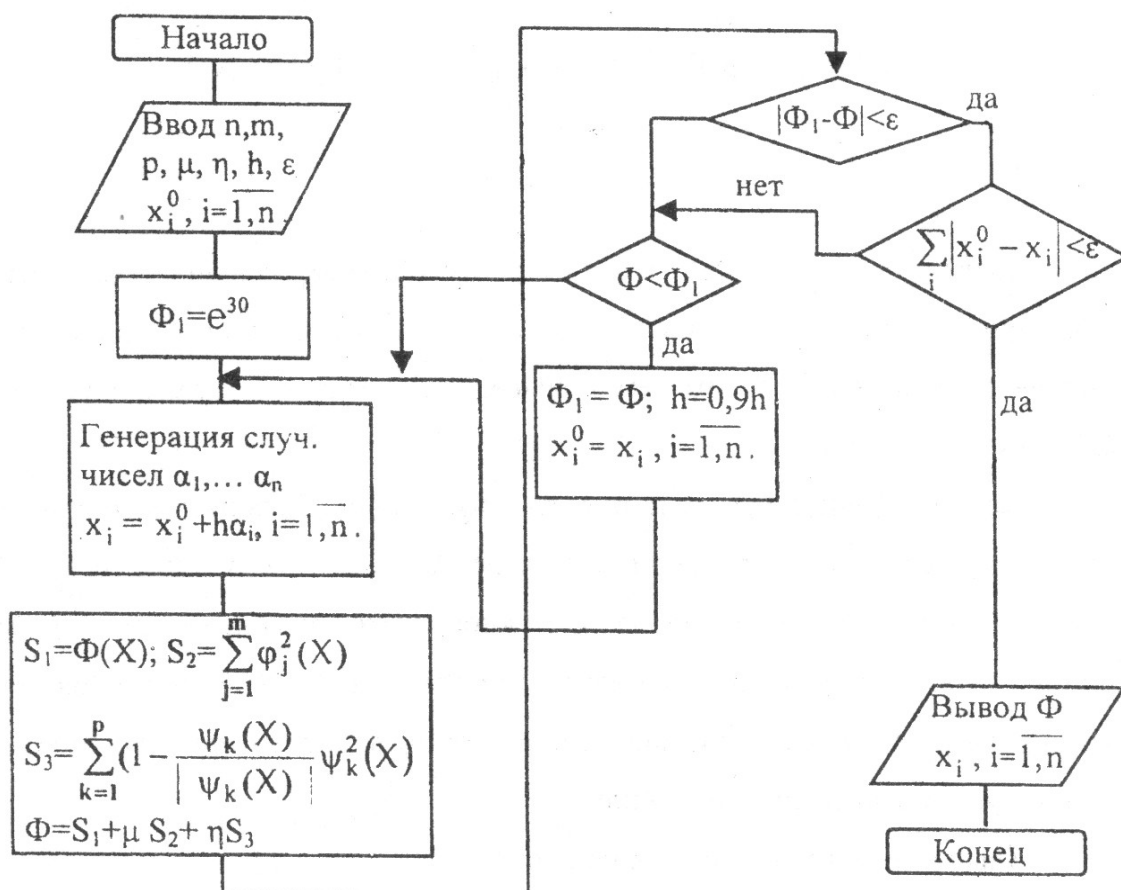


Рисунок 3.15 - Алгоритм метода штрафных функций

Вопросы и упражнения для самоконтроля.

- 1) Что понимают под линиями уровня?
- 2) Какие типы рельефа существуют?
- 3) Какое должно выполняться условие в точке экстремума?
- 4) При каком условии происходит останов метода сканирования?
- 5) В чем состоит метод спуска по координатам?
- 6) Чем отличается метод градиента от метода спуска по координатам?
- 7) Как осуществляется поиск точки минимума в методе градиента?
- 8) Охарактеризуйте метод слепого писка и метод случайных направлений.
- 9) Какие условия налагаются на функцию в методе штрафных функций?

10) Найти точку минимума функции методом сканирования и методом градиента, и выбирая в качестве начальной точки (0,1) и (0,0) ($\epsilon=10^{-3}$, $h=0,1$):

- а) $f(x_1, x_2) = x_1^2 + x_2^2 - 8(x_1 + x_2)$; б) $f(x_1, x_2) = x_1 + 2x_2^2 + 1$;
- в) $f(x_1, x_2) = 3x_1^2 - 4x_1x_2 + 2x_2^2 + 4\sqrt{5}(x_1 + 3x_2)$;
- г) $f(x_1, x_2) = (x_1^2 - x_2^2) + (x_1 - 1)^2$; д) $f(x_1, x_2) = (x_2^3 - 5)^2 \ln x_1$;
- е) $f(x_1, x_2) = -x_1^2 - x_2^2 + x_1x_2 - x_1 + 2x_2$.

4 Решение обыкновенных дифференциальных уравнений

Рассмотрим следующую задачу, называемую задачей Коши:

$$\begin{cases} y' = f(x, y), x \in [a, b] \\ y_0 = y(a). \end{cases} \quad (4.1)$$

Здесь y_0 называется начальным условием. Задача заключается в определении неизвестной функции y на отрезке $[a, b]$. Как следует из рисунка 4.1, различным значениям y_0 соответствуют различные значения y - семейство интегральных кривых.

В отличие от курса математического анализа, определяющего общее решение, здесь рассматривается *частная задача*.

Процесс решения этой задачи называется интегрированием дифференциального уравнения. В большинстве случаев используют приближенные методы интегрирования дифференциального уравнения. Наиболее часто используются методы Рунге-Кутты различных порядков.

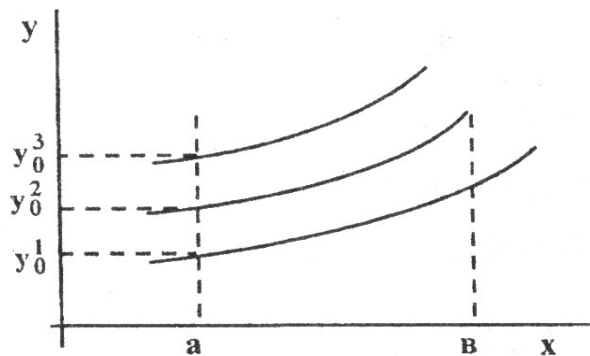


Рисунок 4.1 - Семейство интегральных кривых

4.1 Метод Эйлера (Рунге-Кутта первого порядка)

Этот метод является наиболее простым из используемых методов. Разложим функцию $y(x)$ в точке c в ряд Тейлора:

$$y(x) = y(c) + y'(c)(x - c) + \frac{y''(c)}{2!}(x - c)^2 + \dots$$

и выделим линейную часть.

$$y(x) \approx y(c) + y'(c) \cdot (x - c).$$

Пусть на отрезке $[a, b]$ построена равномерная сетка с шагом h :

$$x_j = a + h \cdot j, \quad j = \overline{0, n}; \quad h = (b - a) / n.$$

Пусть $s = x_j$; $x = x_{j+1}$. Тогда, учитывая (4.1) и полагая $x = x_{j+1}$; $y' = f(x, y(x))$. $x_{j+1} - x_j = h$, имеем $y(x_{j+1}) = y(x_j) + hf(x_j, y(x_j))$ - уравнение прямой.

Геометрическая иллюстрация метода приведена на рисунке 4.2.

Через точку (x_0, y_0) проведем касательную с тангенсом угла наклона $y'(x_0)$:

$$y(x) = y(x_0) + (x - x_0)y'(x_0).$$

В точке x_1 значение касательной равно y_1 . Через точку (x_1, y_1) проведем другую касательную с тангенсом угла наклона $y'(x_1)$, которая позволит определить точку (x_2, y_2) , и т.д.

Обозначив тангенс угла наклона данной прямой k_1 :

$$k_1 = y'_j = f(x_j, y(x_j)),$$

получим рекуррентную формулу

$$\begin{cases} y_{j+1} = y_j + hk_1 \\ k_1 = f(x_j, y_j), y_0 = y(x_0), j = \overline{0, n-1}, \end{cases} \quad (4.2)$$

называемую методом Эйлера.

Погрешность метода $r \sim h$ - первый порядок точности. Поэтому этот метод называют также методом Рунге-Кутты первого порядка.

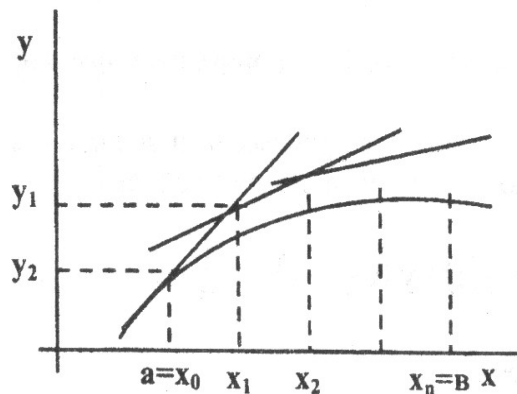


Рисунок 4.2 - Иллюстрация метода Эйлера

Пример. Решить дифференциальное уравнение

$$y' = x + y/2; \quad x \in [1, 3]; \quad y_0 = 0,5$$

(общее аналитическое решение: $y = 6,75e^x - (x + 1)$).

Пусть $h=0,2$.

$$\underline{j=0}: k_1 = x_0 + y_0/2 = 1 + 0,5/2 = 1,25$$

$$y_1 = y_0 + h \cdot k_1 = 0,5 + 0,2 \cdot 1,25 = 0,75$$

$$\underline{j=1}: k_1 = x_1 + y_1/2 = 1,2 + 0,75/2 = 1,57$$

$$y_2 = y_1 + h \cdot k_1 = 0,75 + 0,2 \cdot 1,57 = 1,06$$

$$\underline{j=2}: k_1 = 1,4 + 1,06/2 = 1,93$$

$$y_3 = 1,06 + 0,2 \cdot 1,93 = 1,45 \quad \text{и т.д. до } x_n = 3.$$

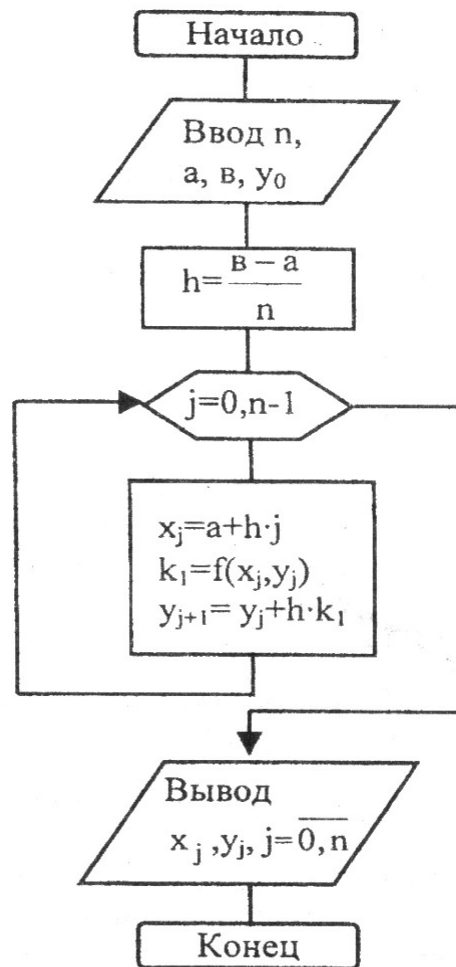


Рисунок 4.3 - Алгоритм метода Эйлера

4.2 Модифицированный метод Эйлера (Рунге-Кутта второго порядка)

По аналогии с методом Эйлера проведем через точку A касательную L_1 с тангенсом угла наклона k_1 (рисунок 4.4). Через точку C с координатами $(x_j + h/2, y_j + h/2 y'_j)$ проведем прямую L_2 с тангенсом угла наклона

$$k_2 = f(x_j + h/2, y_j + h/2 y'_j).$$

Далее через точку A проведем прямую L_3 , параллельную прямой L_2 .
Уравнение этой прямой

$$y = y_j + (x - x_j)k_2.$$

Пусть $x = x_{j+1} = x_j + h$. Тогда $y = y_{j+1}$. Обозначив $k_1 = f(x_j, y_j)$, получим

$$\begin{cases} y_{j+1} = y_j + h \cdot k_2 \\ k_1 = f(x_j, y_j) \\ k_2 = f(x_j + h/2, y_j + h/2 \cdot k_1) \\ y_0 = y(x_0), \quad j = \overline{0, n-1} \end{cases} \quad (4.3)$$

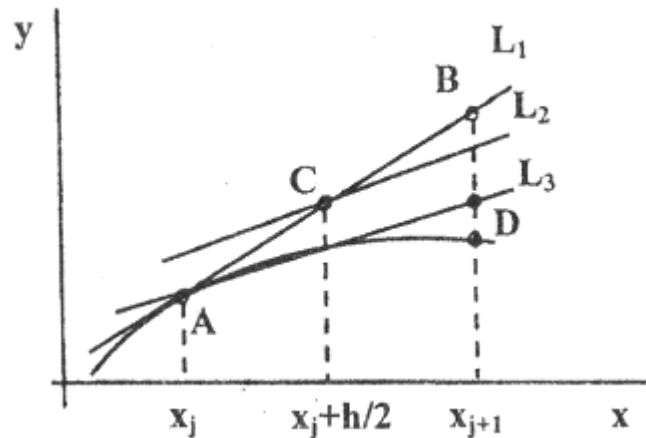


Рисунок 4.4 - Иллюстрация модифицированного метода Эйлера

Из рисунка 4.4 следует, что метод (4.3) обеспечивает большую точность, чем метод (4.2). Действительно, точка D , соответствующая решению (4.3), располагается гораздо ближе к истинному значению функции y , чем точка B , соответствующая решению (4.2).

Погрешность модифицированного метода Эйлера $r \sim h^2$.

Пример. Предыдущий пример решить методом Рунге-Кутты 2-го порядка.

$$\underline{j=0}: k_1 = x_0 + y_0/2 = 1 + 0,5/2 = 1,25$$

$$k_2 = (x_0 + h/2) + (y_0 + h k_1/2) = (1 + 0,2/2) + (1 + 0,2/2) + (0,5 + 0,2 \cdot 1,25/2)/2 = 1,412$$

$$y_1 = y_0 + h \cdot k_2 = 0,5 + 0,2 \cdot 1,412 = 0,78$$

$$\underline{j=1}: k_1 = x_1 + y_1/2 = 1,2 + 0,78/2 = 1,59$$

$$k_2 = (x_1 + h/2) + (y_1 + h k_1/2)/2 = (1 + 0,1) + (0,78 + 0,1 \cdot 1,59)/2 = 1,77$$

$$y_2 = y_1 + h \cdot k_2 = 0,78 + 0,2 \cdot 1,77 = 1,13$$

и т.д.

4.3 Метод Рунге-Кутты четвертого порядка

Аналогичным образом можно получить методы Рунге-Кутты высших порядков. Наиболее часто используют метод Рунге-Кутты четвертого порядка. Приведем его без обоснования используемых формул:

$$\begin{cases} y_{j+1} = y_j + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4) \\ k_1 = f(x_j, y_j) \\ k_2 = f(x_j + h/2, y_j + h/2 \cdot k_1) \\ k_3 = f(x_j + h/2, y_j + h/2 \cdot k_2) \\ k_4 = f(x_j + h, y_j + h \cdot k_3) \\ y_0 = y(x_0), \quad j = \overline{0, n-1}. \end{cases} \quad (4.4)$$

Пример. Решить систему дифференциальных уравнений

$$\begin{cases} y'_1 = -0,1xy_1 & y_{10} = 10 \\ y'_2 = 0,1xy_1 - 2y_2 & y_{20} = 0, \quad x \in [2,8]. \end{cases}$$

Пусть $h=0,5$. В соответствии с (4.4) имеем:

$$\begin{aligned} \underline{j=0}: \quad k_{11} &= -0,1 \cdot 2 \cdot 10 = -2,00 \\ k_{21} &= 0,1 \cdot 2 \cdot 10 - 2 \cdot 0 = 2,00 \\ k_{12} &= -0,1(2+0,25) \cdot (10+0,25(-2)) = -2,14 \\ k_{22} &= 0,1(2+0,25) \cdot (10+0,25(-2)) - 2(0+0,25 \cdot 2) = 1,14 \\ k_{13} &= -0,1 \cdot (2+0,25) \cdot (10+0,25(-2,14)) = -2,13 \\ k_{23} &= 0,1(2+0,25) \cdot (10+0,25(-2,14)) - 2(0+0,25 \cdot 1,14) = 1,56 \\ k_{14} &= -0,1(2+0,5) \cdot (10+0,5(-2,13)) = -2,23 \\ k_{24} &= 0,1(2+0,5) \cdot (10+0,5(-2,13)) - 2(0+0,5 \cdot 1,56) = 0,67 \\ y_{11} &= 10 + 0,5/6(-2-2 \cdot 2,14-2 \cdot 2,13-2,23) = 8,94 \\ y_{21} &= 0 + 0,5/6(2+2 \cdot 1,14+2 \cdot 1,56+0,67) = 0,67. \end{aligned}$$

Таким образом, в узле $x_1=2,5$ значения функций соответственно равны $y_{11}=8,94$, $y_{21}=0,67$. Вычислим далее их значения в узле $x_2=3,0$.

$$\begin{aligned} \underline{j=1}: \quad k_{11} &= -0,1 \cdot 2,5 \cdot 8,94 = -2,24 \\ k_{21} &= 0,1 \cdot 2,5 \cdot 8,94 - 2 \cdot 0,67 = 0,90 \\ k_{12} &= -0,1(2,5+0,25) \cdot (8,94+0,25(-2,24)) = -2,30 \\ k_{22} &= 0,1(2,5+0,25) \cdot (8,94+0,25(-2,24)) - 2(0,67+0,25 \cdot 0,90) = 0,51 \\ k_{13} &= -0,1(2,5+0,25) \cdot (8,94+0,25(-2,30)) = -2,30 \\ k_{23} &= 0,1(2,5+0,25) \cdot (8,94+0,25(-2,30)) - 2(0,67+0,25 \cdot 0,51) = 0,71 \\ k_{14} &= -0,1(2,5+0,5) \cdot (8,94+0,5(-2,30)) = -2,34 \\ k_{24} &= 0,1(2,5+0,5) \cdot (8,94+0,5(-2,30)) - 2(0,67+0,5 \cdot 0,71) = 0,29 \\ y_{12} &= 8,94 + 0,5/6(-2,24-2 \cdot 2,30-2 \cdot 2,30-2,34) = 7,79 \\ y_{22} &= 0,67 + 0,5/6(0,90+2 \cdot 0,51+2 \cdot 0,71+0,29) = 0,97. \end{aligned}$$

Аналогичным образом рассчитываются значения функций в узлах $x_3=3,5, \dots, x_{13}=8,0$.

Погрешность этого метода $r \sim h^4$.

4.4 Метод прогноза и коррекции

В рассмотренных методах (4.2) - (4.4) при расчете y_{j+1} используется информация с j -го узла. Но в нашем распоряжении имеются информации с $j-1$, $j-2$ и более ранних узлов. Хотелось бы их использовать.

Линейной аппроксимацией производной экстраполируем известную информацию на $j+1$ -й узел:

$$\frac{y_{j+1} - y_j}{\Delta x} \approx f(x_j, y_j) \Rightarrow y_{j+1}^{(0)} = y_j + \Delta x f(x_j, y_j). \quad (4.5)$$

Уравнение (4.5) называется *прогнозом*. По нему определяется грубое, первоначальное значение функции в узле $j+1$.

Коррекцию (уточнение) проводят по рекуррентному уравнению

$$y_{j+1}^{(k+1)} = y_j + \frac{\Delta x}{2} (f(x_j, y_j) + f(x_{j+1}, y_{j+1}^{(k)})), \quad k = 0, 1, \dots \quad (4.6)$$

Итерационный процесс (4.6) завершается при выполнении условия

$$\left| y_{j+1}^{(k+1)} - y_{j+1}^{(k)} \right| < \varepsilon.$$

Необходимым условием сходимости метода является

$$h < 2/M, \quad \text{где} \quad M \geq \left| \frac{\partial f}{\partial y} \right|.$$

Если это условие на величину шага не будет выполнено, то процесс (4.6) будет расходиться.

При решении задач с крутым возрастанием функции шаги h выбираются малыми, что приводит к увеличению объема вычислений.

4.5 Решение краевых задач. Метод стрельб

Рассмотрим задачу

$$\begin{cases} y' = f(x, y, z) \\ z' = g(x, y, z) \\ \varphi(y_0, z_0) = 0 \\ \psi(y_n, z_n) = 0 \end{cases} \quad (4.7)$$

В этой задаче зафиксированы значения функций y и z как на левом ($\varphi=0$), так и на правом ($\psi=0$) концах отрезка - краевая задача.

Выберем произвольное значение $y^{(l)}(x_0)$ и из условия $\varphi(y_0, z_0)=0$ определим $z^{(l)}(x_0)$. Эти значения возьмем в качестве начальных условий задачи Коши и проинтегрируем ее любым методом слева направо. При этом поначалу на

правой границе решение наверняка не будет удовлетворять граничному значению ψ . Тогда возвращаемся на начало. Принимаем другие начальные условия задачи Коши $y^{(2)}(x_0)$, $z^{(2)}(x_0)$ и вновь интегрируем. Вновь проверяем, удовлетворяет ли новое решение условию на правой границе, и т.д.

Описанный метод носит название "*метод стрельб*" (рисунки 4.5, 4.6). Коррекцию метода можно проводить, используя, например, метод половинного деления.

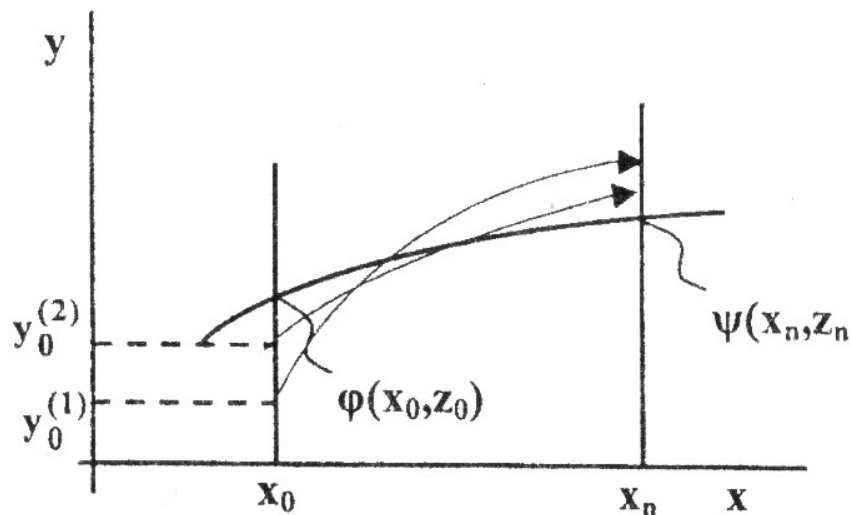


Рисунок 4.5 - Иллюстрация метода стрельб

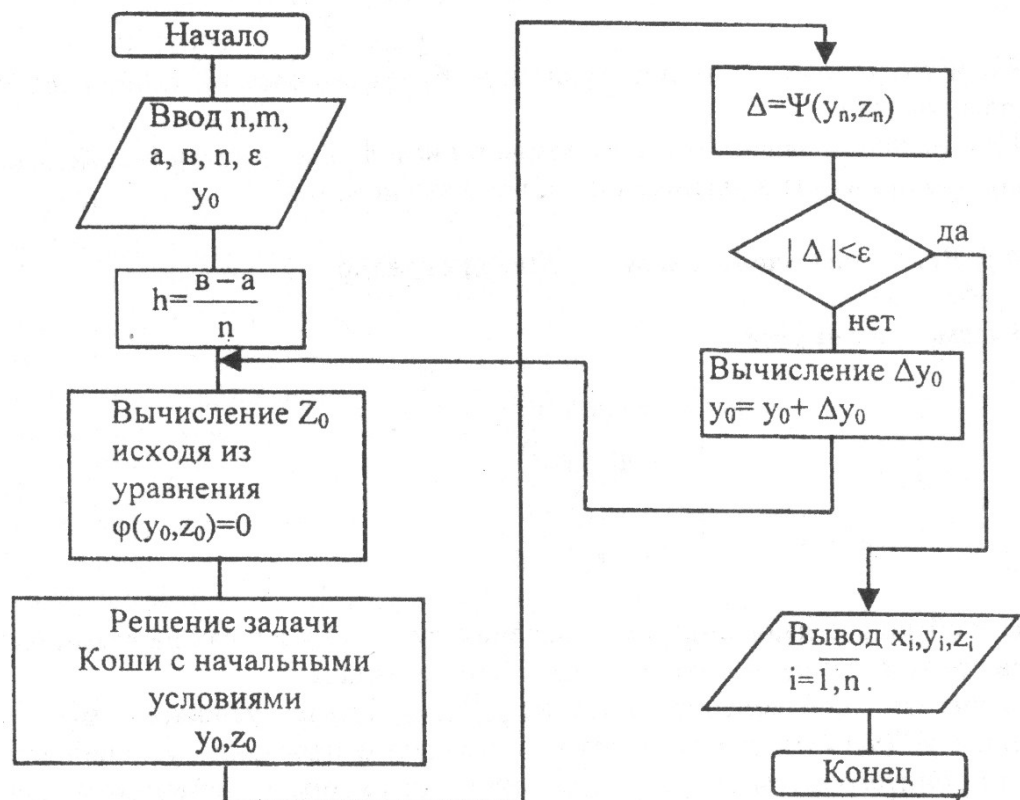


Рисунок 4.6 - Алгоритм метода стрельб

Вопросы и упражнения для самоконтроля.

- 1) Что является решением дифференциального уравнения? Что это означает с геометрической точки зрения?
- 2) В чем заключается сущность метода Эйлера?
- 3) Почему модифицированный метод Эйлера обеспечивает большую точность чем метод Эйлера?
- 4) Чему равна погрешность модифицированного метода Эйлера?
- 5) Какое необходимое условие должно выполняться, чтобы обеспечить сходимость метода прогноза и коррекции?
- 6) В чем состоит метод стрельб?
- 7) Численно решить задачу Коши для обыкновенного дифференциального уравнения

$$\begin{cases} y' = f(x, y), \\ y(x_0) = y_0. \end{cases}$$

на отрезке $[a, b]$ с шагом h :

- а) методом Эйлера; б) методом Рунге-Кутты 2-го порядка; в) методом Рунге-Кутты 4-го порядка; г) методом прогноза и коррекции.

Найти точное решение задачи. Построить графики точного и приближенных решений.

№	$y'=f(x,y)$	a	b	y_0	h	№	$y'=f(x,y)$	a	b	y_0	h
1	$x^2y + y$	0	4	1	0,1	9	$\delta + \cos \frac{y}{\sqrt{11}}$	2,1	3,1	2,5	0,1
2	$\sqrt{4\delta^2 + 1} - 3\delta^2$	2,6	4,6	1,8	0,2	10	$\sin x \cdot \cos^2 y$	-1	3	0	0,1
3	$-y \cos t + \frac{\sin(2t)}{2}$	0	1	0	0,2	11	$-\frac{y}{2x} + x^2$	1	1	1	0,1
4	$\frac{\delta}{\sqrt{1 + \delta^2 + \delta^2}}$	-1	1	0,2	0,1	12	$\sqrt{x^2 + 0,5y^2} + 1$	0	2	1	0,2
5	$\sin(x^2 + y^2)$	0	4	0	0,1	13	$y \cos x - \sin 2x$	2	4	3,5	0,1
6	$\cos(1,5y + x)^2 + 1,4$	1	2	0,9	0,1	14	$\frac{2x-1}{x^2}y + 1$	1	2	1	0,2
7	$\frac{\delta}{\delta + 1} + \delta(\delta + 1)$	0	1	1	0,2	15	$-2xy + xe^{-x^2} \sin x$	0	1	1	0,2
8	$\delta + 2,5\delta^2 + 2$	1	2	0,9	0,1	16	$x + \cos \frac{y}{2}$	-2	-1	3	0,1

5 Решение уравнений в частных производных

5.1 Построение аппроксимирующих разностных схем

Пусть функция $f(x)$ имеет достаточное число ограниченных производных. Разложим функцию в ряд Тейлора:

$$f(x + \Delta x) = f(x) + \Delta x f'(x) + \frac{(\Delta x)^2}{2!} f''(x) + \frac{(\Delta x)^3}{3!} f'''(x) + \dots \quad (5.1)$$

$$f(x - \Delta x) = f(x) - \Delta x f'(x) + \frac{(\Delta x)^2}{2!} f''(x) - \frac{(\Delta x)^3}{3!} f'''(x) + \dots \quad (5.2)$$

Выделив линейную часть ряда (5.1), получим

$$f'(x) \approx \frac{f(x + \Delta x) - f(x)}{\Delta x} \quad \text{- правая разность.}$$

Из линейной части ряда (5.2) следует

$$f'(x) \approx \frac{f(x) - f(x - \Delta x)}{\Delta x} \quad \text{- левая разность.}$$

Если из ряда (5.1) вычесть ряд (5.2) и затем выделить линейную часть, то получим

$$f'(x) \approx \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x} \quad \text{- центральная разность.}$$

Правая и левая разности аппроксимируют производную с первым порядком точности, центральная разность - со вторым порядком точности.

Аналогичным образом можно получить разности

$$f'(x) \approx \frac{f(x + \frac{\Delta x}{2}) - f(x - \frac{\Delta x}{2})}{\Delta x}$$

$$f''(x) \approx \frac{f(x + \Delta x) - 2f(x) + f(x - \Delta x))}{\Delta x^2} \quad \text{и т.д.}$$

5.2 Решение задачи Коши

Рассмотрим задачу Коши для уравнения в частных производных

$$\begin{cases} \frac{\partial y}{\partial t} - \frac{\partial y}{\partial x} = \varphi(x, t), & x \in [x_0, x_n], \quad t \in [0, t_k] \\ y(x, 0) = \psi(x), & x \in [x_0, x_n]. \end{cases} \quad (5.3)$$

Восстановим функцию y , используя разностные аппроксимации производных. Ввиду наличия нескольких комбинаций приближенных аппроксимирующих функций производных эту задачу можно решать несколькими вариантами. Мы рассмотрим лишь два варианта (две схемы).

Схема 1. Используем следующие аппроксимации:

$$\frac{\partial y}{\partial t} \approx \frac{y(x, t + \Delta t) - y(x, t)}{\Delta t}, \quad \frac{\partial y}{\partial x} \approx \frac{y(x + \Delta x) - y(x, t)}{\Delta x}.$$

В рассматриваемой области построим равномерную прямоугольную сетку

$$x_i = x_0 + i\Delta x, \quad i = \overline{0, n}; \quad t_j = 0 + j\Delta t, \quad j = \overline{0, k}.$$

Геометрическое представление взаимного расположения точек аппроксимирующих функций в узлах сетки называется *шаблоном*.

Так, введенным в схеме 1 для решения задачи (5.3) аппроксимирующим функциям соответствует шаблон, приведенный на рисунке 5.1, а.

Для простоты изложения будем значения функции $y(x_i, t_j)$ в узле (i, j) обозначать y_i^j - сеточная функция. Тогда имеем систему линейных уравнений относительно неизвестных y_i^{j+1} :

$$\begin{cases} \frac{y_i^{j+1} - y_i^j}{\Delta t} - \frac{y_{i+1}^j - y_i^j}{\Delta x} = \varphi(x_i, y_j), & i = \overline{0, n}, \quad j = \overline{0, k-1} \\ y_i^0 = \psi(x_i), & i = \overline{0, n} \end{cases} \quad (5.4)$$

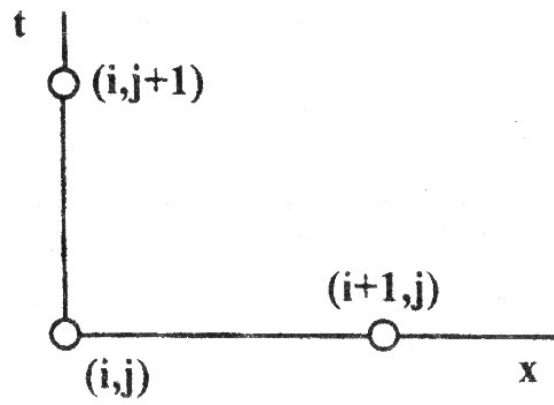
Эта система называется *явной схемой*. Решается она просто. На нулевом шаге известны значения функции в начальный момент времени (заданы начальные условия Ψ).

На первом шаге осуществляется расчет функции на слое $j=1$ (рисунок 5.2,а):

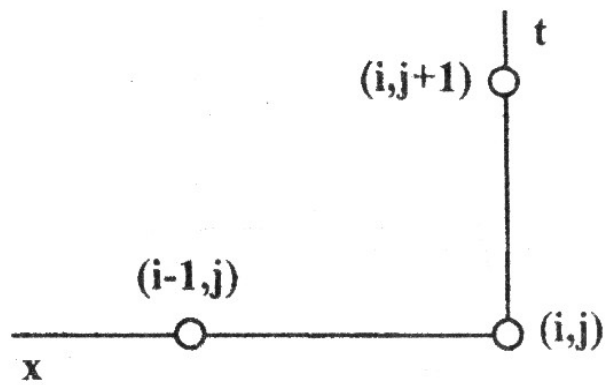
$$y_i^1 = y_i^0 + \Delta t \left[\frac{y_{i+1}^0 - y_i^0}{\Delta x} + \varphi(x_i, 0) \right], \quad i = \overline{0, n-1}. \quad (5.5)$$

Шаг должен завершиться расчетом y_n^1 . Однако для его расчета, в соответствии с уравнением (5.5), кроме значения y_n^0 , должно быть задано значение y_{n+1}^0 . Начальные же условия $y_i^0 = \psi(x_i)$ определены лишь в узлах $(i,0)$, $i = \overline{1, n}$.

Процесс определения y_{n+1}^j называется *конструированием граничных условий* (определение значений функции на граничном слое $i=n+1$). Граничные условия можно, в частности, конструировать экстраполяцией функции по известным ее значениям в узлах сетки. Наиболее часто используют линейную экстраполяцию или допущение, что на границе производная функции равна нулю.

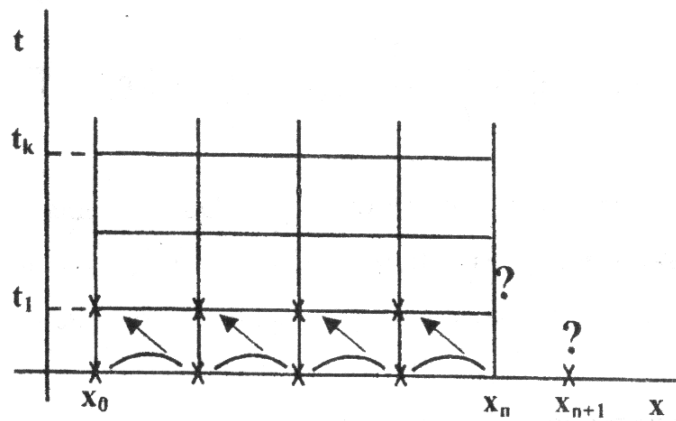


a)

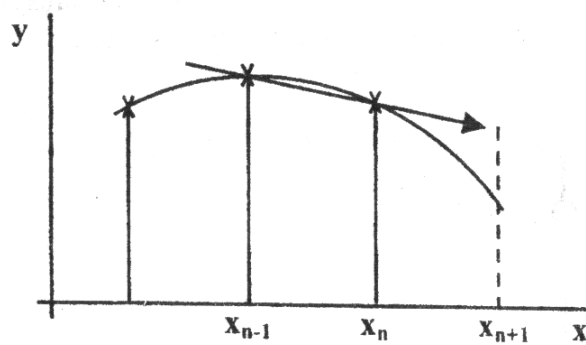


б)

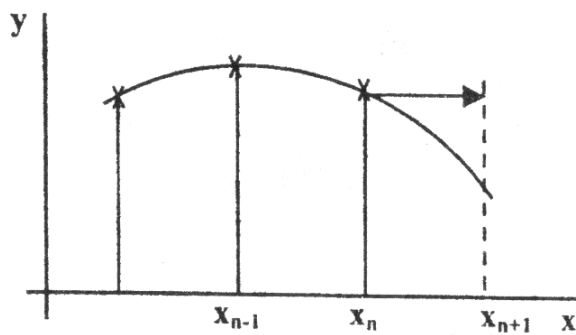
Рисунок 5.1 - Шаблоны решения задачи 5.3: а) - схема 1; б) - схема 2



а)



б)



в)

Рисунок 5.2 - Иллюстрация решения по схеме 1 и определения граничных условий: а) - иллюстрация решения на шаге 1; б), в) - иллюстрация конструирования граничных условий

Линейная экстраполяция (рисунок 5.2,б):

$$y_{n+1}^0 = y_n^0 + (y_n^0 - y_{n-1}^0).$$

Производная функции на границе равна нулю (рисунок 5.2,в):

$$\frac{y_{n+1}^0 - y_n^0}{\Delta x} = 0 \quad \Rightarrow \quad y_{n+1}^0 = y_n^0.$$

Итак, определив значение функции в узле $(n,1)$ на основе y_n^0 и y_{n+1}^0 , завершаем шаг 1 - расчет значений функции на слое $j=1$.

Расчет значений функции на втором шаге осуществляется по уравнениям

$$y_i^2 = y_i^1 + \Delta t \left[\frac{y_{i+1}^1 - y_i^1}{\Delta x} + \varphi(x_i, j\Delta t) \right], \quad i = \overline{0, n}$$

и т.д. по всем $j = \overline{0, k-1}$. При этом на каждом шаге проводится конструирование граничных условий.

После завершения шага $k-1$ определим приближенные значения функции во всех узлах рассматриваемой сетки. Точность вычисления зависит от типа используемых аппроксимирующих функций (типа шаблона) и от точности конструирования граничных условий.

Схема 2. Используем аппроксимации

$$\frac{\partial y}{\partial t} \approx \frac{y(x, t + \Delta t) - y(x, t)}{\Delta t}; \quad \frac{\partial y}{\partial x} \approx \frac{y(x, t) - y(x - \Delta x, t)}{\Delta x}.$$

Шаблон этой схемы приведен на рисунок 5.1,б.

Получим следующую систему линейных уравнений:

$$\begin{cases} \frac{y_i^{j+1} - y_i^j}{\Delta t} - \frac{y_j^i - y_{i-1}^j}{\Delta x} = \varphi(x_i, t_j), & i = \overline{0, n}, \quad j = \overline{0, k-1} \\ y_i^0 = \psi(x_i), & i = \overline{0, n}. \end{cases} \quad (5.6)$$

Эта схема, казалось бы, несущественно отличается от первой. Однако можно показать, что она не пригодна для счета - она не устойчива при любых

$$\frac{\Delta t}{\Delta x} = \text{const.}$$

5.3 Решение уравнения теплопроводности (диффузии)

$$\begin{cases} \frac{\partial y}{\partial t} - \frac{\partial^2 y}{\partial x^2} = \varphi(x, t), & x \in [x_0, x_n], \quad t \in [0, t_k] \\ y(x, 0) = \psi(x), & x \in [x_0, x_n]. \end{cases} \quad (5.7)$$

Рассмотрим две схемы решения задачи.

Схема 1. Используем аппроксимирующие функции производных, соответствующие шаблону, приведенному на рисунке 5.3,а:

$$\begin{cases} \frac{y_i^{j+1} - y_i^j}{\Delta t} - \frac{y_{i+1}^j - 2y_i^j + y_{i-1}^j}{\Delta x^2} = \varphi(x_i, t_j) \\ y_i^0 = \psi(x_i), \quad i = \overline{0, n}, \quad j = \overline{0, k-1}. \end{cases} \quad (5.8)$$

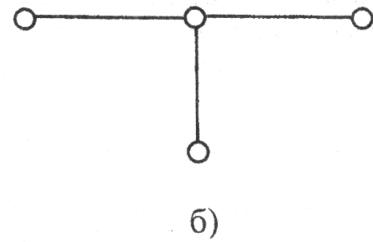
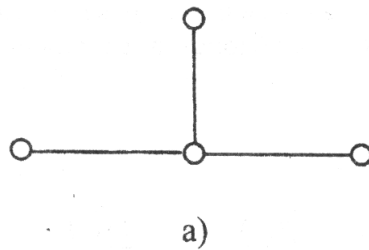


Рисунок 5.3 - Шаблоны решения задачи (5.7): а) - схема 1; б) - схема 2

Эта схема, как и схемы (5.4) и (5.6) решения задачи Коши, называется *явной схемой*. Действительно, из системы (5.8) можно получить уравнения для непосредственного (явного) вычисления неизвестных значений y_i^{j+1} :

$$y_i^{j+1} = y_i^j + \Delta t \left[\frac{y_{i+1}^j - 2y_i^j + y_{i-1}^j}{\Delta x^2} + \varphi(x_i, t_j) \right].$$

$$i = \overline{0, n}, \quad j = \overline{0, k-1}$$

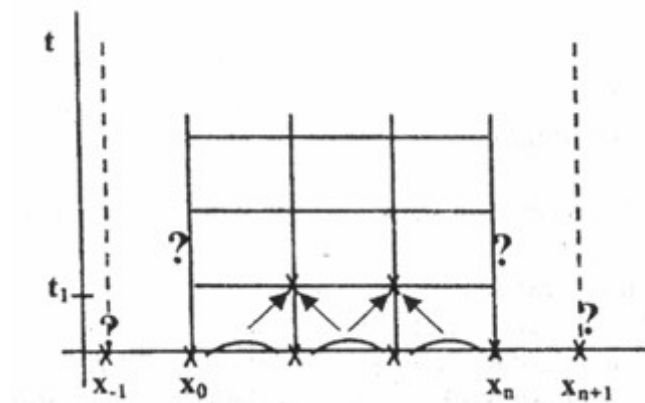


Рисунок 5.4 - Иллюстрация решения на шаге 1

В этой задаче граничные условия необходимо разрешать уже на двух слоях - на слое $i = -1$ и на слое $i = n+1$ (рисунок 5.4).

Схема 2. Для решения задачи используем шаблон, приведенный на рисунок 5.3,б. Тогда система уравнений будет иметь вид

$$\begin{cases} \frac{y_i^{j+1} - y_i^j}{\Delta t} - \frac{y_{i+1}^{j+1} - 2y_i^{j+1} + y_{i-1}^{j+1}}{\Delta x^2} = \varphi(x_i, t_j) \\ y_i^0 = \psi(x_i), \quad i = \overline{0, n}, \quad j = \overline{0, k-1}. \end{cases} \quad (5.9)$$

Данная схема называется *неявной схемой*. Ее нельзя разрешить явно относительно неизвестных y_i^{j+1} , т.к. на $j+1$ -м слое неизвестными являются сразу три значения функции. Для решения задачи выписывают систему уравнений и получают систему линейных уравнений относительно известных y_i^j . При этом матрица системы является трехдиагональной. Решается она с использованием метода прогонки.

Вопросы и упражнения для самоконтроля.

- 1) С каким порядком точности аппроксимируют производную правая, левая и центральная разность?
- 2) Что понимают под шаблоном?
- 3) Дайте понятие явной и неявной разностной схемы для уравнения теплопроводности.

Определить порядок аппроксимации метода

$$\begin{cases} -\frac{(y_{i+1} - 2y_i + y_{i-1}))}{h^2} + (ay_{i+1} + by_i + cy_{i-1}) = f_i + \frac{h^2}{12} f''(x_i), i = 1, \dots, n-1, \\ y_0 = y_n = 0 \end{cases}$$

решения краевой задачи $\begin{cases} -u'' + u = f(x), \\ u(0) = u(1) = 0. \end{cases}$

№	☐	☐	☐	№	☐	☐	☐
1	1/12	5/6	1/12	16	1	0	0
2	-1/12	1	1/12	17	1/2	1/2	0
3	0	3	-2	18	1/4	1/2	1/4
4	1	1	1	19	0	1/3	2/3
5	2/3	-1/3	2/3	20	1	-1	1
6	1	1/3	2/3	21	0	1/2	1/2
7	-1/12	2	-1/12	22	1/2	0	1/2
8	-1	1	1	23	0	-1/3	4/3

Список использованных источников

1. **Бахвалов, Н.С.** Численные методы [Текст]/ Н.П. Жидков, Г.М. Кобельков. - М.: Лаборатория Базовых Знаний, 2001. – 632 с.
2. **Березин, И.С.** Методы вычислений [Текст]: уч. пособие / И.С. Березин, Н.П. Жидков. - М. : Физматгиз, 1962. **Т.1.**- 464 с. **Т. 2.** - 639 с.
3. **Бояринов, А.И.** Методы оптимизации в химической технологии [Текст]: учебное пособие / А.И. Бояринов, В.В. Кафаров; под ред. В.В. Кафарова. - М. : Химия, 1969. - 564 с.
4. **Вычислительная техника в инженерных и экономических расчетах** [Текст]: учеб. для вузов / под ред. А. В. Петрова.- 2-е изд., перераб. и доп. - М. : Высш. школа, 1984. - 320 с.
5. **Волков, Е. А.** Численные методы [Текст] : учеб. пособие / Е.А. Волков. - М. : Наука, 1987. – 248 с.
6. **Гутер, Р.С.** Элементы численного анализ и математической обработки результатов опыта [Текст] / Р.С. Гутер, Б.В. Овчинский.- 2-е изд., перераб. - М. : Наука, 1970. - 432 с.
7. **Демидович, Б.П.** Основы вычислительной математики [Текст] / Б.П. Демидович, И.А. Марон. - М.: Наука, 1970. – 664 с .
8. **Заварыкин, В.М.** Численные методы [Текст]: учеб. пособие для вузов / В.М. Заварыкин, В.Г. Житомирский, М.П. Лапчик. - М.: Просвещение, 1991. – 176 с.
9. **Копченова, Н.В.** Вычислительная математика в примерах и задачах [Текст] / Н.В. Копченова, Марон И.А. - М.: Наука, 1972.- 368 с.
10. **Калиткин, Н. Н.** Численные методы [Текст] : учеб. пособие / Н. Н. Калиткин; под ред. А. А. Самарского. - М. : Наука, 1978. - 512 с.
11. **Мак-Кракен, Д. Д.** Численные методы и программирование на фортране [Текст] : пер. с англ. / Д. Мак-Кракен, У. Дорн.- 2-е изд., стер. - М.: Мир, 1977. - 584 с.
12. **Марчук, Г.И.** Методы вычислительной математики [Текст]: учеб. пособие для вузов по спец."прикл.математика".- 3-е изд., перераб. и доп. - М. : Наука, 1989. – 608 с.
13. **Ракитин, В.И.** Практическое руководство по методам вычислений [Текст] / В.И. Ракитин, В.Е. Первушин. - М.: Высшая школа, 1998.- 383 с.
14. **Турчак, Л.И.** Основы численных методов [Текст] / Л.И. Турчак -М.: Наука, 1987.- 320 с.
15. **Хемлинг, Р.В.** Численные методы [Текст]. Для научных работников и инженеров / Р.В. Хемлинг. -М.: Наука, 1968. – 330 с.
16. **Химмелблау, Д.** Прикладное нелинейное программирование [Текст] / Д. Химмелблау. -М.: Мир. 1975.- 534 с