

В. Д. Шелест  
М. С. Житомирский

# НАЧАЛА ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ

$$\sum_{i=1}^n x_i = \frac{a_1}{a_0}$$

$$\sum_{\substack{i,j \\ 1 \leq i < j \leq n}} x_i x_j = \frac{a_2}{a_0}$$

$$\sum_{\substack{i,j,k \\ 1 \leq i < j < k \leq n}} x_i x_j x_k = \frac{a_3}{a_0}$$

$$\prod_{i=1}^n x_i = (-1)^n \frac{a_n}{a_0}$$

Введение  
в численный  
эксперимент

Санкт-Петербург

Издательство Политехнического университета

2005



Федеральное агентство по образованию

---

САНКТ-ПЕТЕРБУРГСКИЙ  
ГОСУДАРСТВЕННЫЙ ПОЛИТЕХНИЧЕСКИЙ  
УНИВЕРСИТЕТ

---

В. Д. Шелест, М. С. Житомирский

НАЧАЛА  
ВЫЧИСЛИТЕЛЬНОЙ  
МАТЕМАТИКИ

Введение в численный эксперимент

Санкт-Петербург  
Издательство Политехнического университета  
2005

УДК 517.5

ББК 22.19

Ш 426

*Шелест В. Д., Житомирский М. С.* **Начала вычислительной математики.** Введение в численный эксперимент. СПб.: Изд-во СПб ГПУ, 2005. 202 с.

ISBN 5-7422-0069-2

Рассматриваются простые и универсальные способы уточнения корней уравнений с одним неизвестным, приближённого вычисления интегралов и решения систем линейных алгебраических уравнений. Цель книги — изложение методики и практики численного решения задач с помощью компьютера. Даны примеры, созданы надёжные, эффективные и универсальные алгоритмы (по этой причине их можно использовать в математических пакетах) и сформулированы упражнения.

Расчитана на читателей, знакомых с математикой в объёме средней школы — учащихся лицеев, гимназий, колледжей и техникумов. Рекомендуются также студентам младших курсов вузов, преподавателям и учителям, всем пользователям компьютеров.

Ил. 30. Библиогр.: 17 назв.

Печатается по решению редакционно-издательского совета Санкт-Петербургского государственного политехнического университета.

© Шелест В. Д., Житомирский М. С., 2005

© Санкт-Петербургский

государственный политехнический  
университет, 2005

ISBN 5-7422-0069-2

## ОГЛАВЛЕНИЕ

<b>Предисловие</b> .....	5
<b>ГЛАВА 1. Численное решение уравнений с одной переменной</b> .....	7
<b>1. Метод бисекций</b> .....	7
Постановка задачи .....	7
1.1 Нахождение границ интервала, содержащего все корни уравнения .....	10
1.2 Отделение корней уравнения .....	13
1.3 Приближённое значение корня и его оценка .....	22
1.4 Уточнение корня методом бисекций .....	25
<b>2. Метод простых итераций</b> .....	34
Постановка задачи .....	34
2.1 Метод итераций .....	35
2.2 Условие сходимости метода итераций .....	38
2.3 Оценка приближения в методе итераций .....	41
2.4 Выбор уравнения $x = \varphi(x)$ .....	44
<b>3. Метод пропорциональных частей</b> .....	54
3.1 Метод хорд .....	54
3.2 Сходимость метода хорд .....	57
3.3 Оценка погрешности в методе хорд .....	59
<b>4. Метод касательных</b> .....	66
4.1 Метод Ньютона .....	66
4.2 Сходимость метода Ньютона .....	70
4.3 Оценка погрешности в методе Ньютона .....	71
4.4 Модификация метода Ньютона .....	73
<b>ГЛАВА 2. Приближённое вычисление интегралов</b> .....	84
<b>1. Определённый интеграл</b> .....	84
1.1 Понятие интеграла .....	84
1.2 Определение интеграла .....	89
1.3 Точные методы вычисления интегралов .....	91
1.4 Интегрируемость функций .....	94

1.4.1	Суммы Дарбу и их свойства .....	94
1.4.2	Классы интегрируемых функций .....	98
<b>2.</b>	<b>Приближённые методы вычисления интегралов .....</b>	<b>102</b>
2.1	Аналитические методы .....	102
2.2	Численные методы .....	105
2.2.1	Формулы прямоугольников .....	105
2.2.2	Формула трапеций .....	108
2.2.3	Формула парабол .....	112
2.2.4	Практическая оценка погрешности квадратурных формул .....	118
2.2.5	Формулы Гаусса — Лежандра .....	124
<b>ГЛАВА 3.</b>	<b>Решение систем линейных алгебраических уравнений .....</b>	<b>130</b>
<b>1.</b>	<b>Элементарная алгебра матриц .....</b>	<b>130</b>
1.1	Матрицы и векторы .....	130
1.2	Основные действия над матрицами .....	133
1.3	Определители .....	137
<b>2.</b>	<b>Прямые методы .....</b>	<b>144</b>
	Постановка задачи .....	144
2.1	Правило Крамера .....	146
2.1.1	Два уравнения с двумя неизвестными .....	146
2.1.2	Два уравнения с тремя неизвестными .....	148
2.1.3	Три уравнения с тремя неизвестными .....	150
2.2	Метод Гаусса .....	153
<b>3.</b>	<b>Итеративные методы .....</b>	<b>169</b>
3.1	Метод Гаусса — Зейделя .....	169
3.2	Сходимость итеративного процесса .....	172
3.3	Оценка погрешности метода Зейделя .....	174
3.4	Приведение линейной системы к виду, пригодному для итераций .....	175
	<b>Приложение .....</b>	<b>187</b>
	<b>Предметный указатель .....</b>	<b>196</b>
	<b>Список литературы .....</b>	<b>201</b>

## Предисловие

Вычислительная математика — ветвь математики, изучающая методы поиска решения математических задач в виде числового (точного или приближённого) значения. Её истоки лежат в глубокой древности. Началом вычислительной математики можно считать правила вычисления иррациональных чисел.

Современная вычислительная математика состоит из многих разделов; важнейшие из них — вычисление значений функций, численное решение конечных (алгебраических и трансцендентных) уравнений, приближение функций, численное дифференцирование и интегрирование, численное решение дифференциальных и интегро-дифференциальных уравнений, численные методы линейной алгебры, методы оптимизации, методы дискретного анализа. Быстрому развитию вычислительной математики способствовало создание компьютеров. Возникли новые её направления, такие как численные методы теории игр, теории массового обслуживания, минимизации логических функций, комбинаторики, методы решения обратных задач и др.

Численный результат можно получить только посредством арифметических и логических действий, поэтому задачу вычислительной математики формулируют как задачу представления решений (точно или приближённо) в виде последовательности арифметических операций. Значит, численный метод, кроме описания последовательности арифметических действий, должен содержать и оценку погрешности. Точного решения можно добиться при конечном количестве арифметических операций лишь в редких случаях. Чаще результат является пределом бесконечной последовательности операций, и тогда возникают проблемы сходимости метода. При этом интересно не только то,

что метод сходится, но и то, как быстро он это делает, ибо при практической реализации метода важна его эффективность. Её оценивают априори<sup>1</sup> — по количеству операций, необходимых для нахождения результата, и апостериори — по затраченному машинному времени для получения ответа и объёму требуемой при этом памяти компьютера. С эффективностью связана и сложность метода. Параллельно с получением ответа нередко приходится решать вопросы устойчивости результата относительно возмущений исходных данных и округлений при вычислениях.

В заключение остановимся кратко на содержании книги. Она состоит из трёх глав. В *первой её главе* решена задача определения корней конечных (по-другому, скалярных, числовых) уравнений для случая функции одной переменной. В основе изложенных методов лежит замена функции в окрестности корня близкой к ней кривой (например, прямой). *Вторая глава* посвящена определённому интегралу и численному интегрированию, а *третья* — содержит описание методов решения систем линейных алгебраических уравнений. Все методы доведены до численных приложений — даны расчётные схемы и числовые примеры с подробным ходом решения. Разработаны алгоритмы для решения задач на компьютере. Их надёжности, легкочитаемости, эффективности и универсальности уделено особое внимание. Сформулированы упражнения. Они удобны для самостоятельной работы и при проведении лабораторного практикума.

Книга имеет общую цель — познакомить читателя с применением численных методов, подготовить его к решению задач прикладной математики<sup>2</sup>. В добрый путь ...

---

<sup>1</sup> Лат. *a priori* — до опыта, *a posteriori* — из опыта.

<sup>2</sup> Скажем, к коррекции или созданию математических пакетов (например, алгоритмы Calculus Iterate и Approximate Integral меню Maple SCIENTIFIC WORK PLACE неудовлетворительны), к выполнению вычислительного эксперимента и математического моделирования.

Всё есть число.  
Пифагор

## ГЛАВА I. Численное решение уравнений с одной переменной

### 1. Метод бисекций

#### Постановка задачи

Уравнения сыграли важную роль в развитии математики. Они и сегодня вызывают интерес, поскольку встречаются в теоретических и прикладных задачах и, что не менее важно, служат отправным пунктом в изучении таких понятий вычислительной математики, как итеративный процесс, сходимость и оценка приближения.

Методы решения уравнений делятся на *прямые* и *итеративные*, т. е. *методы последовательных приближений*. Прямые методы позволяют представить корни в виде некоторых конечных соотношений. Известны формулы, по которым можно найти решения алгебраических уравнений до четвёртой степени включительно. Уже в XVI веке итальянский математик Кардано опубликовал формулу для решения кубического уравнения<sup>3</sup>

$$t^3 + pt + q = 0,$$

которая имеет вид

$$t = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}.$$

<sup>3</sup>Уравнение  $a_0x^3 + a_1x^2 + a_2x + a_3 = 0$  ( $a_0 \neq 0$ ) приводится к такому виду посредством подстановки  $x = t - \frac{a_1}{3a_0}$ .

Под знаком кубических корней могут оказаться мнимые числа, значит, надо уметь из них извлекать корни<sup>4</sup>. Упомянем здесь также о Н. Х. Абеле и Э. Галуа, доказавших, что общее уравнение степени, выше четвёртой, неразрешимо в радикалах, т. е. нельзя найти корни таких уравнений через коэффициенты уравнения посредством арифметических действий и извлечения корней. В случае неалгебраических уравнений ситуация совсем не простая: решить удаётся лишь некоторые из трансцендентных (показательные, логарифмические, тригонометрические) уравнений.

В условиях, когда формулы “не работают”, важное значение приобретают вычислительные алгоритмы, с помощью которых можно найти корни уравнения с любой наперёд заданной точностью. К этим случаям отнесём и те, в которых хотя и существуют формулы для получения точного решения (формула Кардано), но они столь громоздки, что пользование численными методами предпочтительнее. Не следует думать, что решения, полученные такими методами, в чём-либо уступают точным. Пусть, например, найдены точные корни некоторого уравнения, равные, скажем,  $(1 \pm \sqrt{3})/2$ . “Чистого” математика такой результат, очевидно, удовлетворит; “прикладному” же математику, чтобы воспользоваться этим решением, придётся заменить  $\sqrt{3}$  его приближённым значением с требуемой в задаче точностью.

Здесь, наверное, уместен и следующий пример из линейной алгебры. Решить систему  $n$  линейных алгебраических уравнений с  $n$  неизвестными можно по правилу Крамера или методом последовательного исключения неизвестных (методом Гаусса). С точки зрения “чистого” математика эта задача является тривиальной. Однако при практической реализации названных методов могут возникнуть “подводные рифы”. Так для решения системы по правилу Крамера нужно выполнить  $n \cdot (n + 1)! - 1$  арифметических операций, что

---

<sup>4</sup>В Приложении решения уравнений третьей и четвёртой степеней приведены полностью.

при  $n$ , равном, например, 100, составит приблизительно  $10^{160}$  действий. Решая эту задачу, компьютер с быстродействием  $10^{10}$  операций в секунду будет непрерывно трудиться много лет (заметим, что факториальный рост количества арифметических операций с увеличением размерности задачи называют “проклятием размерности”). Метод Гаусса более эффективен. Для его реализации нужно проделать  $2n(n+1)(n+2)/3 + n(n-1) \Big|_{n=100} = 696700$  операций, а значит, затратить меньше секунды времени процессора. Но этот метод рекомендуют для систем, порядок которых не превышает 200. Дело в том, что ошибки, обусловленные ограниченностью разрядной сетки компьютера (“волчьи ямы округления”), с увеличением количества шагов оказывают такое влияние на конечный результат, что он бывает сильно искажён. Чтобы исправить положение, применяют итеративные методы (при этом ошибки не накапливаются), т. е. отдаётся предпочтение методам последовательных приближений, а не точным методам. В вычислительной математике такой приём типичен.

Мы будем рассматривать уравнение вида

$$f(x) = 0, \quad (1)$$

где  $f(x)$  — функция действительного аргумента, определённая и непрерывная в некотором конечном или бесконечном интервале  $(a; b)$ .

Напомним, что число  $\xi$  из области определения функции называется *корнем* уравнения (1), если  $f(\xi) = 0$ . Наша задача состоит в поиске корней этого уравнения.

Поиск будем выполнять в несколько этапов:

**1 этап**, где определим границы интервала, в котором находятся все корни уравнения (1);

**2 этап** — отделим корни, т. е. установим возможно малые промежутки, в каждом из которых содержится ровно один корень;

**3 этап** — уточним корни, т. е. вычислим их с заданной точностью;

**4 этап** — проведём анализ полученных результатов.

Подробно первые три этапа будут рассмотрены в разделах 1.1 – 1.3 этой главы, а в разделе 1.4 дан конкретный метод решения уравнения (1).

### 1.1 Нахождение границ интервала, содержащего все корни уравнения

К сожалению для уравнения (1) не существует общего способа выполнения первого этапа. В общей форме определение границ корней можно реализовать лишь для алгебраического уравнения в *канонической форме*:

$$a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n = 0, \quad a_0 \neq 0. \quad (2)$$

Прежде чем приступить к доказательству теоремы о границах корней, напомним некоторые факты из алгебры.

**ТЕОРЕМА 1. (Основная теорема алгебры.)** *Уравнение вида (2) (а следовательно, и полином  $P_n(x) = \sum_{i=0}^n a_i x^{n-i}$ ) имеет ровно  $n$  корней, действительных или комплексных, если  $k$ -кратный корень считать за  $k$  корней.*

*Число действительных корней уравнения (2) чётной степени с действительными коэффициентами всегда чётно (например, равно нулю). Если в уравнении чётной степени  $a_n/a_0 < 0$ , то такое уравнение имеет, по крайней мере, два действительных корня разного знака. Уравнение нечётной степени имеет хотя бы один действительный корень того же знака, что и  $(-a_n/a_0)$ . Число положительных корней уравнения (2) не больше числа перемен знаков в ряду коэффициентов полинома  $P_n(x)$  и может отличаться от него на чётное число (правило Декарта). Так как при замене  $x$  на  $-x$  корни уравнения меняют знаки, то с помощью этого правила можно оценить и число отрицательных корней.*

Если ввести обозначения

$$\tilde{a} = \max_{0 \leq i \leq n-1} |a_i|, \quad \tilde{a} = \max_{1 \leq i \leq n} |a_i|,$$

то границы промежутков, в которых расположены все корни уравнения (2), можно найти с помощью следующей теоремы.

**ТЕОРЕМА 2. (Правило кольца.)** *Корни уравнения (2) удовлетворяют двойному неравенству*

$$\frac{|a_n|}{\tilde{a} + |a_n|} \leq |x| \leq 1 + \frac{\tilde{a}}{|a_0|}.$$

*Доказательство.* Сначала найдём промежутки, где нет корней уравнения (2), т. е. решим неравенство  $|f(x)| > 0$ . Воспользовавшись свойствами модуля

$$|u + v| \geq |u| - |v|, \quad \left| \sum_{i=1}^k u_i \right| \leq \sum_{i=1}^k |u_i|, \quad \left| \prod_{i=1}^k u_i \right| = \prod_{i=1}^k |u_i|$$

и формулой

$$u^k - 1 = (u - 1)(u^{k-1} + u^{k-2} + \dots + u + 1),$$

выпишем очевидную цепочку неравенств:

$$\begin{aligned} |f(x)| &\geq |a_0 x^n| - |a_1 x^{n-1} + \dots + a_{n-1} x + a_n| \geq \\ &\geq |a_0 x^n| - (|a_1 x^{n-1}| + \dots + |a_{n-1} x| + |a_n|) = \\ &= |a_0| |x|^n - (|a_1| |x|^{n-1} + \dots + |a_{n-1}| |x| + |a_n|) \geq \\ &\geq |a_0| |x|^n - \tilde{a} (|x|^{n-1} + \dots + |x| + 1) = \\ &= |a_0| |x|^n - \tilde{a} (|x|^n - 1) / (|x| - 1). \end{aligned}$$

Полагая, что  $|x| > 1$ , усилим неравенство, отбросив единицу в числителе дроби. В результате получим

$$|f(x)| \geq |a_0| |x|^n - \tilde{a} |x|^n / (|x| - 1).$$

Следовательно, неравенство  $|f(x)| > 0$  будет выполняться в случае, если

$$|a_0| |x|^n - \tilde{a} |x|^n / (|x| - 1) > 0$$

или

$$|a_0||x| - |a_0| - \tilde{a} > 0.$$

Отсюда находим, что значения  $x$ , для которых  $f(x) \neq 0$ , удовлетворяют неравенству

$$|x| > 1 + \tilde{a}/|a_0|.$$

Значит, все корни уравнения (2) удовлетворяют противоположному неравенству

$$|x| \leq 1 + \tilde{a}/|a_0|.$$

Верхняя граница модулей корней определена. Чтобы найти нижнюю границу, выполним в (2) замену  $x$  на  $1/t$ , после чего получим

$$|t| \leq 1 + \tilde{a}/|a_n|$$

или

$$|x| \geq |a_n|/(\tilde{a} + |a_n|).$$

Теорема доказана.  $\square$

*Замечание 1.* В общем случае переменную  $x$  можно считать комплексной, т. е.  $x = \xi + i\eta$  и тогда  $|x| = \sqrt{\xi^2 + \eta^2}$ . На комплексной плоскости  $\xi O\eta$  корни уравнения (2) расположены в круговом кольце  $r \leq |x| \leq R$ , где

$$r = |a_n|/(\tilde{a} + |a_n|), \quad R = 1 + \tilde{a}/|a_0|.$$

Если  $x \in \mathbf{R}$ , то числа  $r$  и  $R$  являются соответственно нижней и верхней границами положительных корней, а числа  $-R$  и  $-r$  — нижней и верхней границами отрицательных корней уравнения (2).

**ПРИМЕР 1.** Определим область, в которой находятся корни уравнения  $x^3 - 2x - 3 = 0$ .

*Решение.* Здесь  $a_0 = 1$ ,  $a_1 = 0$ ,  $a_2 = -2$ ,  $a_3 = -3$ . Тогда  $\tilde{a} = 2$ ,  $\tilde{a} = 3$ . Согласно теореме 2

$$3/(2+3) \leq |x| \leq 1+3/1, \quad \text{или} \quad |x| \in [0.6; 4]. \quad \triangle$$

Если  $f(x)$  — дробно-рациональная, иррациональная или трансцендентная функция, то важно указать её область определения или, другими словами, ОДЗ — область допустимых значений  $x$ , что нередко позволяет полностью решить первый этап. Установление области определения уравнения важно ещё и потому, что решение уравнения часто сводят к решению неэквивалентного уравнения. В приложениях иногда можно сузить границы корней, так как, например, в трансцендентных уравнениях неизвестные суть частоты колебаний, длины, напряжения, токи, давления, температуры и т. п., а их значения лежат в определённых пределах. Более того, из физического смысла задачи, как правило, известно грубое, приближённое решение, так называемое, *нулевое, начальное приближение корня*. Так что здесь, как говорится, все средства хороши.

**Упражнение 1.** Найдите границы корней следующих уравнений:

- 1)  $-2x^5 + 7.2x^4 - 0.25x + 10 = 0$ ;
- 2)  $0.001p^3 + 0.2p^2 + 3.25p + 43.5 = 0$ ;
- 3)  $x^6 - x^4 + x^2/5 - 1/105 = 0$ ;
- 4)  $x^9 - 3x^7/2 + 27x^5/40 - 57x^3/560 + 2217x/22400 = 0$ ;
- 5)  $x^n \pm x^{n-1} + 2x^{n-2} \pm \dots + (-1)^{n-1}(n-1)x + (-1)^n n = 0$ ;
- 6)  $nx^n \pm (n-1)x^{n-1} + \dots + (-1)^{n-1}x + (-1)^n = 0$ .

## 1.2 Отделение корней уравнения

Перейдём ко второму этапу — разбиению области допустимых значений на отрезки, в каждом из которых содержится один и только один корень. Отделение (изоляция, локализацию) корней можно выполнить двумя способами — *аналитическим* и *графическим*.

**Аналитический способ.** Аналитически корни уравнения (1) можно отделить, используя некоторые свойства функции  $f(x)$ . Здесь будет полезна следующая теорема о существовании корня непрерывной функции.

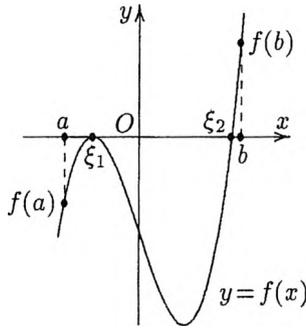


Рис. 1.

**ТЕОРЕМА 3. (Первая теорема Больцано-Коши.)** Если функция  $f(x)$  определена и непрерывна на отрезке  $[a; b]$  и принимает на его концах значения разных знаков, то на  $[a; b]$  содержится, по меньшей мере, один корень уравнения (1).

Геометрически это значит, что если точки графика непрерывной функции находятся по разные стороны от оси  $Ox$ , то кривая пересекает ось абсцисс хотя бы один раз в точке  $x = \xi_2$  (рис. 1).

Требование непрерывности функции  $f(x)$  важно. На рис. 2 приведён график разрывной функции, которая при-

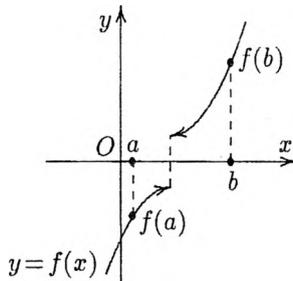


Рис. 2.

нимает на концах отрезка  $[a; b]$  значения разных знаков, но не имеет корней.

Теорема 3 не обеспечивает, однако, единственности корня на отрезке  $[a; b]$ , поэтому наложим на функцию  $f(x)$  до-

полнительные ограничения. Эти ограничения выражаются следующей теоремой о существовании корня.

**ТЕОРЕМА 4.** *Корень  $\xi$  будет единственным, если при выполнении условий теоремы 3 производная  $f'(x)$  существует и сохраняет постоянный знак внутри  $[a; b]$ .*

*Замечание 2.* Теоремы 3 и 4 можно сформулировать несколько в иной форме. Вот формальная запись этих теорем [16]:

**ТЕОРЕМА 3.**

$$\left. \begin{array}{l} f \text{ непр. на } [a; b] \\ f(a)f(b) < 0 \end{array} \right\} \Rightarrow \exists \xi \in (a; b): f(\xi) = 0.$$

**ТЕОРЕМА 4.**

$$\left. \begin{array}{l} f \text{ непр. на } [a; b] \\ f(a)f(b) < 0 \\ \text{sign } f'(x) = \text{const } \forall x \in [a; b] \end{array} \right\} \Rightarrow \exists \xi \in (a; b): f(\xi) = 0.$$

Знакопостоянство  $f'(x)$  означает монотонность функции  $f(x)$  на  $[a; b]$ , и если  $f(a)f(b) < 0$ , то кривая  $y = f(x)$  пересекает ось  $Ox$  внутри  $[a; b]$  в единственной точке (рис. 3). Если же непрерывная на отрезке функция  $f(x)$  монотонна ( $f'(x)$  сохраняет постоянный знак), а для концов отрезка истинно неравенство  $f(a)f(b) > 0$ , то на  $[a; b]$  нет ни одного корня уравнения (1). Из сказанного следует: чтобы отделить корни уравнения (1), сначала надо найти все интервалы монотонности функции  $f(x)$ , а затем из найденных интервалов оставить те, на границах которых  $f(x)$  принимает значения разных знаков. Границами интервалов монотонности непрерывной функции являются точки, в которых производная  $f'(x)$  меняет знак (исключая точки перегиба), т. е. корни уравнения

$$f'(x) = 0. \quad (3)$$

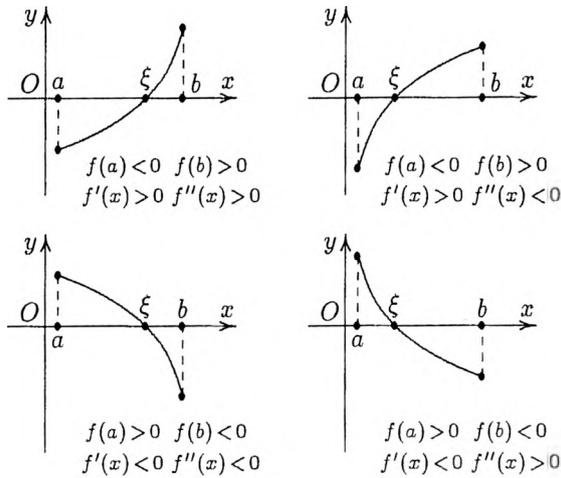


Рис. 3.

Иногда уравнение (3) проще исходного уравнения (1) и задачу отделения корней удаётся решить до конца. Так, например, если уравнение (1) алгебраическое, то уравнение (3) имеет порядок на единицу меньше порядка исходного уравнения (дано уравнение  $P_3(x) = 0$ ; находим корни полинома второй степени:  $P'_3(x) \equiv Q_2(x) = 0$ )<sup>5</sup>.

**ПРИМЕР 2.** Отделим корни уравнения  $x^3 - 2x - 3 = 0$ .

*Решение.* Функция  $f(x) = x^3 - 2x - 3$  и её производная  $f'(x) = 3x^2 - 2$  непрерывны на всей действительной оси. Определим интервалы монотонности функции. Для этого решим уравнение

$$f'(x) = 3x^2 - 2 = 0.$$

Его корни  $x_1 = -\sqrt{2/3}$  и  $x_2 = \sqrt{2/3}$ . Значит, функция монотонна на интервалах  $(-\infty; -\sqrt{2/3})$ ,  $(-\sqrt{2/3}; \sqrt{2/3})$  и

---

<sup>5</sup>  $P_3(x) \equiv \frac{x^3}{3} + \frac{x^2}{2} + x + 1 = 0$ ,  $\xi_1 = -1.153653$ ,  $(\xi_2, \xi_3) \in \mathbb{C}$ ;  
 $Q_2(x) \equiv x^2 + x + 1 = 0$ ,  $d = -3 < 0$ . Перефразируем Конфуция: трудно найти действительный корень в бесконечном интервале, особенно, когда его там нет.

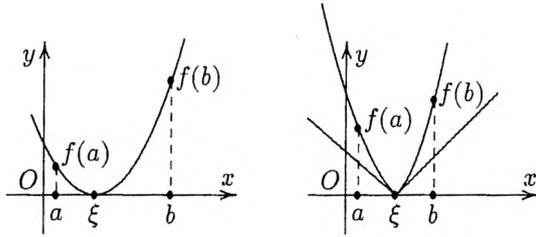


Рис. 4.

$(\sqrt{2/3}; \infty)$  <sup>6</sup>. При этом

$$\begin{aligned} \lim_{x \rightarrow -\infty} f(x) = -\infty < 0, & \quad f(-\sqrt{2/3}) < 0, \\ f(\sqrt{2/3}) < 0, & \quad \lim_{x \rightarrow \infty} f(x) = \infty > 0. \end{aligned}$$

Интервалы монотонности  $(-\infty; -\sqrt{2/3})$  и  $(-\sqrt{2/3}; \sqrt{2/3})$  не содержат корней уравнения, так как функция не меняет знак на их концах. Единственный действительный корень расположен между  $\sqrt{2/3}$  и  $\infty$  (причём  $\sqrt{2/3} \approx 0.817$ ).  $\triangle$

*Замечание 3.* На отрезке  $[a; b]$  может быть корень  $\xi$  уравнения (1) даже в случае, если функция  $f(x)$  на концах отрезка имеет значения одинаковых знаков (рис. 4). В этом случае, очевидно, либо  $f'(\xi) = 0$ , либо  $f(x)$  в точке  $\xi$  не дифференцируема. В обеих ситуациях в точке  $x = \xi$  функция  $f(x)$  достигает экстремума, и следовательно, такой корень является границей интервала монотонности. Корни этого типа в дальнейшем рассматривать не будем.

**Графический способ.** Если функция  $f(x)$  такова, что её график можно построить сравнительно быстро и просто, то для локализации корней уравнения (1) применим графический способ. По графику функции легко судить о том,

<sup>6</sup>На самом деле наша функция монотонна на промежутках  $(-\infty; \sqrt{2/3}]$ ,  $[-\sqrt{2/3}; \sqrt{2/3}]$  и  $[\sqrt{2/3}; \infty)$ . Это следует из её непрерывности.

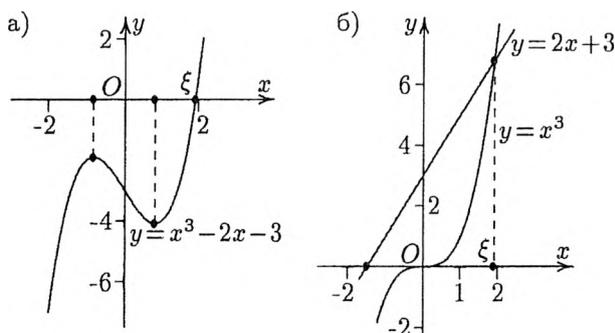


Рис. 5.

где находятся точки его пересечения с осью  $Ox$  и, следовательно, определить отрезки, содержащие изолированные корни уравнения (1).

Если график функции  $y = f(x)$  построить трудно, то выясним, нельзя ли преобразовать уравнение (1) в эквивалентное уравнение

$$\varphi_1(x) = \varphi_2(x), \quad (4)$$

такое, что графики функций  $y = \varphi_1(x)$  и  $y = \varphi_2(x)$  строятся просто. В этом случае корни уравнения (4) суть абсциссы точек пересечения графиков этих функций. В силу эквивалентности полученного уравнения исходному эти корни будут корнями уравнения (1).

**ПРИМЕР 3.** Отделим корни уравнения  $x^3 - 2x - 3 = 0$ .

*Решение.* Построим график функции  $y = x^3 - 2x - 3$  (рис. 5, а) и определим абсциссы точек пересечения этого графика с осью  $Ox$ . Кривая пересекает ось  $Ox$  в одной точке, значит  $\xi \in [1.4; 2.4]$ . Упростим решение. Заменим уравнение  $x^3 - 2x - 3 = 0$  эквивалентным уравнением  $x^3 = 2x + 3$  (рис. 5, б). Строим графики функций  $\varphi_1(x) = x^3$  и  $\varphi_2(x) = 2x + 3$ . Корень  $\xi$  есть точка пересечения кубической параболы и прямой. Он лежит на отрезке  $[1.4; 2.4]$ .

△

Если построение графиков вызывает затруднения (выражение  $f(x)$ ,  $\varphi_1(x)$  или  $\varphi_2(x)$  громоздкое), то сформируем таблицу значений функции с помощью компьютера и по полученным данным отделим корни.

**ПРИМЕР 4.** Отделим корни уравнения  $x^3 - 2x - 3 = 0$  с помощью компьютера.

*Решение.* Сначала составим функциональную схему [17]. Работу по отделению корней будут выполнять пять алгоритмов — головной с именем **ОТДЕЛЕНИЕ КОРНЕЙ** и подчинённые с именами **ГРАНИЦЫ**,  $f$ ,  $f'$ ,  $f''$ . Подчинённые алгоритмы-функции вычисляют значения  $f(x)$ ,  $f'(x)$  и  $f''(x)$ . На основании теоремы 2 один или три действительных корня уравнения  $x^3 - 2x - 3 = 0$  принадлежат отрезкам  $[-4; -0.6]$  или  $[0.6; 4]$ . Будем считать  $a$ , равное  $-4.6$ , левой границей, а  $b$ , равное  $4$ , — правой границей корней. Поиск суженных отрезков, содержащих по одному корню, выполнен *методом равномерного перебора*.

Выберем достаточно малый шаг  $h$ , такой, чтобы не пропустить корней (пусть, например,  $h=1$ ,  $h=0.1$ ). Организуем цикл повторений, в котором в точках  $x_i$ , отстоящих друг от друга на расстоянии  $h$ , определим знаки функций  $f(x_i - h)$  и  $f(x_i)$ . Если они совпадут, то это значит, что на текущем отрезке  $[x_i - h; x_i]$  нет корня уравнения; в противном случае корень наличествует (тогда выведем границы отрезка  $[x_{i-1}; x_i]$  и производные  $f'(x_{i-1})$ ,  $f'(x_i)$ ,  $f''(x_{i-1})$ ,  $f''(x_i)$ ). Как только  $x_i$  превзойдёт  $b$ , проанализируем состояние флажка — значение логической переменной *есть*. Она в случае отсутствия (действительных) корней позволяет вывести информацию об этом.

Приведём проект программы отделения корней.

```

{-----}
алг ОТДЕЛЕНИЕ КОРНЕЙ
    вещ  $a, b, h$ 
начало

```

вывести ' Введите границы  $a$ ,  $b$  и шаг  $h \implies$ '

ввести  $a$ ,  $b$ ,  $h$

вывести ' Ответ:'

вызвать ГРАНИЦЫ( $a$ ,  $b$ ,  $h$ ,  $f$ ,  $f'$ ,  $f''$ )

вывести ' Всё ... '

конец { головного алгоритма }  
 {-----}

алг вещь  $f$ (вещ  $x$ )

знач  $x$

начало

$f \leftarrow (x^2 - 2)x - 3$

возвратиться

конец { алг  $f(x)$  }  
 {-----}

алг вещь  $f'$ (вещ  $x$ )

знач  $x$

начало

$f' \leftarrow 3x^2 - 2$

возвратиться

конец { алг  $f'(x)$  }  
 {-----}

алг вещь  $f''$ (вещ  $x$ )

знач  $x$

начало

$f'' \leftarrow 6x$

возвратиться

конец { алг  $f''(x)$  }  
 {-----}

алг ГРАНИЦЫ(вещ  $a$ ,  $b$ ,  $h$ , алг вещь  $f$ (вещ  $x$ ),  $f'$ (вещ  $x$ ),  
 $f''$ (вещ  $x$ ))

знач  $a$ ,  $b$ ,  $h$

вещ  $x$ ,  $xh$  лог *знак\_слева*, *знак\_x*, *есть*

цел  $n$ ,  $i$

начало

*знак\_слева*  $\leftarrow (f(a) < 0)$  *есть*  $\leftarrow$  *ложь*

$$n \leftarrow [(b - a)/h + 0.5]$$

**вывести** ' Границы отрезков, содержащих корни,  
и производные '

для  $i$  от 1 до  $n$  шаг 1 **выполнять**

$$x \leftarrow a + i \cdot h \quad \text{знак}_x \leftarrow (f(x) < 0)$$

**если**  $\text{знак\_слева}$  не экв  $\text{знак}_x$

**то**

$$xh \leftarrow x - h$$

**вывести** ' левая=',  $xh$ , ' правая=',  $x$ ,  
'  $f$  ( $\lambda$ )=',  $f(xh)$ , '  $f$  ( $\mu$ )=',  $f(x)$ ,  
'  $f'$  ( $\lambda$ )=',  $f'(xh)$ , '  $f'$  ( $\mu$ )=',  $f'(x)$ ,  
'  $f''$  ( $\lambda$ )=',  $f''(xh)$ , '  $f''$  ( $\mu$ )=',  $f''(x)$

$\text{знак\_слева} \leftarrow \text{знак}_x$  **есть**  $\leftarrow$  **истина**

**конец-если** { знаки не совпали: есть корень }

**конец-цикл** { перебора значений  $x$  }

**если не есть**

**то**

**вывести** ' На  $[a; b]$  нет действительных корней',  
' или число действ. корней чётно ( $h$  велик).'

**конец-если** { нет корней }

**возвратиться**

**конец** { алг ГРАНИЦЫ }

---

Алгоритмы были закодированы Турбо-Паскалем версии 7.0 и выполнены на IBM PC. Вот результаты работы: единственный действительный корень уравнения принадлежит отрезку  $[1.4; 2.4]$ ,  $f(1.4) = -3.056$ ,  $f(2.4) = 6.024$ ,  $f'(1.4) = 3.88$ ,  $f'(2.4) = 15.28$ ,  $f''(1.4) = 8.4$ ,  $f''(2.4) = 14.4$ .

△

**Упражнение 2.1.** Отделите аналитически корни уравнений:

1)  $x^4 - 4x^3 + 4x^2 - 4 = 0$ ;

3)  $x \sin x - 1 = 0$ ;

2)  $x^5 - 3x^3 + 6x - 3 = 0$ ;

4)  $x^x = 1$ .

2. Отделите графически корни уравнений<sup>7</sup>:

1)  $1 - 2 \cos x/x = 0$ ;

2)  $x \cos x + 2 = 0$ ;

3)  $\sqrt{2-x^2} - \ln x^2 = 0$ ;

4)  $100 \sin \pi x - x = 0$ ;

5)  $100^{\sin x} - x^2 = 0$ ;

6)  $4 \sin x - 5x^2 - 2x + 3 = 0$ ;

7)  $10^x - \sin x = 0$ ;

8)  $2^{-|x-3|} - |x^2 - x - 6| + 1 = 0$ ;

9)  $\operatorname{ch} kx \cos kx - 1 = 0, \quad k = 1, 2, 3$ ;

10)  $\operatorname{ch} kx \cos kx + 1 = 0, \quad k = 1, 2, 3$ ;

11)  $\operatorname{tg} kx - \operatorname{th} kx = 0, \quad k = 1, 2, 3$ ;

12)  $\left(\sqrt{2+\sqrt{2}}\right)^x + \left(\sqrt{2-\sqrt{2}}\right)^x = 2^x$ ;

13)  $x^x - 1 = 0$ .

3. Отделите корни уравнений с помощью компьютера:

1)  $0.0004p^3 + 0.0008p^2 + 0.104p + 1 = 0$ ;

2)  $10^{-5}p^3 + 17 \cdot 10^{-4}p^2 + 8 \cdot 10^{-2}p + 0.5 = 0$ ;

3)  $0.004p^4 - 0.35p^3 + 4.2p^2 - 30p + 331 = 0$ ;

4)  $x^7 - 7x^5/6 + 119x^3/360 - 149x/6480 = 0$ ;

5)  $\arcsin(\lambda/x) - 2\sqrt{1-(\lambda/x)^2} + 0.047x - 0.094\lambda - 1.844 = 0$ ,  
 $\lambda = 0, 1, 2, \dots, 10$ .

4. Алгоритм ГРАНИЦЫ модифицируйте так, чтобы он в дополнение к границам корней и производным выводил ещё и границы отрезков, для которых выполняется неравенство  $|f(x)| \leq \delta$ , где  $\delta = 10^{-p}$  ( $p = 3, 4, \dots, 7$ ).

Перейдём теперь к оценке результатов отделения корня.

### 1.3 Приближённое значение корня и его оценка

Пусть тем или иным способом удалось отделить корни уравнения (1). Искомый корень  $\xi$  находится на отрезке  $[a; b]$ . В роли приближённого значения корня  $\xi$  можно

---

<sup>7</sup>Синус гиперболический  $\operatorname{sh} x = \frac{e^x - e^{-x}}{2}$ , косинус гиперболический  $\operatorname{ch} x = \frac{e^x + e^{-x}}{2}$  и тангенс гиперболический  $\operatorname{th} x = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x} - 1}{e^{2x} + 1}$ .

принять любое значение  $x^*$  из отрезка  $[a; b]$ . Важно только, какая будет сделана при этом ошибка  $x^* - \xi$  и допустима ли такая ошибка по условию задачи. Точное значение корня  $\xi$  нам неизвестно, следовательно, ошибку  $x^* - \xi$  вычислить нельзя. Тогда найдём оценку сверху для модуля этой погрешности, т. е. сделаем *априорную оценку* (она выполняется до того, как произведены вычисления; более точны *апостериорные оценки*, которые делаются после того, как приближённое решение задачи найдено).

**ТЕОРЕМА 5. (Оценка погрешности приближённого решения.)** Пусть  $\xi$  — точный, а  $x^*$  — приближённый корни уравнения (1), принадлежащие одному и тому же отрезку  $[a; b]$ . Если найдётся такое положительное число  $m_1$ , что для всех  $x \in [a; b]$  выполняется неравенство  $|f'(x)| \geq m_1 > 0$ , то справедлива оценка

$$|x^* - \xi| \leq |f(x^*)|/m_1, \quad (5)$$

где  $m_1$ , в частности, есть наименьшее значение  $|f'(x)|$  при  $x \in [a; b]$ <sup>8</sup>:

$$m_1 = \inf_{x \in [a; b]} |f'(x)|.$$

*Доказательство.* Применим теорему Лагранжа о конечных приращениях (теорему о среднем значении): если функция  $f(x)$  определена и непрерывна на отрезке  $[x^*; \xi]$  и имеет конечную производную в интервале  $(x^*; \xi)$ , то внутри  $(x^*; \xi)$  найдётся хотя бы одна такая точка  $X$ , для которой

$$f(x^*) - f(\xi) = f'(X)(x^* - \xi).$$

Отсюда получим

$$x^* - \xi = (f(x^*) - f(\xi))/f'(X).$$

---

<sup>8</sup>По второй теореме Вейерштрасса, если функция определена и непрерывна на отрезке, то она достигает на нём своих точных верхней и нижней границ. Поэтому в данном случае  $\inf f(x)$  совпадает с наименьшим значением.

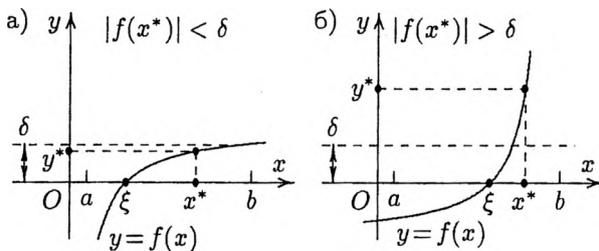


Рис. 6.

Здесь  $X$  — промежуточное значение между  $x^*$  и  $\xi$ , т. е.  $X \in (a; b)$ . Так как  $f(\xi) \equiv 0$  и  $|f'(X)| \geq m_1$ , то взяв по модулю в последнем равенстве, найдём оценку

$$|x^* - \xi| \leq |f(x^*)| / m_1.$$

Теорема доказана.  $\square$

Непонимание смысла этой теоремы может привести к следующей ошибке. Иногда на практике точность найденного приближённого корня  $x^*$  оценивают по тому, насколько хорошо он удовлетворяет уравнению (1). При этом считают, что если число  $|f(x^*)|$  мало, то  $x^*$  является хорошим приближением точного корня  $\xi$ ; если же это число велико, то  $x^*$  — грубое значение корня. Такой подход, как показывает рис. 6, неправильный. Хотя значение  $|f(x^*)|$  на рис. 6, а меньше заданного малого числа  $\delta$ , величина  $x^*$  ещё далека от значения  $\xi$ . На рис. 6, б представлена противоположная ситуация. Несмотря на то, что здесь  $f(x^*) > \delta$ , число  $x^*$  расположено вблизи корня. В первом случае  $m_1 \ll 1$ , а во втором —  $m_1 \gg 1$ . Не следует забывать, что если уравнение (1) умножить на произвольное число  $\lambda \neq 0$ , то получим эквивалентное уравнение  $\lambda f(x) = 0$ , причём значение  $|\lambda f(x^*)|$  можно сделать сколь угодно большим или сколь угодно малым за счёт выбора множителя  $\lambda$ .

При контроле решения алгебраического уравнения может быть полезна *теорема Виета* (см. титульную страницу книги).

**ПРИМЕР 5.** Приближённым корнем уравнения  $f(x) = x^3 - 2x - 3 = 0$  является  $x^* = 1.89$ . Оценим абсолютную погрешность этого корня.

*Решение.* Имеем  $f(x^*) = f(1.89) = -0.02873$ . Так как при  $x = 1.9$  значение  $f(1.9) = 0.059$ , то точный корень  $\xi$  содержится в интервале  $(1.89; 1.9)$ . Производная  $f'(x) = 3x^2 - 2$  на этом промежутке монотонно возрастает. Следовательно,

$$m_1 = \inf_{x \in [1.89; 1.9]} |f'(x)| = |f'(1.89)| = 8.7163.$$

Отсюда по формуле (5) получим

$$|x^* - \xi| \leq 0.02873/8.7163 \approx 0.0033. \quad \triangle$$

**Упражнение 3.** Примените теорему (5) в следующих задачах:

1. Корни уравнений  $2x^3 - 5x^2 + 1 = 0$  и  $4x^3 + 3x^2 - 9x + 2 = 0$  отделены на интервале  $(0; 0.6)$ . Оцените погрешности, которые будут сделаны, если в роли приближённого решения этих уравнений взять середину промежутка.

2. Для уравнения  $0.0005p^3 + 0.105p^2 + 2.16p + 65.3 = 0$  найдено приближённое значение корня  $p = -177$ . Оцените абсолютную погрешность этого решения.

## 1.4 Уточнение корня методом бисекций

Итак, любой корень  $\xi$  уравнения (1) локализован т. е.  $\xi \in [a; b]$ . Теперь уточним его. Для этого применим универсальный способ, являющийся разновидностью *метода проб*. Речь идёт о *методе бисекций* (его другие названия — *метод половинного деления, дихотомии, вилки* [12]). Но прежде чем переходить к описанию метода, сначала докажем вспомогательные утверждения, которые понадобятся в дальнейшем, а затем — теорему существования решения уравнения (1) методом бисекций. Заметим, что доказательства подобных теорем можно разделить на два вида: *конструктивные* и *неконструктивные*. Основу *первых* составляет способ фактического построения искомого решения (его мы и применим при обосновании метода бисекций);

вторые, как правило, основаны на рассуждениях от противного: цепочка логических заключений показывает, что решение обязано существовать, ибо в противном случае получится противоречие.

Докажем сначала две леммы о пределах в неравенствах.

ЛЕММА 1.

$$\left. \begin{array}{l} \lim_{n \rightarrow \infty} x_n = a \\ \lim_{n \rightarrow \infty} y_n = b \\ a < b \end{array} \right\} \Rightarrow \text{с номера } N_0 : x_n < y_n.$$

*Доказательство.* Берём  $\varepsilon = (b - a)/2$ . Тогда с номера  $N_2 : x_n - a < (b - a)/2 \Leftrightarrow x_n < (a + b)/2$ , с номера  $N_1 : (a - b)/2 < y_n - b \Leftrightarrow y_n > (a + b)/2$ . С номера  $N_0 = \sup(N_1, N_2)$  получаем  $x_n < (a + b)/2 < y_n \Rightarrow \Rightarrow x_n < y_n$ . Лемма доказана.  $\square$

ЛЕММА 2.

$$\left. \begin{array}{l} \lim_{n \rightarrow \infty} x_n = a \\ \lim_{n \rightarrow \infty} y_n = b \\ \text{с номера } N_1 : x_n < y_n \end{array} \right\} \Rightarrow a \leq b.$$

*Доказательство.* Пусть  $a > b$ . Тогда по лемме 1 с номера  $N_2 : x_n > y_n$ .

А по условию

с номера  $N_1 : x_n < y_n$ .

Таким образом,

с номера  $N_0 = \sup(N_1, N_2)$  получаем  $\left\{ \begin{array}{l} x_n > y_n \\ x_n < y_n \end{array} \right.$ , чего быть

не может. Лемма доказана.  $\square$

*Замечание 4.* Приведём пример, когда  $x_n < y_n$ , а их пределы равны

$$\begin{aligned} x_n &= 1/n, & y_n &= 2/n, \\ \lim_{n \rightarrow \infty} x_n &= \lim_{n \rightarrow \infty} y_n = 0. \end{aligned}$$

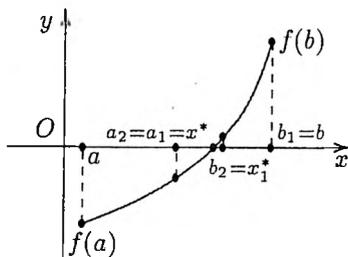


Рис. 7.

*Замечание 5.* Лемма 2 будет, естественно, верна, если в условии задано  $x_n \leq y_n$ .

**Следствие из леммы 2.**

$$\left. \begin{array}{l} x_n \leq b \\ \lim_{n \rightarrow \infty} x_n = a \end{array} \right\} \Rightarrow a \leq b.$$

Теперь перейдём к описанию метода бисекций и доказательству с его помощью теоремы 3 о корне непрерывной функции.

*Доказательство.* Предположим для определённости, что  $f(a) < 0$ , а  $f(b) > 0$  (рис. 7). Разделим точкой  $x^*$  отрезок  $[a; b]$  пополам и в роли начального приближения корня возьмём середину этого отрезка, т. е.  $x^* = (a + b)/2$ . Вычислим  $f(x^*)$ . Если  $f(x^*) = 0$ , то  $\xi = x^*$  является корнем уравнения и теорема доказана. В противном случае,  $f(x^*) \neq 0$ , в качестве нового, суженного отрезка  $[a_1; b_1]$  выберем ту из половин  $[a; x^*]$  или  $[x^*; b]$ , на концах которой функция  $f(x)$  принимает противоположные знаки. В случае, приведённом на рис. 7, оказывается, что  $f(a)f(x^*) > 0$ , а  $f(x^*)f(b) < 0$ , поэтому  $a_1 = x^*$ ,  $b_1 = b$ .

Найденный отрезок снова делим пополам в точке  $x_1^* = (a_1 + b_1)/2$  и получим либо  $f(x_1^*) = 0$ , либо выберем отрезок  $[a_2; b_2]$ , на концах которого  $f(a_2) < 0$ ,  $f(b_2) > 0$ , где, например (рис. 7),  $a_2 = a_1$ , а  $b_2 = x_1^*$ . Будем продолжать этот процесс. Он прервётся на некотором шаге ( $f(x_n^*) = 0$ ) или будет неограниченно продолжаться. В первом случае

вопрос о существовании корня уравнения (1) решён, поэтому рассмотрим *второй* случай.

Неограниченное продолжение процесса даст последовательность вложенных друг в друга отрезков  $[a; b]$ ,  $[a_1; b_1]$ ,  $[a_2; b_2]$ , ..., таких, что:

$$a_n \leq a_{n+1} < b_{n+1} \leq b_n, \quad (6)$$

причём

$$f(a_n) < 0, \quad f(b_n) > 0.$$

Длины отрезков с возрастанием номера  $n$  стремятся к нулю<sup>9</sup>:

$$\lim_{n \rightarrow \infty} (b_n - a_n) = \lim_{n \rightarrow \infty} (b - a)/2^n = 0.$$

Возьмём левые концы отрезков  $(a_n)$ . Согласно (6) они образуют монотонно неубывающую ограниченную последовательность. Она имеет предел, который назовём  $\xi_1$ :

$$\lim_{n \rightarrow \infty} a_n = \xi_1.$$

На основании (6) и лемм 1, 2 получим

$$\xi_1 \leq b_n. \quad (7)$$

Теперь рассмотрим правые концы отрезков  $(b_n)$ . Они образуют монотонно невозрастающую ограниченную последовательность, которая имеет предел  $\xi_2$ :

$$\lim_{n \rightarrow \infty} b_n = \xi_2.$$

Согласно неравенству (7) и лемм 1, 2 эти пределы удовлетворяют неравенству  $\xi_1 \leq \xi_2$ . Итак, имеем

$$a_n \leq \xi_1 \leq \xi_2 \leq b_n$$

и, следовательно,

$$\xi_2 - \xi_1 \leq b_n - a_n = (b - a)/2^n \leq \varepsilon.$$

---

<sup>9</sup>Так как после каждого деления отрезок, содержащий корень, уменьшается вдвое, то после  $n$  делений исходный отрезок сокращается в  $2^n$  раз.

Таким образом, разность  $\xi_2 - \xi_1$  меньше любого наперёд заданного положительного числа  $\varepsilon$ . Это означает, что  $\xi_2 - \xi_1 = 0$ , т. е.  $\xi_1 = \xi_2 = \xi$ .

Найденная точка  $\xi$  — единственная общая точка для всех отрезков построенной последовательности. Докажем, что она является корнем уравнения (1). Согласно определению непрерывности, возможности предельного перехода в неравенствах и учитывая, что  $f(a_n) \leq 0$ , получим

$$f(\xi) = \lim_{n \rightarrow \infty} f(a_n) \leq 0. \quad (8)$$

Аналогично, учитывая, что  $f(b_n) \geq 0$ , имеем

$$f(\xi) = \lim_{n \rightarrow \infty} f(b_n) \geq 0. \quad (9)$$

Из (8) и (9) следует, что  $f(\xi) = 0$ , т. е.  $\xi$  — корень уравнения (1). Теорема доказана.  $\square$

Процесс построения последовательности вложенных стягивающихся отрезков методом бисекций является эффективным вычислительным алгоритмом решения уравнения (1). На  $n$ -м этапе этого процесса получаем неравенство  $a_n \leq \xi \leq b_n$ , которое означает, что число  $a_n$  представляет искомый корень  $\xi$  с *недостатком*, а число  $b_n$  — с *избытком*, с ошибкой, не превосходящей длину отрезка  $\Delta_n = b_n - a_n = (b - a)/2^n$ . При увеличении  $n$  ошибка стремится к нулю по закону геометрической прогрессии со знаменателем  $q = 1/2$ . Если задана точность  $\varepsilon$  ( $\varepsilon > 0$ ), то чтобы её достигнуть, следует выбрать число шагов  $N$ , удовлетворяющее условию

$$N = \left\lceil \ln((b - a)/\varepsilon) / \ln 2 + 0.5 \right\rceil,$$

где  $\ln 2 \approx 0.6931$ .

**ПРИМЕР 6.** Методом бисекций сузим отрезок  $[1.4; 2.4]$ , на котором отделён корень  $\xi$  уравнения  $f(x) = x^3 - 2x - 3 = 0$ , так, чтобы длина полученного отрезка не превышала 0.1.

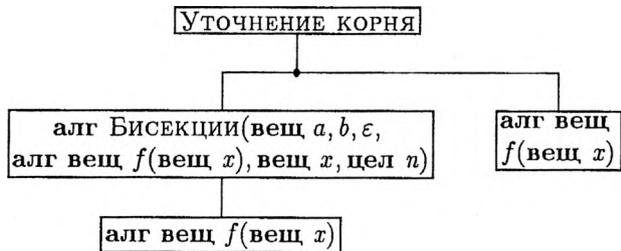


Рис. 8.

*Решение.* Длина исходного отрезка равна 1. Так как на каждом шаге бисекций длина отрезка уменьшается вдвое, то сделаем четыре шага. При этом получим отрезок длиной в  $1/2^4 = 1/16 < 0.1$ . Итак, сначала выясним знаки функции на концах исходного отрезка:  $f(1.4) = -3.056 < 0$ ,  $f(2.4) = 6.024 > 0$ . Теперь приступим к делению:

- 1)  $x^* = (1.4 + 2.4)/2 = 1.9$ ,  $f(1.9) \approx 0.1$ ,  $\xi \in (1.4; 1.9)$ ;
- 2)  $x_1^* = (1.4 + 1.9)/2 = 1.65$ ,  $f(1.65) \approx -1.8$ ,  $\xi \in (1.65; 1.9)$ ;
- 3)  $x_2^* = (1.65 + 1.9)/2 = 1.775$ ,  $f(1.775) \approx -1$ ,  $\xi \in (1.775; 1.9)$ ;
- 4)  $x_3^* = (1.775 + 1.9)/2 = 1.8375$ ,  $f(1.8375) \approx -0.5$ ,  
 $\xi \in (1.8375; 1.9)$ .

Всё: найден отрезок  $[1.8375; 1.9]$ , длина которого  $0.0625 < 0.1$ .  $\triangle$

Метод бисекций легко реализуется на компьютере.

**ПРИМЕР 7.** Методом бисекций с точностью  $\varepsilon = 10^{-7}$  уточним корень уравнения  $x^3 - 2x - 3 = 0$ , отделённый на отрезке  $[1.4; 2.4]$ .

*Решение.* Построим функциональную схему (рис. 8). Она содержит четыре алгоритма: головной с именем УТОЧНЕНИЕ КОРНЯ и три подчинённых с именами БИСКЦИИ (на втором уровне иерархии) и  $f$  (на втором и третьем уровнях).

Приведём текст программы на алгоритмическом языке [17].

{  
 алг УТОЧНЕНИЕ КОРНЯ { Метод бисекций }  
 вещь  $a, b, \varepsilon, x, y$  {  $\varepsilon_{\text{компьютерный}}$  см. с. 52, а также [17] }  
 цел  $n$

начало

вывести ' Введите  $a$  — левую границу,  $b$  — правую границу отрезка и точность  $\varepsilon \implies$ '

ввести  $a, b, \varepsilon$

вывести ' Корень уравнения  $x^3 - 2x - 3 = 0.$ '

' Метод бисекций.'

' Исходные данные.'

'  $a =$ ,  $a$ , '  $b =$ ,  $b$ , '  $\varepsilon =$ ,  $\varepsilon$

вызвать БИСЕКЦИИ ( $a, b, \varepsilon, f, x, n$ )

$y \leftarrow f(x)$

вывести ' Ответ:'

' корень  $x =$ ,  $x$ , ' невязка  $f(x) =$ ,  $y$ ,

' число делений исходного отрезка  $n =$ ,  $n$ ,

' Всё ... '

конец { головного алгоритма }

{

алг вещь  $f(\text{вещ } x)$

... текст этого алгоритма см. на с. 20

конец { алг  $f(x)$  }

{

алг БИСЕКЦИИ (вещ  $a, b, \varepsilon,$

алг вещь  $f(\text{вещ } x),$  вещь  $x,$  цел  $n$ )

знач  $a, b, \varepsilon$  рез  $x, n$

вещ  $a_1, b_1$  лог  $\text{знак}_a$

начало

$\text{знак}_a \leftarrow (f(a) < 0)$   $n \leftarrow 0$

$a_1 \leftarrow a$   $b_1 \leftarrow b$

выполнять

$n \leftarrow n + 1$   $x \leftarrow 0.5(a_1 + b_1)$

если  $\text{знак}_a$  экв ( $f(x) < 0$ )

то

```

    a1 ← x
    иначе
    b1 ← x
    конец-если                                     { выбора отрезка }
    пока b1 - a1 > ε                               { ео цикла по ε10 }
    возвратиться
конец                                             { алг Бисекции }
{-----}

```

Приведём результаты решения (язык программирования TR 7.0): корень  $x = 1.893289$ , при этом  $f(x) = -7.3 \cdot 10^{-7}$ , число делений промежутка  $[1.4; 2.4]$  равно 23.  $\triangle$

**Обсуждение.** Метод бисекций — глобально сходящийся алгоритм. Он прост, неприхотлив (требует только непрерывности функции  $f(x)$  и минимума подготовительной работы за столом), универсален и абсолютно надёжен.

**Упражнение 4. 1.** Методом бисекций уменьшите в 10 раз отрезок  $[0; 0.6]$ , на котором отделены корни уравнений:

$$1) 2x^3 - 5x^2 + 1 = 0; \quad 2) 4x^3 + 3x^2 - 9x + 2 = 0.$$

2. Методом бисекций найдите с точностью  $\varepsilon = 10^{-7}$  корни следующих уравнений:

$$1) \omega^6 - 131\omega^4 + 1344\omega^2 - 2880 = 0;$$

$$2) 99.72z^4 + 2.870565 \cdot 10^{-7}z^2 + 6.674356 \cdot 10^{-16}z - 7600 = 0;$$

$$3) e^x = 1/(x + 1);$$

$$4) \pi \ln \pi x - 1/(\pi x) = 0.$$

3. Методом бисекций найдите первые три положительных корня ( $\varepsilon = 10^{-6}$ ):

$$1) x - \sin 25x = 0;$$

$$2) x^2 + 2x + 1 - \operatorname{tg} \pi x = 0;$$

$$3) \sin x(\operatorname{ch} x + 1) - \operatorname{sh} x(\cos x + 1) = 0;$$

$$4) \alpha - 2(1 + \operatorname{ch} \alpha \cos \alpha)/(\operatorname{ch} \alpha \sin \alpha - \operatorname{sh} \alpha \cos \alpha) = 0;$$

5)  $E - x - RI_0(e^{Dx} - 1) = 0$  при  $E = 2$ ,  $R = 10$ ,  $I_0 = 10^{-8}$  и  $D = 20$ ;

<sup>10</sup>ео — сокращение от end of.

6)  $(E - x)/R - Axe^{-\alpha x} - D(e^{\beta x} - 1) = 0$ , где  $E = 1$ ,  $R = 125$ ,  $A = 0.2718$ ,  $\alpha = 10$ ,  $D = 10^{-8}$  и  $\beta = 20$ .

4. С точностью  $10^{-5}$  найдите первый положительный корень уравнения

$$\alpha^3(\operatorname{ch} \alpha \cos \alpha + 1)/(\operatorname{ch} \alpha \sin \alpha - \operatorname{sh} \alpha \cos \alpha) + E = 0$$

при условии, что  $E$  принимает значения: 0, 20, 65, 100.

5. С точностью  $\varepsilon = 10^{-7}$  решите уравнение

$$\Gamma(x) = 1.5, \quad x \in [2; 3], \quad 1/\Gamma(x) = xe^{Cx} \prod_{k=1}^{\infty} (1 + x/k)e^{-x/k},$$

$$C = 0.577\,215\,664\,901\,532\,5.$$

6. Определите глубину погружения пробкового шара радиуса 50 см, плавающего в воде. Плотность пробки  $\rho = 0.3 \text{ г/см}^3$ ;  $\varepsilon = 10^{-3}$ .

7. Шар радиуса  $R$  разделите параллельными плоскостями на  $n$  равновеликих частей;  $\varepsilon = 10^{-5}$ .

8. Две лестницы, одна в 20 м длины, другая в 30 м, поставлены поперёк улицы и опираются своими концами на противоположные дома. Определите с точностью до 1 см ширину улицы, если точка пересечения лестниц находится на высоте 8 м над землёй.

9. Площадь  $S$  кругового сегмента, дуга которого  $\alpha$ , равна  $R^2(\alpha - \sin \alpha)/2$  ( $\alpha$  есть радианная мера дуги). Найдите сегмент, площадь которого равна 0.2 площади круга (найти сегмент — значит определить угловую меру его дуги);  $\varepsilon = 10^{-3}$ .

10. Через точку, лежащую на диаметре полукруга, проведите прямую, делящую площадь полукруга в отношении  $m : n$ ;  $\varepsilon = 10^{-4}$ .

11. Определите, описание какого алгоритма приведено ниже:

$$x_0' = a, \quad h_0 = (b - a)/2, \quad s_0 = \operatorname{sign} f(a),$$

$$x_n = x_{n-1} + s_0 \cdot \operatorname{sign} f(x_{n-1}) \cdot h_{n-1},$$

$$h_n = h_{n-1}/2, \quad n \in \mathbf{N}.$$

## 2. Метод простых итераций

### Постановка задачи

Пусть дано уравнение

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

где функция  $f(x)$  определена и непрерывна всюду на отрезке  $[a; b]$ , содержащем действительный корень уравнения. Требуется вычислить его с любой наперёд заданной точностью  $\varepsilon$ .

Многие методы численного решения уравнения (10) основаны на идее последовательных приближений. Ниже будет получена *рекуррентная* формула<sup>11</sup> метода простой итерации  $x_{n+1} = \varphi(x_n)$ ,  $n \in \mathbf{Z}_0$ . Как видно, она позволяет выразить  $(n+1)$ -й член последовательности через значение её  $n$ -го члена. При наличии рекуррентной формулы последовательность  $(x_n)$  определяется выбором нулевого члена  $x_0$ . Эта последовательность называется *итеративной*, причём *итерация*<sup>12</sup> есть результат неоднократного применения какой-либо математической операции (в данном случае вычислений по рекуррентной формуле). Другими словами, *итерацией* называется каждый шаг в итеративном процессе уточнения начального приближения  $x_0$  корня. В результате итераций находится последовательность приближённых значений  $(x_n)$  корня  $\xi$ . Если эти значения с ростом  $n$  приближаются к  $\xi$ , то говорят, что итеративный процесс *сходится*.

При построении последовательности  $(x_n)$  возникают два вопроса:

1. Можно ли итеративный процесс продолжать неограниченно и будут ли числа  $x_n$  принадлежать отрезку  $[a; b]$ ?

2. Если итеративный процесс бесконечен, то как ведут себя числа  $x_n$  при  $n \rightarrow \infty$ ?

<sup>11</sup> Лат. *recurrens* — возвращающийся.

<sup>12</sup> Лат. *iteratio* — повторение.

Итак, важно знать, сходится ли последовательность  $(x_n)$  к корню уравнения (10), и, если сходится, то какова скорость сходимости. Напомним некоторые факты, которые потребуются в дальнейшем. Речь идёт о признаке сходимости, базирующемся только на свойствах членов данной последовательности:

*для сходимости последовательности  $(x_n)$  необходимо и достаточно, чтобы она удовлетворяла условию Коши, т. е. при любом  $\varepsilon > 0$  существует такой номер  $N$ , что для всех  $p \in \mathbf{Z}_0$  выполняется неравенство  $|x_{n+p} - x_n| < \varepsilon$ .*

Заметим, что последовательность, удовлетворяющая условию Коши, называется *последовательностью Коши* (по-другому, *фундаментальной последовательностью*).

## 2.1 Метод итераций

Итак, уточним корень уравнения (10) методом итераций. Для этого заменим данное уравнение эквивалентным уравнением вида

$$x = \varphi(x), \quad (11)$$

где функция  $\varphi(x)$  непрерывна на  $[a; b]$ . Возьмём в роли начального приближения какое-либо значение  $x_0 \in [a; b]$  и подставим его в правую часть уравнения (11). Вычислим значение  $\varphi(x_0)$ . Получим первое приближение корня, которое назовём  $x_1$ . Подставим теперь  $x_1$  в  $\varphi(x)$ , найдём новое число — второе приближение  $x_2 = \varphi(x_1)$ . Продолжая этот процесс, построим последовательность чисел  $(x_n)$ , определяемых рекуррентной формулой

$$\begin{cases} x_{n+1} = \varphi(x_n), & n \in \mathbf{Z}_0, \\ x_0 = \frac{a+b}{2}. \end{cases} \quad (12)$$

Если последовательность  $(x_n)$  сходится, т. е. существует конечный предел  $\lim_{n \rightarrow \infty} x_n = \xi$ , то, переходя к пределу в

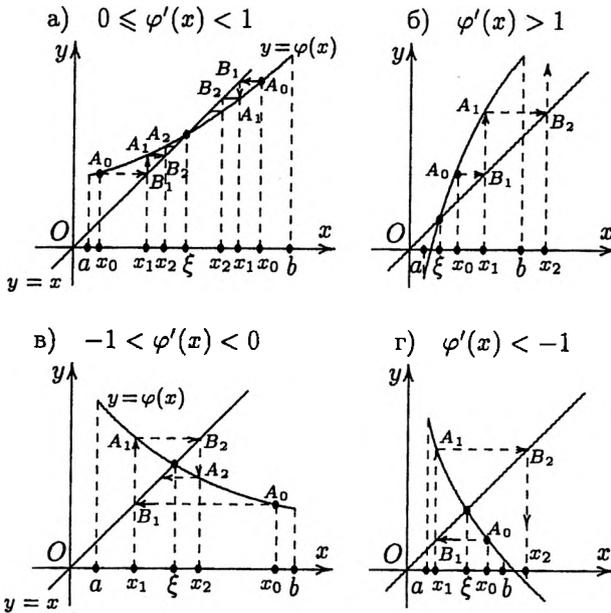


Рис. 9.

равенстве (12) при  $n \rightarrow \infty$ , найдём, что

$$\lim_{n \rightarrow \infty} x_{n+1} = \varphi(\lim_{n \rightarrow \infty} x_n)$$

или

$$\xi = \varphi(\xi).$$

Это равенство означает: предел  $\xi$  является корнем уравнения (11), а следовательно, и уравнения (10). В силу сходимости процесса (12) этот корень можно вычислить (при достаточно большом  $n$ ) с любой наперёд заданной точностью  $\varepsilon$ .

Приведём геометрическую интерпретацию метода. Действительный корень уравнения (11) является абсциссой точки пересечения графиков функций  $y = x$  и  $y = \varphi(x)$ . Построим эти графики. Здесь возможны два случая:  $\varphi'(x) \geq 0$  (рис. 9, а, б) и  $\varphi'(x) < 0$  (рис. 9, в, г).

Возьмём на оси  $Ox$  произвольную точку с абсциссой  $x_0$ , лежащую на отрезке  $[a; b]$ . Проведём через неё вертикаль-

ную прямую до пересечения с кривой  $y = \varphi(x)$  в точке  $A_0(x_0, \varphi(x_0))$ . Затем, учитывая (11), через точку  $A_0$  проведём горизонтальную прямую до пересечения с прямой  $y = x$  в точке  $B_1$ . Но  $\varphi(x_0)$  есть новое значение  $x_1$ , следовательно, точка  $B_1$  имеет координаты  $(x_1, \varphi(x_0))$ . Далее проведём через точку  $B_1$  вертикальную прямую. Она пересечёт кривую  $y = \varphi(x)$  в точке  $A_1(x_1, \varphi(x_1))$ . Продолжая процесс построения, в результате получим ломаную с бесконечным множеством звеньев  $A_0B_1A_1B_2A_2\dots$  и бесконечным множеством вершин, причём вершины  $A_0, A_1, A_2, \dots$  лежат на кривой, а вершины  $B_1, B_2, \dots$  — на биссектрисе первого координатного угла. Абсциссы вершин этой ломаной  $x_0, x_1, x_2, \dots$  суть не что иное, как члены итеративной последовательности  $(x_n)$ .

На рис. 9, а вблизи корня  $\xi$  функция  $\varphi(x)$  возрастает и выполнено условие  $0 < \varphi'(x) < 1$ , а итеративная последовательность либо убывает (если  $x_0 > \xi$ ), либо возрастает ( $x_0 < \xi$ ). Ломаная линия имеет вид *лестницы* с бесконечным множеством ступенек. В обоих случаях последовательность  $(x_n)$  сходится к корню  $\xi$  монотонно (*одностороннее приближение*).

На рис. 9, в в окрестности корня  $\xi$  функция  $\varphi(x)$  убывает и истинно неравенство  $-1 < \varphi'(x) < 0$ . Ломаная имеет вид бесконечной *спирали*, сжимающейся при приближении к точке пересечения кривой и биссектрисы (значения  $x_0, x_1, x_2, \dots$  “обёртывают” корень  $\xi$ ; точка  $\xi$  — *центр итерации*). Последовательность сходится к корню, но она немонотонна: чередуются члены последовательности, большие и меньшие (*двустороннее приближение*).

На рис. 9, б и г в окрестности корня  $|\varphi'(x)| > 1$ . Процесс итераций расходится: как бы близко к корню ни взять нулевое приближение  $x_0$ , каждое последующее значение отстоит дальше от корня  $\xi$ , чем предыдущее. В такой ситуации метод итераций неприменим.

Если же окажется, что вблизи корня  $|\varphi'(x)| = 1$ , то по-

лучим бесконечный итеративный процесс (“холостой ход”, “вечное повторение действий”): при  $\varphi'(x) = 1$  все приближения равны между собой, а при  $\varphi'(x) = -1$  равны между собой значения с чётными номерами  $x_0, x_2, \dots, x_{2k}, \dots$  и равны с нечётными —  $x_1, x_3, \dots, x_{2k+1}, \dots$ . Конечно, здесь в рассуждениях игнорируются погрешности округлений, т. е. предполагается, что последовательные приближения находятся точно.

## 2.2 Условие сходимости метода итераций

Прежде чем применять любой итеративный метод, необходимо знать, будет ли сходиться последовательность  $(x_n)$  к корню уравнения (11) (а следовательно, и уравнения (10)) при возрастании  $n$ ?

**Определение 1.** Метод приближённого решения уравнения (10) называется *сходящимся*, если последовательность приближённых значений корня  $(x_n)$  сходящаяся, т. е. имеется конечный предел, равный точному значению корня:  $\lim_{n \rightarrow \infty} x_n = \xi$ .

Выведём *достаточные* условия сходимости процесса итераций, т. е. условия, которые служат гарантией того, что последовательные значения  $x_n$  будут приближаться к корню уравнения (11). Отметим, что эти условия не будут *необходимыми*, так как существуют уравнения, для которых они не выполняются, но тем не менее посредством итераций (12) можно уточнить корень.

**ТЕОРЕМА 6. (О сходимости итеративной последовательности и единственности решения.)** Пусть  $\xi$  — корень уравнения (11) и пусть функция  $\varphi(x)$  определена и дифференцируема на отрезке  $[a; b]$ , причём для  $x \in [a; b]$  все её значения  $\varphi(x) \in [a; b]$ . Тогда, если существует такое положительное число (постоянная Липшица, коэффициент сжатия)  $q < 1$ , что при  $x \in [a; b]$  выполняется неравен-

ство

$$|\varphi'(x)| \leq q < 1, \quad (13)$$

то на отрезке  $[a; b]$  уравнение (11) имеет единственный корень  $x = \xi$  и процесс итераций (12) сходится к этому корню независимо от выбора начального приближения  $x_0 \in [a; b]$ .

Сформулированная теорема имеет простой смысл. Скажем, что функция  $\varphi$  осуществляет отображение точки  $x$  на точку  $y = \varphi(x)$ . Тогда условие (13) означает, что отображение  $\varphi$  является сжимающим: расстояние между точками  $x_1$  и  $x_2$  больше, чем расстояние между их изображениями  $y_1 = \varphi(x_1)$  и  $y_2 = \varphi(x_2)$ . Корень  $\xi$  — неподвижная точка отображения  $\varphi$  (он преобразуется сам в себя:  $\xi = \varphi(\xi)$ ). Значит, каждый шаг в процессе (12), сжимая расстояние, должен приближать члены итеративной последовательности  $(x_n)$  к неподвижной точке  $\xi$ .

Перейдём к доказательству теоремы 6.

*Доказательство.* Сначала докажем, что метод сходится. Выберем произвольным образом число  $x_0 \in [a; b]$  и по формуле (12) построим последовательность точек  $x_1, x_2, \dots, x_n, \dots$ , которые в силу условий теоремы также принадлежат отрезку  $[a; b]$ . Найдём разность двух последовательных приближений:

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

Используя теорему Лагранжа о приращении функции, получим

$$x_{n+1} - x_n = \varphi'(x^*)(x_n - x_{n-1}),$$

где  $x^* \in (x_{n-1}; x_n)$ . В последнем равенстве возьмём по модулю и воспользуемся условием (13). В результате получим неравенство

$$|x_{n+1} - x_n| \leq q(x_n - x_{n-1}).$$

Отсюда, полагая  $n = 1, 2, 3, \dots$ , шаг за шагом находим

$$\begin{aligned} |x_2 - x_1| &\leq q|x_1 - x_0|, \\ |x_3 - x_2| &\leq q|x_2 - x_1| \leq q^2|x_1 - x_0|, \\ &\dots\dots\dots \\ |x_{n+1} - x_n| &\leq q|x_n - x_{n-1}| \leq q^n|x_1 - x_0|. \end{aligned} \quad (14)$$

Оценка (14) позволяет доказать *фундаментальность* последовательности  $(x_n)$ . Действительно, пусть  $p$  — любое натуральное число. Тогда

$$x_{n+p} - x_n = \sum_{k=1}^p (x_{n+k} - x_{n+k-1}),$$

и согласно (14) имеем

$$\begin{aligned} |x_{n+p} - x_n| &\leq |x_1 - x_0| \sum_{k=1}^p q^{n+k-1} = \\ &= |x_1 - x_0| q^n (1 - q^p) / (1 - q) \leq |x_1 - x_0| q^n / (1 - q), \end{aligned}$$

т. е.

$$|x_{n+p} - x_n| \leq |x_1 - x_0| q^n / (1 - q), \quad k, p \in \mathbf{N}. \quad (15)$$

Так как правая часть неравенства (15) стремится к нулю при  $n \rightarrow \infty$  (ибо при  $0 < q < 1$   $\lim_{n \rightarrow \infty} q^n = 0$ ) и не зависит от  $p$ , последовательность  $(x_n)$  является фундаментальной, значит существует предел

$$\lim_{n \rightarrow \infty} x_n = \xi.$$

Переходя в (12) к пределу при  $n \rightarrow \infty$  и учитывая непрерывность функции  $\varphi(x)$ , получим  $\xi = \varphi(\xi)$ , т. е.  $\xi$  — корень уравнения (11).

Теперь докажем *единственность* решения по методу итераций. Предположим противное: пусть  $\zeta$  — ещё один корень уравнения (11), принадлежащий отрезку  $[a; b]$ . Тогда  $\zeta = \varphi(\zeta)$ . Найдём разность

$$\zeta - \xi = \varphi(\zeta) - \varphi(\xi).$$

Применив теорему Лагранжа, получим

$$\zeta - \xi = \varphi'(x^*)(\zeta - \xi)$$

или

$$(\zeta - \xi)(1 - \varphi'(x^*)) = 0,$$

где  $x^* \in (\zeta; \xi)$ . Так как разность  $1 - \varphi'(x^*)$  не равна нулю (см. условие (13)), то  $\zeta - \xi = 0$ , т. е.  $\zeta = \xi$ . Следовательно, корень  $\xi$  единственный. Теорема доказана.  $\square$

*Замечание 6.* В роли начального приближения  $x_0$  можно брать любое число из отрезка  $[a; b]$ , например,  $x_0 = (a+b)/2$ , а в роли постоянной Липшица  $q$  — наибольшее значение модуля производной функции  $\varphi(x)$  на отрезке  $[a; b]$ :

$$q = \sup_{x \in [a; b]} |\varphi'(x)| \quad (16)$$

(напомним, что у функции, заданной на отрезке,  $\sup$  совпадает с наибольшим значением). Если при решении уравнения (11) окажется, что  $0 < q < 1$ , то условие (13) выполнено. Величина  $q$  характеризует скорость сходимости метода итераций, или, другими словами, качество выбора функции  $\varphi(x)$ .

### 2.3 Оценка приближения в методе итераций

Решение задач на компьютере сопряжено с погрешностями трёх видов: первая, *неустраняемая погрешность* или *погрешность исходной информации* (получается в результате суммирования погрешности математической модели и погрешности исходных данных), вторая, *погрешность метода* или *погрешность аппроксимации* (порождается вынужденной заменой точного решения приближённым) и третья, *вычислительная погрешность* или *погрешность округления* (результат конечности разрядной сетки сумматора). Исследуем *погрешность аппроксимации* на примере метода итераций.



где  $\varepsilon$  — заданная предельная абсолютная погрешность вычисления корня. Тогда в силу (17) будет выполняться неравенство  $|\xi - x_{n+1}| \leq \varepsilon$ , т. е.  $\xi = x_{n+1} \pm \varepsilon$ .

Однако, если  $0 < q \leq 0.5$ , то  $(1 - q)/q \geq 1$  и оценку (17) можно заменить неравенством

$$|\xi - x_{n+1}| \leq |x_{n+1} - x_n|. \quad (19)$$

В этом случае из неравенства

$$|x_{n+1} - x_n| \leq \varepsilon \quad (20)$$

следует, что

$$|\xi - x_{n+1}| \leq \varepsilon. \quad (21)$$

Неравенство (20) можно использовать для оценки погрешности, если на отрезке  $[a; b]$  производная  $\varphi'(x)$  отрицательна, т. е.  $-1 < \varphi'(x) < 0$ . В этом случае имеет место *спиралевидная сходимость* (рис. 10, б), т. е. последовательные приближения  $x_n$  и  $x_{n+1}$  всегда находятся по разные стороны от корня  $\xi$ .

**2.** Для оценки погрешности приближения  $x_{n+1}$  можно воспользоваться и *общим неравенством*

$$|\xi - x_{n+1}| \leq |f(x_{n+1})|/m_1, \quad (22)$$

где  $m_1 = \inf_{x \in [a; b]} |f'(x)|$ .

**ПРИМЕР 8.** Уточним с тремя верными десятичными знаками корень уравнения  $x^3 - 2x - 3 = 0$ , отделённый на отрезке  $[1.4; 2.4]$ .

*Решение.* Преобразовать данное уравнение к виду  $x = \varphi(x)$  можно различными способами, например,

$$\begin{aligned} x &\neq \sqrt[3]{2x + 3}, & x &= \sqrt{2 + 3/x}, \\ x &= (x^3 - 3)/2, & x &= x^3 - x - 3, \\ x &= (2x + 3)/x^2, & x &= (x^3 + x - 3)/3 \end{aligned}$$

и т. д., но условию (13) удовлетворяют только первые два уравнения. Действительно, возьмём, например,  $\varphi(x) = \sqrt[3]{2x + 3}$ . На отрезке  $[1.4; 2.4]$

$$0 < \varphi'(x) = 2 / (3\sqrt{(2x + 3)^2}) < 1,$$

причём  $q = \sup_{x \in [1.4; 2.4]} |\varphi'(x)| = \varphi'(1.4) \approx 0.207$ . Поэтому расчёты можно вести с точностью до четырёх десятичных знаков и в соответствии с правилом (20) прервать вычисления, как только два последовательных приближения совпадут. При этом выполнится неравенство  $|x_{n+1} - x_n| \leq 0.00005$ .

Итак, имеем расчётную формулу

$$x_{n+1} = \sqrt[3]{2x_n + 3}, \quad n \in \mathbf{Z}_0.$$

В роли начального приближения  $x_0$  возьмём значение выражения  $(1.4 + 2.4)/2$ . Приведём результаты расчётов:

$$\begin{array}{lll} x_0 = 1.9000, & x_1 = 1.8945, & x_2 = 1.8935, \\ x_3 = 1.8933, & x_4 = 1.8933, & \end{array}$$

отсюда следует, что  $\xi = 1.8933 \pm 0.00005$ .  $\triangle$

*Замечание 7.* Так как сходимость метода выше для малых значений  $|\varphi'(x)|$ , выгодно подобрать, по возможности, функцию  $\varphi(x)$  так, чтобы её производная в окрестности корня  $\xi$  была близка к нулю. В указанном смысле выбор в роли  $\varphi(x)$  выражения  $\sqrt{2 + 3/x}$  менее удачен, чем только что рассмотренный вариант. Действительно,  $\varphi'(x) = -1.5/(x\sqrt{(2x + 3)x})$  и  $q = \sup_{x \in [1.4; 2.4]} |\varphi'(x)| = \varphi'(1.4) \approx 0.376 > 0.207$ .

## 2.4 Выбор уравнения $x = \varphi(x)$

Преобразование уравнения (10) к виду (11) можно реализовать бесконечным множеством способов, но интересны лишь те варианты, для которых выполняется условие (13). Здесь важно ещё, каково значение коэффициента сжатия

$q$  при этом получится (чем меньше липшицева константа, тем быстрее процесс последовательных приближений сходится).

Укажем некоторые достаточно *общие приёмы* приведения уравнения (10) к виду (11), обеспечивающие выполнение условия (13).

1. Для построения необходимого уравнения полезно воспользоваться *свойством взаимно обратных функций*. Речь идёт о следующем. Пусть уравнение (11) таково, что на отрезке локализации корня истинно неравенство  $|\varphi'(x)| \geq r > 1$ . Тогда процесс итераций для этого уравнения расходится. Однако, если уравнение  $x = \varphi(x)$  заменить эквивалентным уравнением  $x = \psi(x)$ , где  $\psi(x) = \varphi^{-1}(x)$  — обратная функция, то получим сходящийся итеративный процесс, так как

$$|\psi'(x)| = |1/\varphi'(\psi(x))| \leq 1/r = q < 1.$$

**ПРИМЕР 9.** Возьмём уравнение  $x^3 - 2x - 3 = 0$ .

*Решение.* Перейдём к эквивалентному уравнению  $x = (x^3 - 3)/2$ . Это отображение не является сжатым. Действительно,  $\varphi'(x) = 3x^2/2$ , поэтому  $\varphi'(x) \geq 2.94$  при  $x \in [1.4; 2.4]$ . Переход же к обратной функции  $\psi(x) = \sqrt[3]{2x+3}$  даст сходящийся итеративный процесс (см. пример 8).  $\triangle$

2. Укажем достаточно *общий приём* приведения уравнения (10) к виду (11), обеспечивающий выполнение условия (13). Пусть  $\xi \in [a; b]$ , причём

$$0 < m_1 \leq f'(x) \leq M_1 \quad \text{при } x \in [a; b], \quad (23)$$

где в роли  $m_1$  и  $M_1$  возьмём соответственно числа

$$m_1 = \inf_{x \in [a; b]} |f'(x)|, \quad M_1 = \sup_{x \in [a; b]} |f'(x)|$$

(если производная  $f'(x) < 0$ , то вместо уравнения (10) решаем уравнение  $-f(x) = 0$ ).

Эквивалентность уравнений (10) и  $x = x - \lambda f(x)$ , где  $\lambda > 0$ , очевидна. Подберём константу  $\lambda$  так, чтобы в окрестности корня  $\xi$  выполнялось неравенство

$$0 \leq \varphi'(x) = 1 - \lambda f'(x) \leq q < 1.$$

Отсюда на основании (23) получим

$$0 \leq 1 - \lambda M_1 \leq 1 - \lambda m_1 \leq q < 1.$$

Следовательно, если взять

$$\lambda = 1/M_1 \text{ и } q = 1 - m_1/M_1 < 1, \quad (24)$$

то условие (13) будет выполнено.

**ПРИМЕР 10.** Преобразуем уравнение  $x^3 - 2x - 3 = 0$ ,  $\xi \in [1.4; 2.4]$  к приведённому виду (11) при выполнении неравенства (13).

*Решение.* Функция  $f(x) = x^3 - 2x - 3$  возрастает на отрезке локализации корня, так как  $f'(x) = 3x^2 - 2 > 0$  при  $x \in [1.4; 2.4]$ . Значит,  $m_1 = \inf_{x \in [1.4; 2.4]} |f'(x)| = |f'(1.4)| = 3.88$ , а  $M_1 = \sup_{x \in [1.4; 2.4]} |f'(x)| = |f'(2.4)| = 15.28$ . Тогда в силу (24)

найдем, что  $\lambda = 1/M_1 = 1/15.28$ ,  $q = 1 - m_1/M_1 \approx \approx 0.7461$ . Расчётные формулы метода итераций для уравнения  $x^3 - 2x - 3 = 0$  примут следующий вид:

$$x_{n+1} = x_n - \Delta_n, \quad \Delta_n = (x_n^3 - 2x_n - 3)/15.28, \quad x_0 = 1.9, \quad (25)$$

$$|\Delta_n| \leq E, \quad E = \varepsilon(1 - q)/q, \quad q = 0.7461, \quad \varepsilon = 10^{-7}, \quad n \in \mathbf{Z}_0. \quad (26)$$

Формулы (25) и (26) будут реализованы ниже в алгоритме УТОЧНЕНИЕ КОРНЯ.  $\triangle$

**Обсуждение.** Изложенный метод уточнения корня уравнения (10) прост и обладает свойством *самоисправляемости*, а именно: если выполнены условия теоремы 6, то метод итераций сходится при любом выборе начального значения  $x_0$  из отрезка  $[a; b]$ , так как ошибка в вычислениях, не

выводящая за границы этого отрезка, не влияет на конечный результат (если только она не сделана на последнем шаге), ибо ошибочное значение можно рассматривать как новое начальное приближение  $x_0$  в итеративном процессе (12). В худшем случае ошибка приведёт к увеличению количества шагов уточнения. Свойство самоисправляемости усиливает надёжность метода.

Метод итераций легко программируется для использования на компьютере.

**ПРИМЕР 11.** Решим задачу из примера 10 на компьютере.

*Решение.* Построим функциональную схему программы. Задачу будут решать пять алгоритмов — головной (назовём его УТОЧНЕНИЕ КОРНЯ) и подчинённые: ITER (уточняет корень  $x$ ), DELTA (вычисляет поправку  $\Delta_n$ ),  $f$  (контролирует вычисления, находит невязку) и  $f'$  (вычисляет  $f'(x)$ ). Надёжности алгоритмов последовательных приближений уделим особое внимание, так как вполне возможна ситуация, когда вследствие ошибок (неверно отделён корень  $x$ , неправильно выбрано начальное приближение  $x_0$ , число  $\epsilon$  слишком мало, неверно построена функция  $\varphi$ , т. е.  $\Delta_n$ , описки в командах и т. п.) алгоритм ITER будет выполняться “вечно”. Чтобы такое не произошло, будем задавать число  $n_{\max}$  — максимальное количество итеративных шагов, за которое планируем уточнить корень  $x$ . Это число выберем не слишком малым и не слишком большим. В учебных задачах оно равно десяткам, сотням, а в прикладных, практических задачах — сотням, тысячам.

Кроме того будем использовать переменную-флажок  $n_{\text{end}}$ , которая получит текущее значение  $n$ , если выход из цикла приближений совершится при верном неравенстве (26), или  $n_{\max} + 1$ , если поправка по модулю всё ещё больше  $E$ , а выполненное число итераций  $n$  превзойдёт на единицу заданное число  $n_{\max}$ . Поэтому аргументами (значениями) алгоритма ITER будут **вещь**  $x_0$ ,  $\epsilon$  и **цел**  $n_{\max}$ , функция **алг** **вещ** DELTA, а результатами — **вещ**  $x$  и **цел**  $n_{\text{end}}$ .

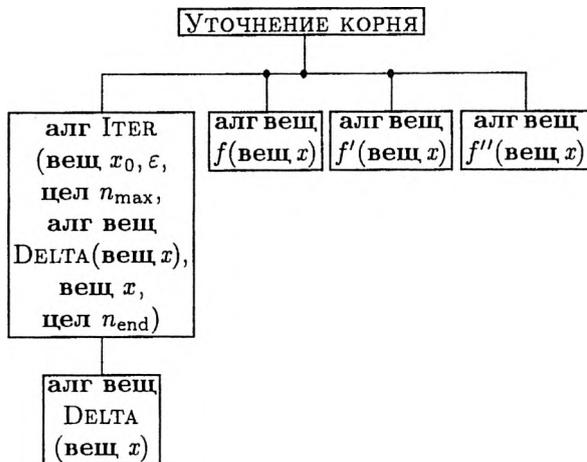


Рис. 11.

Приведём схему иерархии (см. рис. 11). Поясим её<sup>13</sup>. Головной алгоритм передаст аргументы алгоритму ITER, а получив от него продукцию, проанализирует её и независимо от результата анализа (корень уточнён или не уточнён), проконтролирует значение  $x$  и выведет: корень  $x$ , число сделанных итеративных шагов  $n$  и значение  $y = f(x)$ . Эти результаты (“отпечатки пальцев”) могут быть информацией к размышлению. Что мы знаем о корне? Из графического решения приближённое значение корня равно 1.9. Его можно уточнить, вычислив  $x_1$  по формуле  $x_1 = (3 + x_0(17.28 - x_0^2)) / 15.28$ . Первое приближение равно 1.8961. Однако будет лучше, если возложить контроль решения на компьютер. Так мы и поступим: обратимся к алгоритму  $f$ , подставим корень  $x$  в уравнение (10) и определим, насколько близка к нулю невязка (не забывая при этом, что если уравнение (10) умножить на отличное

<sup>13</sup>Эта функциональная схема годна для всех итеративных методов, но алгоритм  $f''$  потребуется только в реализациях методов хорд и касательных.

от нуля значение  $\Lambda$ , то получим эквивалентное уравнение  $\Lambda f(x) = 0$ , причём число  $|\Lambda f(x)|$  можно сделать сколь угодно малым за счёт выбора множителя  $\Lambda$ ).

По значению  $n_{\text{end}}$  будем судить, как отработал алгоритм ITER. При  $\varepsilon = 10^{-7}$  ждём, что процесс итераций сойдётся за десяток-второй шагов. Значение  $n_{\text{max}}$  возьмём с запасом, положив его равным, например, 100. Если окажется, что конечное значение  $n$  равно  $n_{\text{max}} + 1$ , то неравенство (26) не выполнилось ни разу и следует искать причину этого. Таким образом, переменная  $n_{\text{end}}$  играет роль сигнальной переменной (флажка).

Значение  $y = f(x)$  тоже расскажет о многом. Так, если невязка  $f(x) \approx 0$  и  $n_{\text{end}} = n_{\text{max}} + 1$ , то скорее всего, что слишком малы были входные данные  $\varepsilon$  и (или)  $n_{\text{max}}$ . Тогда, например, при медленной сходимости процесса итераций неравенство (26) просто не успело выполниться.

Алгоритмы ITER,  $f$  и  $f'$  находятся на втором уровне иерархии. Они управляются головным алгоритмом. Алгоритм ITER, получив данные извне, подготовит цикл приближений, а затем, уже внутри цикла, обратится к алгоритму DELTA, вычислит поправку  $\Delta_n$ , найдёт очередное приближение к корню, оценит погрешность и либо повторит цикл уточнения корня, либо совершит специальный выход из тела цикла.

Представим проекты всех алгоритмов. Порядок их разработки — операционный [17]: сначала спроектируем головной алгоритм, затем  $f'$ , ITER, DELTA и, наконец,  $f$ .

```

{-----}
алг Уточнение корня { Метод итераций }
  конст вещ  $\varepsilon_{\text{компьютерный}} \leftarrow 1.2 \cdot 10^{-7}$  {см. с. 52, а также [17]}
  вещ  $a, b, f'_a, \text{sgn}, f'_b, m_1, M_1, \lambda, q, \varepsilon, E, x_0, x, y$ 
  общий вещ  $C$ 
  цел  $n_{\text{max}}, n$ 
начало
  вывести '  $a, b$  — границы отрезка и

```

$n_{\max}$  — тах число итераций  $\implies$ 

ввести  $a, b, n_{\max}$

 $f'_a \leftarrow f'(a)$ 

если  $f'_a > 0$  то  $sgn \leftarrow 1$  иначе  $sgn \leftarrow -1$  конец-если

 $f'_a \leftarrow |f'_a| \quad f'_b \leftarrow |f'(b)|$ 

если  $f'_a > f'_b$

то

 $m_1 \leftarrow f'_b \quad M_1 \leftarrow f'_a$ 

иначе

 $m_1 \leftarrow f'_a \quad M_1 \leftarrow f'_b$ 

конец-если

 $\lambda \leftarrow 1/M_1 \quad q \leftarrow 1 - m_1\lambda \quad C \leftarrow sgn \cdot \lambda$ 

пока истина выполнять

вывести  $\varepsilon$  — точность  $\implies$

ввести  $\varepsilon$

 $E \leftarrow ((1-q)/q)\varepsilon$ 

если  $E > \varepsilon_{\text{компьютерный}}$

то

прервать  $\{ \text{break (TP, C++) или exit (F90)} \}$

иначе

вывести сообщение об ошибке,  $E$

конец-если

конец-цикл

$\{ \text{ ввода } \varepsilon \}$

вывести заголовки и значения  $a, b, \varepsilon, E, n_{\max}$

 $x_0 \leftarrow 0.5(a + b)$ 

вызвать ITER( $x_0, E, n_{\max}, \text{DELTA}, x, n$ )

если  $n = n_{\max} + 1$

то

вывести сообщение об ошибке и значение  $n_{\max}$

конец-если

 $y \leftarrow f(x)$ 

вывести заголовок выходных данных,  $x, y, n$

конец

$\{ \text{ головного алгоритма } \}$

{ \_\_\_\_\_ }

алг вещь  $f'(вещ\ x)$



сло итераций  $n_{\max}$ . Их можно варьировать в ходе решения задачи на компьютере, чтобы получить полную картину процесса уточнения корня. В роли  $\epsilon$  сначала можно взять грубое значение, например,  $10^{-1}$  (чтобы хоть что-то получить и убедиться: “программа дошла-таки до конца”), а затем  $\epsilon$  можно постепенно уменьшать, положив  $\epsilon = 10^{-3}$ ,  $10^{-5}$ ,  $10^{-7}$ . Такой подход позволит путём оценки машинного времени уточнения корня правильно выбрать допустимое значение  $\epsilon$ .

Результаты работы ТР-программы при  $x_0 = 1.9$ ,  $\epsilon = 10^{-7}$  и  $n_{\max} = 100$  таковы: корень  $x = 1.893289$ ,  $f(1.893289) = 2 \cdot 10^{-7}$ , число итераций  $n = 15$ .  $\triangle$

Отметим, что алгоритм ITER универсален. Его мы будем использовать для уточнения корня уравнения (11) любым итеративным методом.

В *заключение* приведём алгоритм, позволяющий правильно выбрать входное значение максимально возможной точности  $\epsilon$  для вашего компьютера:

алг EPSILON	{ Компьютерный $\epsilon$ IBM PC (см. [17]): }
вещ $\epsilon$ , $E$	{ single, float, real(4): $1.20 \cdot 10^{-7}$ }
начало	{ real: $9.10 \cdot 10^{-13}$ }
$\epsilon \leftarrow 1.0$	{ double, real(8): $2.23 \cdot 10^{-16}$ }
выполнять	{ extended, long double: $1.09 \cdot 10^{-19}$ }
$\epsilon \leftarrow 0.5\epsilon$	$E \leftarrow \epsilon + 1.0$
пока $E > 1.0$	{ eo цикла по $\epsilon$ }
вывести ' eps компьютерный =', $2.0 \cdot \epsilon$	
конец	{ алг Epsilon }

При типе данных двойной точности в этом алгоритме надо использовать объявление *двточн*  $\epsilon$ ,  $E$ , а константы 1.0, 0.5 и 2.0 представить в форме с двойной точностью.

**Упражнение 5.** 1. Докажите теорему о поведении последовательности  $(x_n)$ :

**ТЕОРЕМА.** Пусть в некоторой окрестности  $[a; b]$  корня  $\xi$

уравнения (11) производная  $\varphi'(x)$  сохраняет постоянный знак и выполняется неравенство (13). Тогда:

1) если производная  $0 \leq \varphi'(x) < 1$ , то последовательные приближения  $x_{n+1} = \varphi(x_n)$ ,  $n \in \mathbf{Z}_0$ ,  $x_0 \in [a; b]$  сходятся к корню  $\xi$  монотонно;

2) если же  $-1 < \varphi'(x) < 0$ , то приближения осциллируют около корня (указание: примените теорему Лагранжа [7] и метод математической индукции).

2. Предложите собственный способ приведения уравнения (10) к виду (11), пригодному для итераций [8].

3. Для каких положительных значений  $x$  сходится последовательность  $\varphi_{n+1}(x) = x^{\varphi_n(x)}$ ,  $n \in \mathbf{Z}_0$ ,  $\varphi_0(x) = 1$ ,  $\varphi_1(x) = x$  (задача Ганса Эрманна)?

4. Проиллюстрируйте (геометрически) решение методом итераций следующих уравнений:

$$1) x - 1 = 0; \quad 3) ax - b = 0, \quad a \neq 0.$$

$$2) 2x - 1 = 0;$$

Рассмотрите также случай “холостого хода”.

5. Методом итераций с точностью  $\varepsilon = 10^{-7}$  уточните корни следующих уравнений:

$$1) 0.0004p^3 + 0.0008p^2 + 0.104p + 1 = 0;$$

$$2) 2 \cdot 10^{-9}p^4 - 2 \cdot 10^{-5}p^3 + 3 \cdot 10^{-3}p^2 - 13 \cdot 10^{-2}p + 100 = 0;$$

$$3) 2^x = x + 3;$$

$$4) 4x = 2^x;$$

$$5) \sqrt{x^2 - 5x + 6} = \sqrt[3]{9x - 10 - 2x^2};$$

$$6) 1 - \frac{k^2}{3 \cdot 4} + \frac{k^4}{3 \cdot 4 \cdot 7 \cdot 8} - \frac{k^6}{3 \cdot 4 \cdot 7 \cdot 8 \cdot 11 \cdot 12} + \dots = 0, \quad k_0 = 4;$$

$$7) x - \frac{x^3}{3} + \frac{x^5}{10} - \dots + (-1)^{n-1} \frac{x^{2n-1}}{(n-1)!(2n-1)} \pm \dots = 0.4431185, \quad x_0 = 0.44.$$

6. Методом итераций, используя равенство  $\sin 54^\circ = \cos 36^\circ$ , найдите  $\sin 18^\circ$  и  $\sin 6^\circ$ ;  $\varepsilon = 10^{-7}$ .

7. Методом итераций решите задачу об Ахиллесе и черепахе: через сколько микросекунд Ахиллес догонит черепаху, если перед стартом расстояние между ними было 1000 шагов (Ахиллес за одну секунду делает 10 шагов, а черепаха при этом проползает расстояние в 1 шаг).

8. Как глубоко погрузится в воду ( $\rho = 1 \frac{\text{г}}{\text{см}^3}$ ) или ртуть ( $\rho = 13.55 \frac{\text{г}}{\text{см}^3}$ ) ледяной шар ( $\rho = 0.917 \frac{\text{г}}{\text{см}^3}$ ) радиуса  $r = 100$  см.

9. Резервуар для нефти представляет собой лежащий цилиндр радиуса  $r$  м. Длина резервуара равна  $\ell$  м. Какова высота уровня нефти, если резервуар заполнен на треть своей ёмкости? Определите высоты всех уровней, отвечающих наполнению 1, 2, ..., 99%.

10. Решите задачу о козе: определите длину верёвки, если для выпаса козы используется половина лужайки (лужайка представляет собой круг радиуса  $r$ , а столб, к которому крепится привязь, находится на границе круга).

### 3. Метод пропорциональных частей

#### 3.1 Метод хорд

Дано уравнение

$$f(x) = 0, \quad (27)$$

где функция  $f(x)$  определена и непрерывна всюду на отрезке  $[a; b]$ , содержащем действительный корень уравнения. Уточним его с точностью  $\epsilon$  методом хорд. Для этого предположим, что на отрезке локализации корня функция  $f(x)$  имеет непрерывные производные  $f'(x)$  и  $f''(x)$ , сохраняющие постоянные знаки. Приведём геометрическое обоснование метода.

**Первый случай.** Рассмотрим сначала случай, когда первая и вторая производные имеют одинаковые знаки при  $x \in [a; b]$ , т. е.  $f'(x)f''(x) > 0$  (рис. 12).

Идея метода хорд состоит в следующем: допустим, что на достаточно малом отрезке  $[a; b]$  функция  $f(x)$  меняется почти линейно. Тогда кривую  $y = f(x)$  на отрезке  $[a; b]$  заменим стягивающей её хордой, проходящей через точки  $A_0(a, f(a))$  и  $B(b, f(b))$ . В роли начального приближения  $x_0$  возьмём  $a$ , а в роли первого приближения  $x_1$  точного

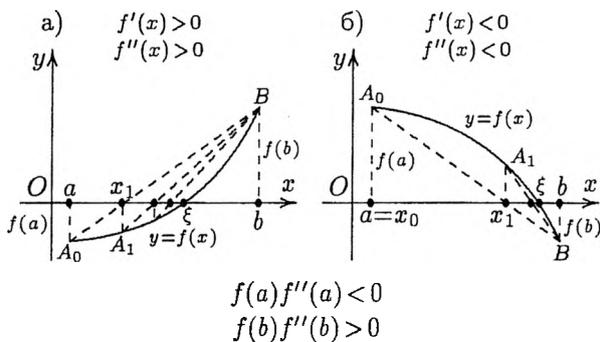


Рис. 12.

корня  $\xi$  — точку пересечения хорды с осью абсцисс. Уравнение хорды  $A_0B$  есть

$$(x - x_0)/(b - x_0) = (y - f(x_0))/(f(b) - f(x_0)).$$

Подставив в него  $y = 0$  и  $x = x_1$ , найдём первое приближение

$$x_1 = x_0 - (b - x_0)f(x_0)/(f(b) - f(x_0)). \quad (28)$$

Затем, проведя хорду через точки  $A_1(x_1, f(x_1))$  и  $B(b, f(b))$ , найдём следующее приближение:

$$(x - x_1)/(b - x_1) = (y - f(x_1))/(f(b) - f(x_1)),$$

$$y = 0, \quad x = x_2,$$

$$x_2 = x_1 - (b - x_1)f(x_1)/(f(b) - f(x_1))$$

и т. д.

Таким образом, если приближение  $x_n$  найдено, то  $(n + 1)$ -ое приближение получим по формуле

$$x_{n+1} = x_n - (b - x_n)f(x_n)/(f(b) - f(x_n)), \quad n \in \mathbf{Z}_0. \quad (29)$$

Заметим, что все хорды имеют общую неподвижную точку  $B(b, f(b))$ , которая характеризуется тем, что в ней знак функции совпадает со знаком её второй производной т. е.  $f(b)f''(b) > 0$  (рис. 12). При этом начальное приближение

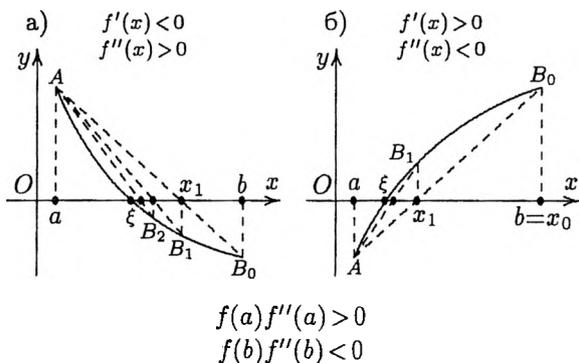


Рис. 13.

$x_0 = a$  и все последующие приближения лежат по левую сторону от корня  $\xi$  и выполняются неравенства

$$f(a)f''(a) < 0,$$

$$a = x_0 < x_1 < x_2 < \dots < x_n < x_{n+1} < \dots < \xi < b,$$

т. е. приближения образуют монотонно возрастающую ограниченную последовательность.

**Второй случай.** В случае, представленном на рис. 13, первая и вторая производные имеют противоположные знаки, так что выполняется неравенство  $f'(x)f''(x) < 0$ . Здесь хорды проводим последовательно через точки  $A(a, f(a))$  и  $B_0(b, f(b))$ ,  $A(a, f(a))$  и  $B_1(x_1, f(x_1))$  и т. д. В результате получим следующее правило для подсчёта  $(n + 1)$ -го приближения:

$$x_{n+1} = x_n - (a - x_n)f(x_n)/(f(a) - f(x_n)), \quad n \in \mathbf{Z}_0. \quad (30)$$

Все хорды имеют общую (неподвижную) точку  $A(a, f(a))$ , в которой знак функции совпадает со знаком её второй производной, т. е.  $f(a)f''(a) > 0$ . Начальным приближением  $x_0$  является  $b$ . Все приближения лежат справа от корня  $\xi$  и выполняются неравенства

$$f(b)f''(b) < 0,$$

$$a < \xi < \dots < x_{n+1} < x_n < \dots < x_2 < x_1 < x_0 = b,$$

т. е. приближения образуют монотонно убывающую ограниченную последовательность.

Обобщая эти результаты, заключаем:

1) формулы (29) и (30) можно объединить в одну

$$x_{n+1} = x_n - (X - x_n)f(x_n)/(f(X) - f(x_n)), \quad n \in \mathbf{Z}_0; \quad (31)$$

2) неподвижен тот конец  $X$ , для которого знак функции  $f(x)$  совпадает со знаком её второй производной  $f''(x)$

$$f(X)f''(X) > 0; \quad (32)$$

3) последовательные приближения (*подвижные точки*) лежат по ту сторону корня  $\xi$ , где функция  $f(x)$  имеет знак, противоположный знаку её второй производной;

4) в роли *начального приближения*  $x_0$  следует взять границу отрезка  $[a; b]$ , противоположную неподвижной точке  $X$ , т. е. для  $x_0$  должно выполняться неравенство

$$f(x_0)f(X) < 0 \quad \text{или} \quad f(x_0)f''(x_0) < 0. \quad (33)$$

*Замечание 8.* Метод хорд имеет ещё одно имя — *метод пропорциональных частей*. Поясним происхождение этого названия. Рассмотрим для определённости случай, представленный на рис. 12, а. Вместо того, чтобы делить отрезок  $[a; b]$ , например, пополам (как это делалось в методе бисекций), разделим его, что, может быть, естественнее, в отношении  $-f(a) : f(b)$ . Из подобия треугольников имеем

$$-f(a)/(x_1 - a) = f(b)/(b - x_1),$$

откуда при  $a = x_0$  получим формулу (28).

### 3.2 Сходимость метода хорд

Докажем теорему о сходимости метода.

**ТЕОРЕМА 7. (О сходимости и единственности решения методом хорд.)** Пусть корень  $\xi$  уравнения (27) отделён на отрезке  $[a; b]$ , т. е.  $f(a)f(b) < 0$ , причём  $f(x)$  вместе с  $f'(x)$  и  $f''(x)$  непрерывны, а  $f'(x)$  и  $f''(x)$  для всех

$x \in [a; b]$  не меняют знака, то, взяв в роли  $X \in [a; b]$  значение, удовлетворяющее неравенству (32), и исходя из начального приближения  $x_0 \in [a; b]$ , удовлетворяющего неравенству (33), можно найти методом хорд (31) единственный корень  $\xi$  уравнения (27) с любой степенью точности  $\varepsilon$ .

*Доказательство.* Пусть для определённости  $f(a) < 0$ ,  $f(b) > 0$ ,  $f'(x) > 0$ ,  $f''(x) > 0$ ,  $x_0 = a$ ,  $X = b$  (рис. 12, а) (в случае  $f''(x) < 0$  следует взять вместо уравнения (27) уравнение  $-f(x) = 0$ ). Приближённые значения корня в методе хорд получаются по формулам

$$\begin{aligned} x_1 &= a - (b - a)f(a)/(f(b) - f(a)), \\ x_2 &= x_1 - (b - x_1)f(x_1)/(f(b) - f(x_1)), \\ &\dots\dots\dots \\ x_{n+1} &= x_n - (b - x_n)f(x_n)/(f(b) - f(x_n)). \end{aligned} \quad (34)$$

В рассматриваемом случае кривая обращена выпуклостью вниз и лежит ниже своей хорды  $A_0B$ . Очевидно, что

$$(b - a)f(a)/(f(b) - f(a)) < 0.$$

Значит  $a < x_1 < b$  и точка  $x_1$  лежит левее точки  $\xi$ , поэтому  $f(x_1) < 0$ . Так как функция  $f(x)$  возрастающая, то разность  $f(b) - f(x_1) > 0$  и, следовательно,

$$(b - x_1)f(x_1)/(f(b) - f(x_1)) < 0.$$

Тогда из формулы для  $x_2$  найдём, что  $x_2 > x_1$ . Аналогично можно показать, что  $x_3 > x_2$  и вообще  $b > x_{n+1} > x_n$  при любом  $n \in \mathbf{Z}_0$ . Итак, итеративная последовательность  $(x_n)$  монотонно возрастает и ограничена сверху. Тогда (*теорема Вейерштрасса*) она сходится, т. е. имеет конечный предел. Обозначим его через  $\eta$ . Значит

$$\lim_{n \rightarrow \infty} x_{n+1} = \eta \leq \xi < b.$$

Докажем теперь, что  $\eta = \xi$  (т. е. предел равен точному корню  $\xi$ ). Перейдём к пределу в формуле (34). Имеем

$$\lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} \left( x_n - (b - x_n)f(x_n)/(f(b) - f(x_n)) \right).$$

Так как в силу непрерывности функции  $\lim_{n \rightarrow \infty} f(x_n) = f(\eta)$ , то на основании теорем о пределах получим

$$\eta = \eta - (b - \eta)f(\eta)/(f(b) - f(\eta)).$$

Отсюда

$$(b - \eta)f(\eta)/(f(b) - f(\eta)) = 0$$

и поэтому  $f(\eta) = 0$ , что означает:  $\eta$  является корнем уравнения (27). Так как  $a < \eta < b$ , а корень уравнения (27) отделён на отрезке  $[a; b]$ , то  $\eta = \xi$ . Итак, метод хорд сходится и уточнён единственный корень.  $\square$

### 3.3 Оценка погрешности в методе хорд

Из сходимости метода хорд следует, что при  $n \rightarrow \infty$  модуль ошибки  $|x_{n+1} - \xi| \rightarrow 0$ , значит, методом хорд можно найти приближённое значение корня уравнения (27) с любой заданной степенью точности.

1. Выведем *неравенство*, позволяющее оценивать абсолютную погрешность приближённого значения  $x_{n+1}$ , если известны два последовательных приближения. Предположим, что выполнены условия теоремы 7 и пусть

$$0 < m_1 \leq |f'(x)| \leq M_1 < \infty,$$

где  $m_1 = \inf_{x \in [a; b]} |f'(x)|$ ,  $M_1 = \sup_{x \in [a; b]} |f'(x)|$ .

Запишем формулу (31) в виде

$$(x_{n+1} - x_n)(f(x_n) - f(X)) = -(x_n - X)f(x_n)$$

или

$$-f(x_n) = (x_{n+1} - x_n)(f(x_n) - f(X))/(x_n - X).$$

Учитывая, что  $f(\xi) = 0$ , получим равенство

$$f(\xi) - f(x_n) = (x_{n+1} - x_n)(f(x_n) - f(X))/(x_n - X).$$

Применив к нему теорему Лагранжа, найдём, что

$$(\xi - x_n)f'(\eta_1) = (x_{n+1} - x_n)(x_n - X)f'(\eta_2)/(x_n - X),$$

где  $\eta_1 \in (x_n; \xi)$ ,  $\eta_2 \in (x_n; X)$ , или

$$(\xi - x_n)f'(\eta_1) = (x_{n+1} - x_n)f'(\eta_2).$$

Вычтя из обеих частей последнего равенства число  $x_{n+1}f'(\eta_1)$  и взяв по модулю в получившемся соотношении, запишем его в виде

$$|\xi - x_{n+1}| = |f'(\eta_2) - f'(\eta_1)| |x_{n+1} - x_n| / |f'(\eta_1)|. \quad (35)$$

Так как  $f'(x)$  сохраняет постоянный знак на отрезке  $[a; b]$ , причём абсциссы  $\eta_1, \eta_2 \in (a; b)$ , то очевидно, что

$$|f'(\eta_2) - f'(\eta_1)| \leq M_1 - m_1.$$

Поэтому из равенства (35) выводим оценку

$$|\xi - x_{n+1}| \leq |x_{n+1} - x_n| (M_1 - m_1)/m_1 \leq \varepsilon \quad (36)$$

или

$$|x_{n+1} - x_n| \leq \varepsilon m_1 / (M_1 - m_1). \quad (37)$$

Если при  $x \in [a; b]$  имеет место неравенство  $M_1 \leq 2m_1$ , то  $m_1/(M_1 - m_1) \geq 1$  и из условия (36) следует

$$|\xi - x_{n+1}| \leq |x_{n+1} - x_n|. \quad (38)$$

Таким образом, как только станет выполняться неравенство

$$|x_{n+1} - x_n| \leq \varepsilon, \quad (39)$$

где  $\varepsilon$  — заданная предельная абсолютная погрешность, то гарантировано, что  $|\xi - x_{n+1}| \leq \varepsilon$ .

2. Точность приближения можно оценить и по *общей формуле*

$$|\xi - x_{n+1}| \leq |f(x_{n+1})|/m_1, \quad (40)$$

где  $m_1 = \inf_{x \in [a; b]} |f'(x)|$ .

ПРИМЕР 12. Методом хорд уточним корень уравнения  $x^3 - 2x - 3 = 0$ ,  $\xi \in [1.4; 2.4]$ ;  $\varepsilon = 10^{-3}$ .

*Решение.* Имеем  $f(x) = x^3 - 2x - 3$ ,  $f'(x) = 3x^2 - 2$ ,  $f''(x) = 6x$ . Так как на отрезке  $[1.4; 2.4]$   $f(1.4) < 0$ ,  $f(2.4) > 0$ ,  $f''(x) > 0$  (здесь налицо случай рис. 12, а), то в силу неравенств (32) и (33) положим  $x_0 = 1.4$ ,  $X = 2.4$ , а  $m_1 = \inf_{x \in [1.4; 2.4]} |f'(x)| = f'(1.4) = 3.88$ . Оценку точности выполним по (40). Расчётная формула (31) примет вид

$$x_{n+1} = x_n - (2.4 - x_n)f(x_n)/(f(2.4) - f(x_n)), \quad n \in \mathbf{Z}_0,$$

где  $f(x_n) = x_n^3 - 2x_n - 3$ ,  $f(2.4) = 6.024$ .

Выполнив вычисления, получим

$$\begin{aligned} n = 0, \quad x_1 &= 1.4 - (2.4 - 1.4)(-3.056)/ \\ &\quad / (6.024 - (-3.056)) = 1.7366; \\ n = 1, \quad x_2 &= 1.7366 - (2.4 - 1.7366)(-1.236)/ \\ &\quad / (6.024 - (-1.236)) = 1.8495; \\ n = 2, \quad x_3 &= 1.8495 - (2.4 - 1.8495)(-0.3725)/ \\ &\quad / (6.024 - (-0.3725)) = 1.8816; \\ n = 3, \quad x_4 &= 1.8816 - (2.4 - 1.8816)(-0.1015)/ \\ &\quad / (6.024 - (-0.1015)) = 1.8902. \\ n = 4, \quad x_5 &= 1.8902 - (2.4 - 1.8902)(-0.027)/ \\ &\quad / (6.024 - (-0.027)) = 1.8925; \\ n = 5, \quad x_6 &= 1.8925 - (2.4 - 1.8925)(-0.0069)/ \\ &\quad / (6.024 - (-0.0069)) = 1.8931. \end{aligned}$$

Так как  $f(x_6) \approx -0.0018$ , то на основании (40) имеем

$$|\xi - x_6| \leq 0.0018/3.88 = 0.0005.$$

Таким образом,  $\xi = 1.8931 \pm 0.0005$ .  $\triangle$

**Обсуждение.** Метод хорд — частный случай метода итераций. У них общая (линейная) скорость сходимости и свойство самоисправляемости. Однако метод хорд проще, так как при его применении отпадает необходимость предварительного преобразования уравнения (27) к приведённому виду, удобному для итераций. При этом успех применения метода хорд гарантирован. Можно рассчитывать на его быструю сходимость, если в окрестности корня  $f(x)$  близка к линейной функции.

Метод хорд легко реализуется на компьютере.

**ПРИМЕР 13.** Методом хорд уточним корень уравнения из примера 12, используя компьютер,  $\varepsilon = 10^{-7}$ .

*Решение.* Оценку погрешности выполним в соответствии с неравенством (37). Имеем  $m_1 = f'(1.4) = 3.88$ ,  $M_1 = \sup_{x \in [1.4; 2.4]} |f'(x)| = f'(2.4) = 15.28$ . Расчётные формулы метода примут вид

$$x_{n+1} = x_n - \Delta_n,$$

$$\Delta_n = (X - x_n)f(x_n)/(f(X) - f(x_n)), \quad n \in \mathbf{Z}_0,$$

$$|\Delta_n| \leq E, \quad E = \varepsilon m_1 / (M_1 - m_1),$$

$$x_0 = a = 1.4, \quad X = b = 2.4, \quad m_1 = f'(a) = 3.88,$$

$$M_1 = f'(b) = 15.28, \quad n_{\max} = 100, \quad \varepsilon = 10^{-7}.$$

Функциональная схема приведена на рис. 11. Эта задача по способу решения и по необходимым ресурсам полностью подобна задаче решения того же уравнения методом итераций, поэтому вряд ли требуются какие-то пояснения.

Приведём проект программы.

```

{-----}
алг УТОЧНЕНИЕ КОРНЯ                                     { Метод хорд }
  конст вещ  $\varepsilon_{\text{компьютерный}} \leftarrow 1.2 \cdot 10^{-7}$ 
  вещ  $a, b, f'_a, f'_b, m_1, M_1, \varepsilon, E, f_a, x_0, x, y$ 
  общий вещ  $X, f_X$ 
  цел  $n_{\max}, n$ 

```

**начало**

**вывести**  $a, b$  — границы отрезка и

$n_{\max}$  — max число итераций  $\implies$ :

**вести**  $a, b, n_{\max}$

$f'_a \leftarrow |f'(a)|$      $f'_b \leftarrow |f'(b)|$

**если**  $f'_a > f'_b$

**то**

$m_1 \leftarrow f'_b$      $M_1 \leftarrow f'_a$

**иначе**

$m_1 \leftarrow f'_a$      $M_1 \leftarrow f'_b$

**конец-если**

**пока истина выполнять**

**вывести**  $\varepsilon$  — точность  $\implies$

**вести**  $\varepsilon$

$E \leftarrow (m_1 / (M_1 - m_1)) \varepsilon$

**если**  $E > \varepsilon_{\text{компьютерный}}$

**то**

**прервать**

**иначе**

**вывести** сообщение об ошибке,  $E$

**конец-если**

**конец-цикл**

{ ввода  $\varepsilon$  }

**вывести** заголовки и значения  $a, b, \varepsilon, E, n_{\max}$

$f_a \leftarrow f(a)$

**если**  $f_a \cdot f''(a) > 0$

**то**

$X \leftarrow a$      $f_X \leftarrow f_a$      $x_0 \leftarrow b$

**иначе**

$X \leftarrow b$      $f_X \leftarrow f(b)$      $x_0 \leftarrow a$

**конец-если**

**вызвать** ITER( $x_0, E, n_{\max}, \text{DELTA}, x, n$ )

**если**  $n = n_{\max} + 1$

**то**

**вывести** сообщение об ошибке и значение  $n_{\max}$

**конец-если**

$y \leftarrow f(x)$

**вывести** заголовок выходных данных,  $x, y, n$

**конец** { головного алгоритма }  
 { \_\_\_\_\_ }

**алг вещь**  $f'$  (**вещ**  $x$ )

... текст этого алгоритма см. на с. 20

**конец** { алг  $f'(x)$  }  
 { \_\_\_\_\_ }

**алг вещь**  $f$  (**вещ**  $x$ )

... текст этого алгоритма см. на с. 20

**конец** { алг  $f(x)$  }  
 { \_\_\_\_\_ }

**алг вещь**  $f''$  (**вещ**  $x$ )

... текст этого алгоритма см. на с. 20

**конец** { алг  $f''(x)$  }  
 { \_\_\_\_\_ }

**алг ITER**( ...

текст этого алгоритма см. на с. 51

**конец** { алг Iter }  
 { \_\_\_\_\_ }

**алг вещь** DELTA (**вещ**  $x$ )

знач  $x$

вещ  $f_x$

общий **вещ**  $X, f_X$

**начало**

$f_x \leftarrow (x^2 - 2)x - 3$

DELTA  $\leftarrow (X - x)f_x / (f_X - f_x)$

**возвратиться**

**конец** { алг Delta( $x$ ) }  
 { \_\_\_\_\_ }

Отметим, что в алгоритмах УТОЧНЕНИЕ КОРНЯ и DELTA используются *общие* (*внешние, глобальные*) данные — это неподвижный конец  $X$  и значение функции  $f(X)$ .

Кодирование алгоритмов было выполнено языком программирования Фортран 90. Результаты решения тако-

вы:  $x = 1.893289$ ,  $f(1.893289) = -4.4 \cdot 10^{-8}$ , число итераций  $n = 14$ .  $\triangle$

**Упражнение 6.** 1. Методом хорд с точностью  $\varepsilon = 10^{-7}$  уточните корни следующих уравнений:

- 1)  $x - 1 = 0$ ;
- 2)  $\alpha(\alpha^4 - 9.74\alpha^3 + 28.91\alpha^2 - 27.83\alpha + 7.04) = 0$ ;
- 3)  $\alpha^4 - 27.8\alpha^3 + 56.07\alpha^2 - 20.5\alpha + 1.4 = 0$ ;
- 4)  $e^x = 1/(x - 1)$ ;
- 5)  $x - \sin(\pi x/2) = 0$ ;
- 6)  $x^2 - 2x + 1 - \operatorname{ctg} \pi x = 0$ ;
- 7)  $x \cdot 2^{3x} - 1 = 0$ ;
- 8)  $x \cdot 3^x = 3.(3)$ ;
- 9)  $x^x + 3x = 5$ .

2. Найдите координаты точек пересечения гиперболы  $y = 1/(x - 1)$  и иррациональной функции  $|y| = \sqrt{x^2 - 1}$ ;  $\varepsilon = 10^{-7}$ .

3. Вычислите  $\cos 20^\circ$ , если  $\cos 60^\circ = 1/2$ ;  $\varepsilon = 10^{-7}$ .

4. Исходя из равенства  $\cos 90^\circ = 0$ , найдите  $\operatorname{tg} 6^\circ$ ;  $\varepsilon = 10^{-7}$ .

5. Определите с точностью до миллиметра глубину погружения в воду ( $\rho = 1 \text{ г/см}^3$ ) шара из пчелиного воска ( $r=100 \text{ см}$ ,  $\rho = 0.955 \text{ г/см}^3$ ).

6. Д. Вегстейн предложил итеративный метод, сходящийся независимо от величины  $|\varphi'(x)|$ . Расчётная формула Вегстейна [4]:

$$\tilde{x}_{n+1} = x_{n+1} - \frac{(x_{n+1} - x_n)(x_{n+1} - \tilde{x}_n)}{x_{n+1} - x_n - \tilde{x}_n + \tilde{x}_{n-1}},$$

где  $x_n = \varphi(\tilde{x}_{n-1})$ ,  $n \in \mathbf{N}$ ,  $\tilde{x}_0 = x_0$ ,  $\tilde{x}_1 = x_1$ ,  $|\tilde{x}_{n+1} - \tilde{x}_n| \leq \varepsilon |\tilde{x}_n|$ ,  $\xi \approx \tilde{x}_{n+1}$ .

Этим методом уточните корень уравнения  $x^3 - 2x - 3 = 0$ . Здесь  $\xi \in [1.4; 2.4]$ ,  $x = \varphi(x)$ ,  $\varphi(x) = (x^3 - 3)/2$ ,  $\varphi'(x) = 1.5x^2$ , и несмотря на то, что условие сходимости метода итераций не выполнено ( $\varphi'(x) \geq 2.94$  при  $x \in [1.4; 2.4]$ ),

алгоритм Д. Вегстейна (алг WEGSTEIN) даст прекрасный результат [17].

## 4. Метод касательных

### 4.1 Метод Ньютона

Пусть корень  $\xi$  уравнения

$$f(x) = 0 \quad (41)$$

локализован на отрезке  $[a; b]$ , причём функция  $f(x)$  определена и непрерывна, а её производные  $f'(x)$  и  $f''(x)$  непрерывны и сохраняют постоянные знаки при  $x \in [a; b]$ . Уточним корень уравнения (41) методом Ньютона (в зарубежной литературе его называют *методом Ньютона — Рафсона*).

Метод Ньютона — это ещё один способ преобразования уравнения (41) к виду, пригодному для построения процесса итераций. Идея его проста. Геометрически алгоритм Ньютона состоит в том, что дуга кривой  $y = f(x)$  заменяется касательной к этой кривой (отсюда и другие названия метода — *метод касательных*, *метод линеаризации*). Здесь возможны два случая.

**Первый случай.** Рассмотрим для определённости случай, когда *первая и вторая производные имеют одинаковые знаки*, т. е.  $f'(x)f''(x) > 0$  при  $x \in [a; b]$ . Пусть при этом  $f(a) < 0$ ,  $f(b) > 0$  (рис. 14, а) или  $f(a) > 0$ ,  $f(b) < 0$  (рис. 14, б). Возьмём произвольную точку  $x_0 \in [a; b]$ , например,  $x_0 = b$ , и напишем уравнение касательной к графику функции  $f(x)$  в точке  $B_0(x_0, f(x_0))$ :

$$y - f(x_0) = f'(x_0)(x - x_0).$$

Графики функции  $f(x)$  и её касательной близки около точки касания, поэтому ждём, что точка пересечения касательной с осью абсцисс (точка  $x_1$ ) будет близка к корню  $\xi$ .

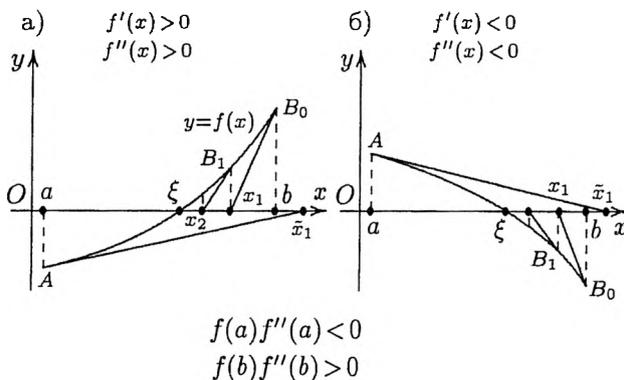


Рис. 14.

Для определения  $x_1$  ( $y = 0$ ,  $x = x_1$ ) имеем уравнение:

$$-f(x_0) = f'(x_0)(x_1 - x_0).$$

Таким образом, получим

$$x_1 = x_0 - f(x_0)/f'(x_0).$$

Теперь корень уравнения (41) находится на суженном отрезке  $[a; x_1]$  и выполняется неравенство:

$$a < \xi < x_1 < x_0 = b.$$

Повторим проделанную только что процедуру. Проведём касательную к графику функции  $f(x)$  в точке  $B_1(x_1, f(x_1))$  и найдём  $x_2$  — точку пересечения касательной с осью  $Ox$  ( $y = 0$ ,  $x = x_2$ ). Имеем

$$x_2 = x_1 - f(x_1)/f'(x_1), \quad (42)$$

причём  $\xi \in [a; x_2]$  и выполняется неравенство

$$a < \xi < x_2 < x_1 < x_0 = b.$$

Продолжая этот процесс, получим последовательность приближённых значений  $(x_n)$ , определяемую рекуррентной формулой

$$x_{n+1} = x_n - f(x_n)/f'(x_n), \quad n \in \mathbf{Z}_0. \quad (43)$$

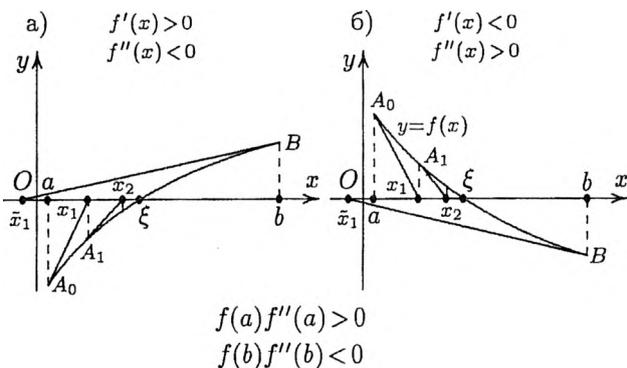


Рис. 15.

Из рис. 14 видно, что каждый последующий член последовательности  $(x_n)$  расположен ближе к корню  $\xi$ , чем предыдущий, однако все  $x_n$  остаются больше точного корня  $\xi$ , т. е.  $x_{n+1}$  есть приближённое значение корня  $\xi$  с избытком:

$$a < \xi < \dots < x_{n+1} < x_n < \dots < x_2 < x_1 < x_0 = b.$$

Заметим, что если провести касательную к кривой  $y = f(x)$  в точке  $A(a, f(a))$ , то касательная может пересечь ось абсцисс в точке  $\tilde{x}_1$ , не принадлежащей отрезку  $[a; b]$ .

**Второй случай.** В случае, изображённом на рис. 15, первая и вторая производные имеют противоположные знаки, так что  $f'(x)f''(x) < 0$ . Здесь выделим два подслучая:  $f(a) < 0$ ,  $f(b) > 0$  (рис. 15, а) и  $f(a) > 0$ ,  $f(b) < 0$  (рис. 15, б).

Если провести касательную к кривой в точке  $B$ , то касательная может пересечь ось  $Ox$  вне отрезка  $[a; b]$  (точка  $\tilde{x}_1$ ). Значит, “хорошим” начальным значением  $x_0$  является то, для которого выполнено неравенство

$$f(x_0)f''(x_0) > 0. \quad (44)$$

Этому условию (его называют *условием Фурье* [4]) отвечает абсцисса точки  $A_0$ , поэтому положим  $x_0 = a$ .

Итак, проведём через точку  $A_0(x_0, f(x_0))$  касательную к кривой до пересечения с осью  $Ox$ :

$$y - f(x_0) = f'(x_0)(x - x_0).$$

Абсциссу точки пересечения назовём  $x_1$ :

$$x_1 = x_0 - f(x_0)/f'(x_0).$$

Корень  $\xi$  теперь находится на отрезке  $[x_1; b]$  и при этом выполняется неравенство

$$a = x_0 < x_1 < \xi < b.$$

Снова строим касательную к кривой, но уже в точке  $A_1(x_1, f(x_1))$ :

$$y - f(x_1) = f'(x_1)(x - x_1).$$

Абсциссу точки пересечения касательной с осью  $Ox$  назовём  $x_2$ :

$$x_2 = x_1 - f(x_1)/f'(x_1).$$

Корень принадлежит суженному отрезку  $[x_2; b]$ , причём истинно неравенство

$$a = x_0 < x_1 < x_2 < \xi < b.$$

Продолжая процесс построения, найдём, что

$$x_{n+1} = x_n - f(x_n)/f'(x_n), \quad n \in \mathbf{Z}_0$$

и

$$a = x_0 < x_1 < x_2 < \dots < x_n < x_{n+1} < \dots < \xi < b,$$

т. е. каждый последующий член последовательности лежит ближе к точному корню  $\xi$ , чем предыдущий. Значит,  $x_{n+1}$  — приближённое значение корня  $\xi$  с *недостатком*.

Сравнивая оба варианта, заметим, что они отличаются только выбором начального приближения: в *первом* варианте за  $x_0$  взят конец отрезка  $b$ , а во *втором* — конец  $a$ . Итак, при выборе начального приближения корня следует руководствоваться следующим правилом (44): в роли  $x_0$

берём тот конец отрезка  $[a; b]$ , в котором знак функции совпадает со знаком её второй производной.

## 4.2 Сходимость метода Ньютона

Выясним *достаточные* условия сходимости процесса Ньютона.

**ТЕОРЕМА 8. (О сходимости метода Ньютона и единственности корня.)** Если на концах отрезка  $[a; b]$  функция  $f(x)$  принимает значения разных знаков, причём  $f'(x)$  и  $f''(x)$  непрерывны, отличны от нуля и сохраняют определённые знаки при  $x \in [a; b]$ , то, исходя из начального приближения  $x_0 \in [a; b]$ , удовлетворяющего условию (44), можно методом Ньютона (формула (43)) вычислить единственный на этом отрезке корень уравнения (41) с любой степенью точности  $\epsilon$ .

*Доказательство.* Так как по условию теоремы  $f(a)f(b) < 0$  и производная  $f'(x)$  знакопостоянна при  $x \in [a; b]$ , то на этом отрезке находится корень уравнения и притом только один. Обозначим его через  $\xi$ .

Для определённости в дальнейшем будем считать, что  $f(a) < 0$ ,  $f(b) > 0$ ,  $f'(x) > 0$  и  $f''(x) > 0$  (см. рис. 14, а). Выберем нулевое приближение  $x_0 \in [a; b]$  так, чтобы выполнялось неравенство (44). Получим  $x_0 = b$ . Исходя из него, найдём приближённые значения корня  $\xi$ :

$$\begin{aligned} x_1 &= b - f(b)/f'(b), \\ x_2 &= x_1 - f(x_1)/f'(x_1), \\ &\dots\dots\dots \\ x_{n+1} &= x_n - f(x_n)/f'(x_n). \end{aligned}$$

Так как  $f(b)/f'(b) > 0$ , то имеем неравенство  $a < x_1 < b$ . Из условия  $f''(x) > 0$  следует, что кривая выпукла вниз, т. е. расположена выше касательной, значит точка  $x_1$  лежит

правее точки  $\xi$  и, следовательно,  $f(x_1) > 0$ . Ввиду положительности производной  $f'(x)$  получим, что  $f(x_1)/f'(x_1) > 0$ . Тогда из уравнения (42) следует, что  $x_2 < x_1$ .

Аналогично покажем, что  $x_3 < x_2$ , и вообще  $x_{n+1} < x_n$  для любого  $n$ . Значит, последовательность приближённых значений корня  $(x_n)$  является монотонно убывающей. Так как любое приближение корня принадлежит отрезку  $[a; b]$ , то  $x_n > a$  для любого  $n$ , т. е. последовательность  $(x_n)$  ограничена снизу. Тогда по *теореме Вейерштрасса* она сходится, т. е. имеет конечный предел. Пусть  $\lim_{n \rightarrow \infty} x_n = \eta$ . Перейдём к пределу в равенстве (43) при  $n \rightarrow \infty$ :

$$\lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} (x_n - f(x_n)/f'(x_n)). \quad (45)$$

В силу непрерывности функции  $f(x)$  и её производной  $f'(x)$  имеем  $\lim_{n \rightarrow \infty} f(x_n) = f(\eta)$  и  $\lim_{n \rightarrow \infty} f'(x_n) = f'(\eta)$ . После перехода в (45) к пределу получим

$$\eta = \eta - f(\eta)/f'(\eta),$$

или

$$f(\eta)/f'(\eta) = 0.$$

Отсюда  $f(\eta) = 0$ , т. е.  $\eta$  является корнем уравнения (41). Так как  $\eta \in [a; b]$ , а корень  $\xi \in [a; b]$  уравнения (41) единственный, то  $\eta = \xi$ . Итак, получим, что  $\lim_{n \rightarrow \infty} x_n = \xi$ , значит метод Ньютона сходится и теорема доказана.  $\square$

### 4.3 Оценка погрешности в методе Ньютона

Из сходимости метода Ньютона следует, что при  $n \rightarrow \infty$  абсолютная ошибка  $|\xi - x_{n+1}| \rightarrow 0$ , т. е. этот метод позволяет уточнить корень уравнения (41) с любой наперёд заданной степенью точности.

1. Для оценки погрешности  $(n+1)$ -го приближения можно воспользоваться *общей формулой*

$$|\xi - x_{n+1}| \leq |f(x_{n+1})|/m_1, \quad (46)$$

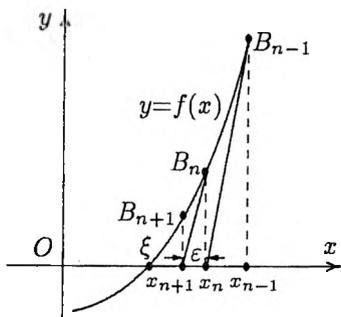


Рис. 16.

где  $m_1 = \inf_{x \in [a; b]} |f'(x)|$ .

2. В Приложении посредством формулы Тейлора [7] получено ещё одно неравенство для оценки погрешности приближения  $x_{n+1}$ :

$$|\xi - x_{n+1}| \leq (x_{n+1} - x_n)^2 M_2 / (2m_1),$$

где  $M_2 = \sup_{x \in [a; b]} |f''(x)|$ , откуда следует, что

$$|x_{n+1} - x_n| \leq \sqrt{(2m_1/M_2)\epsilon}. \quad (47)$$

Заметим, что в общем случае равенство с точностью до  $\epsilon$  двух последовательных приближений  $x_n$  и  $x_{n+1}$  ( $|x_{n+1} - x_n| \leq \epsilon$ ) вовсе не гарантирует, что с той же точностью совпадает с  $\xi$  приближённое значение  $x_{n+1}$  (рис. 16).

Если же на отрезке  $[a; b]$  выполняется условие

$$2m_1/M_2 \geq 1, \quad (48)$$

то из неравенства

$$|x_{n+1} - x_n| \leq \sqrt{\epsilon}$$

следует справедливость неравенства

$$|\xi - x_{n+1}| \leq \epsilon,$$

т. е. совпавшие старшие разряды приближений  $x_n$  и  $x_{n+1}$  верны. Кроме того, можно убедиться, что в случае выполнения неравенства (48) при применении метода Ньютона количество верных цифр искомого корня  $\xi$  удваивается на каждом шаге. Налицо имеем квадратично сходящийся процесс итераций.

**ПРИМЕР 14.** Методом Ньютона уточним корень уравнения  $x^3 - 2x - 3 = 0$ , отделённый на отрезке  $[1.4; 2.4]$ .

*Решение.* Имеем  $f(x) = x^3 - 2x - 3$ ,  $f'(x) = 3x^2 - 2$ ,  $f''(x) = 6x$ ,  $f(1.4) = -3.056 < 0$ ,  $f(2.4) = 6.024 > 0$ . Обе производные на данном отрезке положительны. В силу условия (44) начальное приближение  $x_0$  возьмём на правом конце отрезка:  $x_0 = b = 2.4$ . Тогда получим  $f'(2.4) = 15.28$ ,  $m_1 = f'(1.4) = 3.88$ ,  $M_2 = f''(2.4) = 14.4$ . По формуле (43) найдём  $x_1 = x_0 - f(x_0)/f'(x_0) = 2.4 - 6.024/15.28 = 2.0058$ . Далее вычислим

$$x_2 = 2.0058 - 1.0582/10.0697 = 2.0058 - 0.1051 = 1.9007,$$

$$x_3 = 1.9007 - 0.0652/8.838 = 1.9007 - 0.0074 = 1.8933,$$

$$x_4 = 1.8933 - 0.0001/8.7538 = 1.8933 - 0.00001 = 1.8933$$

и т. д. Кстати, здесь значение  $2m_1/M_2$  равно 0.5389.  $\triangle$

#### 4.4 Модификация метода Ньютона

Если производная  $f'(x)$  мало меняется на отрезке  $[a; b]$ , то в формуле (43) можно положить  $f'(x_n) \approx f'(x_0)$ . В результате получим расчётную формулу *упрощённого* метода Ньютона.

$$x_{n+1} = x_n - f(x_n)/f'(x_0), \quad n \in \mathbf{Z}_0. \quad (49)$$

Геометрически такая модификация означает, что касательные в точках  $B_n(x_n, f(x_n))$ ,  $n \in \mathbf{N}$  заменяются прямыми, параллельными касательной, проведённой к кривой  $y = f(x)$  в точке  $B_0(x_0, f(x_0))$  (рис. 17). Это, с одной сто-

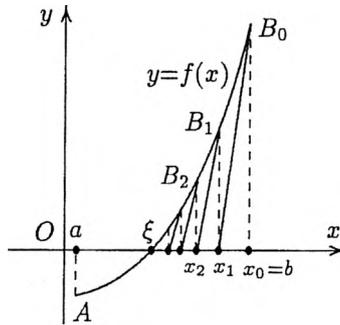


Рис. 17.

роны, приводит к увеличению количества итеративных шагов, необходимых для достижения заданной точности (вместо квадратичной имеем линейную сходимость). С другой стороны, правило (49) избавляет от необходимости вычислять на каждом шаге итераций значение  $f'(x)$ . Значит, такая модификация может быть полезна, когда  $f'(x)$  представляет собой громоздкое выражение. Тогда время, потраченное на “лишние” приближения, как правило, компенсируется уменьшением времени вычисления каждого приближения.

Можно доказать, что если выполнены условия теоремы 8, то упрощённый метод Ньютона сходится.

**ПРИМЕР 15.** Упрощённым методом Ньютона решите задачу из примера 14.

*Решение.* Выберем  $x_0 = 2.4$ , тогда

$$f'(x_0) = 3x_0^2 - 2 \Big|_{x_0=2.4} = 15.28.$$

Расчётная формула (49) примет вид

$$x_{n+1} = x_n - (x_n^3 - 2x_n - 3)/15.28.$$

Используя её, получим

$$\begin{aligned}x_1 &= 2.4 - 6.024/15.28 = 2.0058, \\x_2 &= 2.0058 - 1.0582/15.28 = 1.9365, \\x_3 &= 1.9365 - 0.3889/15.28 = 1.911, \\x_4 &= 1.911 - 0.1568/15.28 = 1.9007\end{aligned}$$

и т. д.  $\triangle$

**Обсуждение.** Метод Ньютона прост, обладает свойством самоисправляемости и имеет высокую скорость сходимости. Из формулы (43) видно, что чем больше значение  $|f'(x)|$  в окрестности корня  $\xi$ , тем меньше поправка, которую нужно вычесть из  $n$ -го приближения, чтобы получить  $(n+1)$ -е приближение. Поэтому метод Ньютона следует применять, когда вблизи корня график функции имеет большую крутизну. Он широко используется, так как быстро сходится и легко переносится на компьютер. В современных вычислительных системах этот метод применяется для (*программной*) реализации операций деления и извлечения квадратного корня.

**ПРИМЕР 16.** Используя оценку (47), решим задачу из примера 14 на компьютере;  $\varepsilon = 10^{-7}$ .

*Решение.* Запишем расчётные формулы для данного случая:

$$\begin{aligned}x_{n+1} &= x_n - \Delta_n, \quad \Delta_n = f(x_n)/f'(x_n), \quad n \in \mathbf{Z}_0, \\|\Delta_n| &\leq E, \quad E = \sqrt{(2m_1/M_2)\varepsilon}, \\x_0 &= b = 2.4, \quad m_1 = f'(1.4) = 3.88, \quad M_2 = f''(2.4) = 14.4, \\n_{\max} &= 100, \quad \varepsilon = 10^{-7}.\end{aligned}$$

Функциональную схему программы см. на с. 48.

Приведём проект программы.

{  
 алг УТОЧНЕНИЕ КОРНЯ { Метод Ньютона }  
 конст вещ  $\varepsilon_{\text{компьютерный}} \leftarrow 1.2 \cdot 10^{-7}$

вещ  $a, b, f'_a, f'_b, m_1, f''_a, f''_b, M_2, \varepsilon, E, x_0, x, y$   
цел  $n_{\max}, n$

**начало**

**вывести** '  $a, b$  — границы отрезка и

$n_{\max}$  — тах число итераций  $\implies$ '

**ввести**  $a, b, n_{\max}$

$f'_a \leftarrow |f'(a)|$      $f'_b \leftarrow |f'(b)|$

**если**  $f'_a > f'_b$  **то**  $m_1 \leftarrow f'_b$  **иначе**  $m_1 \leftarrow f'_a$  **конец-если**

$f''_a \leftarrow |f''(a)|$      $f''_b \leftarrow |f''(b)|$

**если**  $f''_a > f''_b$  **то**  $M_2 \leftarrow f''_a$  **иначе**  $M_2 \leftarrow f''_b$  **конец-если**

**пока истина выполнять**

**вывести** '  $\varepsilon$  — точность  $\implies$ '

**ввести**  $\varepsilon$

$E \leftarrow \sqrt{(2m_1/M_2)\varepsilon}$

**если**  $E > \varepsilon_{\text{компьютерный}}$

**то**

**прервать**

**иначе**

**вывести** сообщение об ошибке,  $E$

**конец-если**

**конец-цикл**

{ ввода  $\varepsilon$  }

**вывести** заголовки и значения  $a, b, \varepsilon, E, n_{\max}$

**если**  $f(a) \cdot f''(a) > 0$

**то**

$x_0 \leftarrow a$

**иначе**

$x_0 \leftarrow b$

**конец-если**

**вызвать** ИТЕР( $x_0, E, n_{\max}, \text{DELTA}, x, n$ )

**если**  $n = n_{\max} + 1$

**то**

**вывести** сообщение об ошибке и значение  $n_{\max}$

**конец-если**

$y \leftarrow f(x)$

**вывести** заголовок выходных данных,  $x, y, n$

```

конец                                     { головного алгоритма }
{-----}
алг вещь  $f'$ (вещ  $x$ )
    ... текст этого алгоритма см. на с. 20
конец                                     { алг  $f'(x)$  }
{-----}
алг вещь  $f''$ (вещ  $x$ )
    ... текст этого алгоритма см. на с. 20
конец                                     { алг  $f''(x)$  }
{-----}
алг вещь  $f$ (вещ  $x$ )
    ... текст этого алгоритма см. на с. 20
конец                                     { алг  $f(x)$  }
{-----}
алг ITER( ...
    текст этого алгоритма см. на с. 51
конец                                     { алг Iter }
{-----}
алг вещь DELTA (вещ  $x$ )
    знач  $x$ 
начало
    DELTA  $\leftarrow ((x^2 - 2)x - 3) / (3x^2 - 2)$ 
    возвратиться
конец                                     { алг Delta( $x$ ) }
{-----}

```

Алгоритмы были закодированы языком программирования C++ (версия 4.5). Результаты решения следующие:  $x = 1.893289$ ,  $f(1.893289) = 7 \cdot 10^{-9}$ , количество итераций  $n = 4$ .  $\triangle$

**Заключение.** Мы познакомились с четырьмя классическими численными методами решения уравнения  $f(x) = 0$  — методами бисекций, итераций, хорд и касательных. Наряду с ними используются ещё более 20 методов [4]. Ситуация, когда одну и ту же задачу можно решать разными методами, типична в практике вычислений. Тогда

естествен вопрос о сравнении методов между собой.

При оценке методов будем учитывать *следующие факторы* ([12], [17]):

- 1) надёжность, универсальность;
- 2) простота организации вычислительного процесса и контроля за точностью;
- 3) скорость сходимости, эффективность;

Проанализируем с этой точки зрения изложенные выше методы (кодирование формул без понимания их сути часто лишь компрометирует вычислительную математику).

1) Универсальным является метод бисекций. Он абсолютно надёжен и неприхотлив (требует от функции  $f(x)$  только непрерывности, а для своего старта — минимума подготовительной работы). Итеративные методы более “тепличны”: они накладывают на функцию  $f(x)$  достаточно сильные ограничения, в частности, знакопостоянство первой,  $f'(x)$ , а то и второй,  $f''(x)$ , производной.

2) Все методы предельно просты и лаконичны, но метод бисекций и здесь обладает преимуществом: стартовать в расчётах можно, начиная с любого конца отрезка  $[a; b]$ , на котором изолирован точный корень  $\xi$ . Процесс бисекций всегда сходится. Он для контроля на каждом шаге имеет двустороннюю оценку, по которой легко судить о достигнутой точности. Сходимость же итеративных методов определяется многими факторами (см. теоремы о сходимости).

3) Как правило, наибольшей скоростью сходимости обладает метод Ньютона, а медленнее всех сходится метод бисекций (сравните параметры алгоритмов “число шагов уточнения”: см. сс. 77, 65, 52 и 32).

Итак, ответ на вопрос о наилучшем методе решения уравнения  $f(x) = 0$  неоднозначен. Он зависит от свойств функции  $f(x)$  и, в соответствии с этим, каким характеристикам метода мы придаём наибольшее значение. На практике чаще всего используется метод Ньютона (в силу высокой скорости сходимости).

При решении практических задач бывает трудно, а то и

невозможно проверить те ограничения, которые накладываются итеративными методами, поэтому нередко применимость метода проверяют “экспериментально”: начинают расчёты и следят за поведением первых значений последовательности  $(x_n)$ ; если видно, что процесс сходится, то уточнение продолжают; в противном случае выявляют причину расходимости и, может быть, выбирают другое начальное приближение и даже — другой метод решения (например, комбинированный). Такой подход носит достаточно общий характер. В прикладной математике и физике нередко используются вычислительные схемы, не получившие строгого обоснования, но они тем не менее успешно применяются при решении сложных задач (вспомним известный афоризм: “чистая” математика делает то, что можно, так, как нужно, а прикладная — то, что нужно, так, как можно).

**Упражнение 7. 1.** Модифицируйте алгоритм ITER — постройте алгоритм ITERATION, который при оценке погрешности использовал бы неравенство (46).

2. Методом Ньютона решите следующие уравнения ( $\varepsilon = 10^{-7}$ ):

1)  $x + 1 = 0$ ;

2)  $x^4 - 2x - 4 = 0$ ;

3)  $x^4 + (x - 2)^2 - 8 = 0$ ;

4)  $12\lambda^3 - 36\lambda^2 + 299\lambda - 264 = 0$ ;

5)  $0.001p^5 + 0.02p^4 + 0.5p^3 + 4p^2 + 10p + 40 = 0$ ;

6)  $\sin x = x^2/|x| - 0.1234567$ ;

7)  $\sqrt{-x^2 + 3x - 2} + \sqrt{x^2 - 4x + 3} = \sqrt{2}(1 - \sqrt{x})$ ;

8)  $x \cdot 2^{2x} = 1$ ;

9)  $x^x = 64$ ;

10)  $5x = 2^x$ ;

11)  $5^x + 12^x - 18^x = 0$ ;

12)  $x^x - 3.(34)x = 5.(67)$ ;

13)  $\ln 2/\pi \cdot \sin(2\pi \ln x/\ln 2) + 1/5 = 0$ .

3. Методом Ньютона решите следующие системы уравнений ( $\varepsilon = 10^{-7}$ ):

- |   |   |
|---|---|
| 1) $\sin x = y + 1.32,$<br>$\cos y = x - 0.85;$ | 3) $y(x - 1) = 1,$<br>$x^2 = y^2 + 1;$                            |
| 2) $2x^3 - y^2 - 1 = 0,$<br>$xy^3 - y - 4 = 0;$ | 4) $5x - 6y + 20 \lg x + 16 = 0,$<br>$2x + y - 10 \lg y - 4 = 0.$ |

4. К какому из корней  $0, \pm 1$  уравнения  $x^3 - x = 0$  сходится процесс Ньютона, если начинать с произвольного значения  $x_0$ ?

5. Решите методом Ньютона следующие задачи ( $\varepsilon = 10^{-7}$ ):

- 1) вычислите корень  $n$ -й степени из числа  $a$ ;
- 2) найдите обратную величину  $x = 1/a$ ;
- 3) вычислите обратную величину квадратного корня  $x = 1/\sqrt{a}$ .

6. Исходя из равенства  $\sin 90^\circ = 1$ , вычислите  $\sin 18^\circ$ ;  $\varepsilon = 10^{-7}$ .

7. Зная, что  $\operatorname{tg} 60^\circ = \sqrt{3}$ , вычислите  $\operatorname{tg} 20^\circ$ ;  $\varepsilon = 10^{-7}$ .

8. Меняя параметр  $\alpha$  от  $0$  до  $\pi$  с шагом  $\pi/20$ , исследуйте методом Ньютона поведение первого положительного корня уравнения ( $\varepsilon = 10^{-7}$ )

$$3 \cos x \cos(\alpha - x) = 2 \sin^2 \alpha x.$$

9. Постройте процесс Ньютона для функции

$$f(x) = \ln 2/\pi \cdot \sin(2\pi \ln x / \ln 2) + 1,$$

где  $x_0 = 2^k$ ,  $k \in \mathbf{Z}_0$ .

10. Определите с точностью до миллиметра глубину погружения в воду ( $\rho = 1\text{г/см}^3$ ) деревянного шара радиуса  $r = 100$  см, если шар состоит из:

- |  |                                       |
|--|---------------------------------------|
| 1) бамбука, $\rho = 0.4\text{г/см}^3;$ | 4) берёзы, $\rho = 0.7\text{г/см}^3;$ |
| 2) сосны, $\rho = 0.45\text{г/см}^3;$  | 5) дуба, $\rho = 0.8\text{г/см}^3.$   |
| 3) кедра, $\rho = 0.55\text{г/см}^3;$  |                                       |

Используйте гибридный метод (методы хорд + Ньютона).

11. Пусть рассматривается случай, представленный на рис. 14, а. Возьмём на графике  $y = f(x)$  точку  $(x_0, f(x_0))$  и проведём через неё прямую, уравнение которой есть

$$y - f(x_0) = k(x - x_0).$$

Из этого уравнения получите расчётные формулы метода хорд, Ньютона и бисекций.

12. Пусть  $\varepsilon=10^{-15}$ ,  $n_{\max}=100$ . Решите следующие задачи:

1) методом итераций уточните наименьший положительный нуль функции

$$y = x^3 - 14.4621x^2 + 60.6959x - 70.9238;$$

2) методом хорд уточните ближайший к точке  $O(0,0)$  нуль функции

$$y = x^3 - 10.2374x^2 - 91.2105x + 492.560;$$

3) модифицированным методом касательных уточните наибольший отрицательный нуль функции

$$y = x^3 - 5.57496x^2 - 193.022x - 633.105;$$

4) методом хорд уточните ближайший к  $-1$  нуль функции

$$y = x^3 + 23.7997x^2 + 44.9378x - 703.378;$$

5) модифицированным методом касательных уточните наименьший положительный нуль функции

$$y = x^3 + 9.57496x^2 - 243.672x + 773.651;$$

6) методом итераций уточните ближайший к точке  $P(-1, -1)$  нуль функции

$$y = x^3 + 20.2374x^2 - 131.210x - 843.923;$$

7) методом хорд уточните ближайший к  $1$  нуль функции

$$y = x^3 + 38.4621x^2 + 364.594x + 914.196;$$

8) методом касательных уточните наибольший положительный нуль функции

$$y = x^3 - 13.3667x^2 + 39.8645x - 20.6282;$$

9) модифицированным методом касательных уточните ближайший к  $\pi$  нуль функции

$$y = x^3 + 2.65804x^2 - 28.0640x + 21.9032;$$

10) методом хорд уточните ближайший к  $e$  нуль функции

$$y = x^3 - 6.49510x^2 - 31.2543x + 23.1782;$$

11) методом касательных уточните ближайший к точке  $P(-\pi, e)$  нуль функции

$$y = x^3 + 9.92960x^2 - 17.8390x - 24.4532;$$

12) методом итераций уточните ближайший к  $-1/e$  нуль функции

$$y = x^3 + 6.09510x^2 - 35.3942x - 25.7283;$$

13) методом хорд уточните ближайший к  $1/\pi$  нуль функции

$$y = x^3 - 10.3296x^2 + 20.3641x + 27.0033;$$

14) модифицированным методом касательных уточните наименьший отрицательный нуль функции

$$y = x^3 - 1.65804x^2 - 34.3767x - 28.2783;$$

15) методом итераций уточните ближайший к точке  $P(-\pi, -\pi)$  нуль функции

$$y = x^3 + 14.7667x^2 - 54.3542x + 29.5533;$$

16) методом хорд уточните ближайший к оси  $Oy$  нуль функции

$$y = x^3 - 14.9667x^2 + 56.4241x - 30.8283;$$

17) методом касательных уточните ближайший к точке  $P(1, 2)$  нуль функции

$$y = x^3 - 5.09510x^2 - 45.7440x + 32.1033;$$

18) методом итераций уточните ближайший к  $\sqrt{2}$  нуль функции

$$y = x^3 - 12.9621x^2 + 50.7295x - 57.8098;$$

19) методом хорд уточните ближайший к  $1/\sqrt{2}$  нуль функции

$$y = x^3 - 4.49967x^2 + 24.4975x + 59.5583;$$

20) методом касательных уточните ближайший к  $2\pi$  нуль функции

$$y = x^3 - 4.99467x^2 + 25.7594x + 58.6693;$$

21) методом итераций уточните ближайший к  $1/e$  нуль функции

$$y = x^3 - 14.4721x^2 - 91.2105x + 562.833;$$

22) методом хорд уточните наименьший по модулю нуль функции

$$y = x^3 + 7.5046x^2 + 11.0866x + 3.8239;$$

23) модифицированным методом касательных уточните наибольший по модулю нуль функции

$$y = x^3 + 10.0128x^2 + 26.6918x - 0.479798;$$

24) методом итераций уточните ближайший к точке  $P(\pi, \pi)$  нуль функции

$$y = x^3 + 1.0864x^2 - 8.5604x + 17.3658;$$

25) методом касательных уточните ближайший к точке  $P(-\pi, -e)$  нуль функции

$$y = x^3 + 10.2543x^2 - 5.8976x - 3.9475;$$

26) модифицированным методом касательных уточните наименьший положительный нуль функции

$$y = x^3 + 8.9958x^2 - 15.6291x - 1.4688;$$

27) методом итераций уточните ближайший к точке  $P(-e, \pi)$  нуль функции

$$y = x^3 - 14.4611x^2 + 47.3564x - 63.4360;$$

28) методом хорд уточните наибольший по модулю нуль функции

$$y = x^3 - 39.2424x^2 - 11.6349x + 76.9317;$$

29) методом итераций уточните ближайший к нулю корень функции

$$y = x^3 - 19.7997x^2 + 28.9378x + 562.833;$$

30) методом касательных уточните наименьший по модулю нуль функции

$$y = x^3 - 1.4451x^2 - 86.6617x + 202.8204;$$

31) методом итераций уточните наименьший положительный нуль функции

$$y = x^3 + 2.7443x^2 + 12.4447x - 17.4259;$$

32) методом хорд уточните ближайший к точке  $P(\pi, e)$  нуль функции

$$y = x^3 - \pi x^2 + e^2 x + 10(\pi - e);$$

33) методом Вегстейна уточните второй нуль функции

$$y = x^3 - 2x^2 - 3x + 1.$$

13. Найдите с точностью  $\varepsilon=10^{-15}$  корни полинома

$$P_{20}(x) = (x-1)(x-2)\dots(x-20) \equiv x^{20} - 210x^{19} + \dots + 20! = 0.$$

14. Предложите свой метод уточнения корня уравнения  $f(x)=0$ .

## ГЛАВА II. Приближённое вычисление интегралов

### 1. Определённый интеграл

#### 1.1 Понятие интеграла

Интеграл — одно из важнейших математических понятий. Решения задач математики, механики и физики, а также финансов и статистики, экономики и других областей приложений нередко выражаются через определённые интегралы (см. [12]).

**ПРИМЕР 1.** К примеру, подсчитаем *путь*, пройденный материальной точкой *при неравномерном движении* за время  $T_1 \leq t \leq T_2$ . Пусть точка перемещается прямолинейно со скоростью  $v(t)$ , где  $v$  — непрерывная функция от  $t$ .

*Решение.* Для решения сформулированной задачи<sup>1</sup> выберем между начальным  $T_1$  и конечным  $T_2$  моментами времени несколько промежуточных моментов  $T_1 < t_1 < t_2 < \dots < t_{n-1} < T_2$ . Кроме того для единообразия обозначений положим  $t_0 = T_1$  и  $t_n = T_2$  (рис. 1).

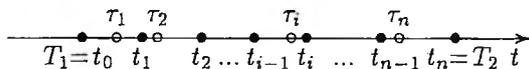


Рис. 1.

Итак, исходный временной промежуток  $[T_1; T_2]$  разбит на  $n$  элементарных отрезков  $[t_{i-1}; t_i]$  ( $i = 1, 2, \dots, n$ ). Рассмотрим один из них. Выберем на нём произвольный момент

<sup>1</sup>Если точка движется равномерно ( $v(t) = v = \text{const}$ ), ответ гтов:  $s = v(T_2 - T_1)$ .

времени  $\tau_i$ ,  $t_{i-1} \leq \tau_i \leq t_i$  и подсчитаем соответствующее ему значение скорости  $v_i = v(\tau_i)$ . Если отрезок времени  $[t_{i-1}; t_i]$  достаточно короток, то скорость  $v_i$  изменяется в нём мало. Пренебрежём изменением скорости. Будем считать её постоянной  $v(t) \approx v_i = v(\tau_i)$ ,  $t_{i-1} \leq \tau_i \leq t_i$ . В результате заданное неравномерное движение на отрезке  $[t_{i-1}; t_i]$  заменится равномерным, а при таком движении точка за временной промежуток  $\Delta t_i = t_i - t_{i-1}$  переместится на расстояние  $\Delta s_i \approx v(\tau_i)\Delta t_i$ .

Чтобы подсчитать длину всего пути, сделаем указанную процедуру для каждого отрезка времени и результаты сложим:

$$s \approx v(\tau_1)\Delta t_1 + v(\tau_2)\Delta t_2 + \dots + v(\tau_n)\Delta t_n. \quad (1)$$

Заметим, что приближённое значение суммы (1) не является постоянным: разным разбиениям промежутка  $[T_1; T_2]$  на части и разному выбору моментов времени  $\tau_i$  на каждом элементарном отрезке соответствуют разные значения  $s$ .

Будем сужать временные промежутки  $[t_{i-1}; t_i]$  ( $i = 1, 2, \dots, n$ ). При этом отклонение движения от равномерного для каждого элементарного отрезка времени уменьшается. Так как  $v$  есть непрерывная функция от  $t$ , то в пределе при неограниченном измельчении разбиения ( $n \rightarrow \infty$ ,  $\max_{1 \leq i \leq n} \Delta t_i \rightarrow 0$ ) приближённое равенство (1) трансформируется в точную формулу для пройденного пути. Предел суммы (1) (если он существует) называется *определённым интегралом* от функции  $v(t)$  по отрезку  $[T_1; T_2]$  и обозначается

$$s = \int_{T_1}^{T_2} v(t) dt. \quad (2)$$

△

**ПРИМЕР 2.** Решим ещё задачу. Подсчитаем *работу переменной силы*  $F$  при перемещении материальной точки

вдоль оси  $Ox$  из положения  $x = a$  в положение  $x = b$ .  $F$  является непрерывной функцией от  $x$ , определённой на отрезке  $[a; b]$ . Пусть направление силы совпадает с направлением движения точки.

*Решение.* Сформулированная задача с математической точки зрения похожа на предыдущую, поэтому для её решения<sup>2</sup> применим аналогичный подход.

Выберем на  $[a; b]$  несколько вспомогательных точек  $a < x_1 < x_2 < \dots < x_{n-1} < b$  и дополнительно положим  $x_0 = a$  и  $x_n = b$ . Тогда исходный отрезок  $[a; b]$  будет разбит на  $n$  элементарных отрезков  $[x_{i-1}; x_i]$  ( $i = 1, 2, \dots, n$ ). Возьмём на каждом из них по произвольной точке  $\xi_i$ ,  $x_{i-1} \leq \xi_i \leq x_i$  и вычислим значения  $F_i = F(\xi_i)$ . Если отрезок  $[x_{i-1}; x_i]$  мал, то и сила изменяется на нём незначительно. Будем приближённо считать её постоянной:  $F(x) \approx F(\xi_i)$ ,  $x \in [x_{i-1}; x_i]$ . Так как  $F$  есть непрерывная функция от  $x$ , то при достаточно малой длине отрезка  $[x_{i-1}; x_i]$  работа по перемещению материальной точки из положения  $x_{i-1}$  в положение  $x_i$  равна  $\Delta A_i \approx F(\xi_i)\Delta x_i$ , где длина  $\Delta x_i = x_i - x_{i-1}$ . Суммируя все  $\Delta A_i$  ( $i = 1, 2, \dots, n$ ), подсчитаем полную работу

$$A \approx F(\xi_1)\Delta x_1 + F(\xi_2)\Delta x_2 + \dots + F(\xi_n)\Delta x_n. \quad (3)$$

Формула (3) похожа на формулу (1). Точность приближённого равенства будет тем выше, чем короче отрезки, на которые разбит исходный отрезок  $[a; b]$ . В пределе при неограниченном измельчении ( $n \rightarrow \infty$ ,  $\max_{1 \leq i \leq n} \Delta x_i \rightarrow 0$ ) из приближённой формулы получим точную формулу для подсчёта выполненной работы в виде *интеграла* от функции  $F(x)$ :

$$A = \int_a^b F(x) dx. \quad (4)$$

△

<sup>2</sup>Если сила постоянна, ответ даётся формулой  $A = F(b - a)$ .

Действуя аналогично, можно найти *массу прямолинейного стержня* длины  $L_1 \leq l \leq L_2$  с линейной плотностью  $\delta(l)$ :

$$m = \int_{L_1}^{L_2} \delta(l) dl, \quad (5)$$

или, например, *количество электричества*, которое переносится через поперечное сечение проводника током  $I(t)$  за время  $T_1 \leq t \leq T_2$ :

$$Q = \int_{T_1}^{T_2} I(t) dt. \quad (6)$$

Наконец, решим последнюю задачу. Исторически понятие интеграла теснейшим образом связано с *квадратурой*, т. е. с задачей об определении площади плоской фигуры.

**ПРИМЕР 3.** Вычислим *площадь криволинейной трапеции* — фигуры (рис. 2), ограниченной графиком некото-

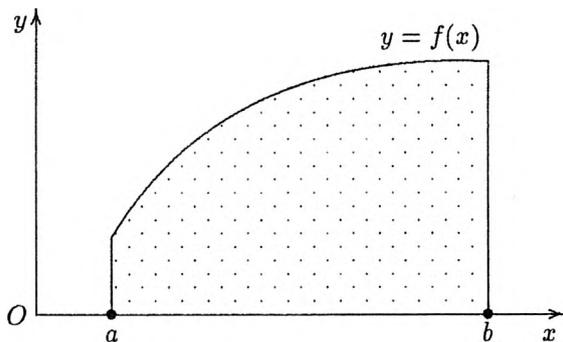


Рис. 2.

рой непрерывной функции, заданной на отрезке  $[a; b]$  (при этом пусть  $f(x) \geq 0 \forall x \in [a; b]$ ), отрезком  $[a; b]$  оси  $Ox$  и двумя вертикалями  $x = a$ ,  $x = b$ .

*Решение.* Для решения этой задачи так же, как в предыдущих случаях, разобьём произвольным образом отрезок  $[a; b]$  точками деления  $x_i$  ( $i = 0, 1, \dots, n$ ;  $x_0 = a$ ,  $x_n = b$ )

на части. Затем на каждом элементарном отрезке возьмём произвольную точку  $\xi_i \in [x_{i-1}; x_i]$  ( $i = 1, 2, \dots, n$ ), найдём в ней значение  $f(x)$  и построим прямоугольник высоты  $f(\xi_i)$ . В результате получим ступенчатую фигуру (“ступенчатый многоугольник”; рис. 3), состоящую из прямоугольников.

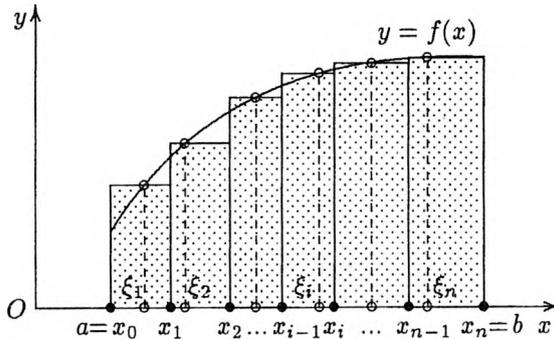


Рис. 3.

Её площадь

$$\tilde{S} = f(\xi_1)\Delta x_1 + f(\xi_2)\Delta x_2 + \dots + f(\xi_n)\Delta x_n \quad (7)$$

будем считать некоторым приближением площади криволинейной трапеции. Увеличим число делений исходного отрезка  $[a; b]$ . Заметим, что точность аппроксимации криволинейной трапеции ступенчатым многоугольником при измельчении разбиения улучшается (сравните рис. 3 и 4). В силу непрерывности функции  $f$  объединение построенных прямоугольников при большом  $n$ , а, значит, при малых  $\Delta x_i$ , почти совпадает с криволинейной трапецией. Поэтому естественно ожидать, что существует предел

$$S = \lim_{\substack{n \rightarrow \infty \\ \max \Delta x_i \rightarrow 0}} \sum_{i=1}^n f(\xi_i)\Delta x_i, \quad (8)$$

т. е. *интеграл*

$$S = \int_a^b f(x) dx, \quad (9)$$

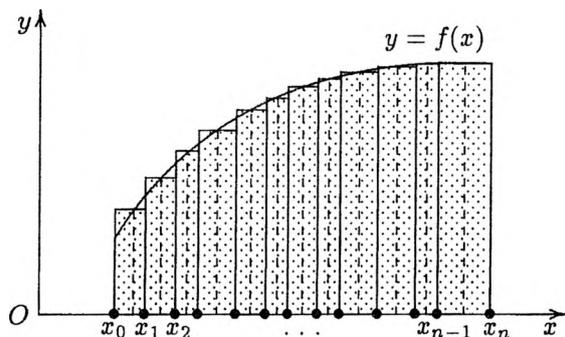


Рис. 4.

являющийся искомой площадью<sup>3</sup>.  $\triangle$

## 1.2 Определение интеграла

Характер задач, решение которых записывается через интеграл, ясен из рассмотренных выше примеров. Приведём теперь определение интеграла.

Пусть некоторая функция  $f(x)$  задана на отрезке  $[a; b]$ . Разделим его точками  $x_i$ ,  $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$  на элементарные отрезки  $[x_{i-1}; x_i]$  ( $i = 1, 2, \dots, n$ ). Выполненную операцию назовём *разбиением*. Обозначим её символом  $Z(x_i)$ . Подсчитаем длины полученных отрезков  $\Delta x_i = x_i - x_{i-1}$  и найдём  $\Delta = \Delta(Z) = \max_{1 \leq i \leq n} \Delta x_i$  — длину наибольшего отрезка при разбиении  $Z(x_i)$ . На каждом из отрезков  $[x_{i-1}; x_i]$  выберем по произвольному числу  $\xi_i$ ,  $x_{i-1} \leq \xi_i \leq x_i$  ( $i = 1, 2, \dots, n$ ) и вычислим значения произведений  $f(\xi_i)\Delta x_i$ . Найденные величины сложим:

$$I(x_i, \xi_i) = f(\xi_1)\Delta x_1 + f(\xi_2)\Delta x_2 + \dots + f(\xi_n)\Delta x_n. \quad (10)$$

Сумму (10) назовём *интегральной суммой* относительно разбиения  $Z$ . Её значение зависит от выбора точек разбиения  $x_i$  и точек  $\xi_i$ . Заметим, что суммы (1), (3) и (7) из п. 1 также являются интегральными.

<sup>3</sup>Если  $f(x) \leq 0 \forall x \in [a; b]$ , то площадь  $S$  равна  $-\int_a^b f(x) dx$ .

Выясним поведение интегральных сумм при неограниченном измельчении разбиения, т. е. при  $\Delta \rightarrow 0$ . Следуя [12], сформулируем два определения.

**Определение 1.** Число  $\mathcal{I}$  называется *пределом интегральных сумм*  $I(x_i; \xi_i)$  при  $\Delta \rightarrow 0$ , если для произвольной точности  $\varepsilon > 0$  можно указать такое  $\delta(\varepsilon)$ , что при любом разбиении  $Z(x_i)$ , удовлетворяющем условию  $\Delta < \delta$ , выполняется неравенство

$$|\mathcal{I} - I(x_i; \xi_i)| < \varepsilon \quad (11)$$

независимо от выбора точек  $\xi_i \in [x_{i-1}; x_i]$ .

**Определение 2.** Если для функции  $f(x)$ , заданной на отрезке  $[a; b]$ , существует предел  $\mathcal{I}$  интегральных сумм при  $\Delta \rightarrow 0$ , то она называется *интегрируемой* на данном отрезке, а число  $\mathcal{I}$  — *определённым интегралом* от этой функции по отрезку  $[a; b]$ , который записывается так<sup>4</sup>:

$$\mathcal{I} = \int_a^b f(x) dx. \quad (12)$$

Здесь числа  $a$  и  $b$  суть *нижний* и *верхний пределы* интегрирования; *знак интеграла*  $\int$  есть стилизованная буква  $S$  — начальная буква латинского слова *Summa*; функция  $f$  — *подынтегральная функция*, а  $x$  — *переменная интегрирования*.

Этому определению “на языке  $\varepsilon - \delta$ ” равносильно определение “на языке последовательностей”: пусть отрезок  $[a; b]$  разбит на части сначала одним способом, затем — вторым, третьим способом и т. д.; функция  $f(x)$  называется *интегрируемой* на  $[a; b]$ , если при любой последовательности  $Z_n$

---

<sup>4</sup>Приведённое определение принадлежит Риману. Он не только его впервые сформулировал, но и исследовал область его применения. По этой причине сумму (10) нередко называют *римановой суммой*; мы же говорим об *интегральной сумме*, желая, таким образом, подчеркнуть её связь с интегралом.

разбиений, для которых  $\lim_{n \rightarrow \infty} \Delta(Z_n) = 0$ , последовательность  $I(Z_n)$  соответствующих интегральных сумм сходится независимо от выбора внутренних точек  $\xi$  к одному и тому же предельному значению (которое и есть *определённый интеграл*).

Таким образом, если интегрируемость  $f(x)$  известна, то достаточно найти предел последовательности  $I(Z_n)$  для какой-нибудь последовательности разбиений  $Z_n$ , удовлетворяющей условию  $\lim_{n \rightarrow \infty} \Delta(Z_n) = 0$ .

**ПРИМЕР 4.** Возьмём функцию, равную постоянной на заданном отрезке, т. е. пусть, например,  $f(x) = C \forall x \in [a; b]$ . Сначала докажем, что она интегрируема, а затем вычислим интеграл  $\mathcal{I} = \int_a^b C dx$ .

*Решение.* Выполним произвольное разбиение  $Z$  отрезка  $[a; b]$ . Выберем точки  $\xi_i$  и составим интегральную сумму

$$I(Z) = C(\Delta x_1 + \Delta x_2 + \dots + \Delta x_n) = C(b - a).$$

Она не зависит ни от разбиения, ни от выбора точек  $\xi_i$ . Значит, существует предел интегральных сумм при  $\Delta(Z) \rightarrow 0$ , т. е. при  $n \rightarrow \infty$ , равный той же величине

$$\mathcal{I} = \lim_{n \rightarrow \infty} I(Z) = \int_a^b C dx = C(b - a). \quad \Delta$$

### 1.3 Точные методы вычисления интегралов

Решаемые задачи нередко сводятся к вычислению интегралов. При этом, как правило, самая трудная часть дела уже позади (см. примеры 1, 2 и 3). Все методы вычисления интегралов делятся на *точные* и *приближённые*.

*Один точный метод* вычисления был предложен Ньютоном и Лейбницем после открытия ими связи между операциями дифференцирования и интегрирования<sup>5</sup>. Они установили, что значение *определённого интеграла* выражает-

<sup>5</sup>Исторически интеграл появился независимо от производной.

ся разностью двух значений, при  $x = b$  и при  $x = a$ , любой первообразной функции  $F^6$ :

$$\int_a^b f(x) dx = F(x) \Big|_a^b = F(b) - F(a). \quad (13)$$

Эта формула верна для произвольной функции  $f$ , непрерывной на отрезке  $[a; b]$ .

ПРИМЕР 5. Вычислим интеграл  $\mathcal{I} = \int_0^1 e^x dx$ .

*Решение.* Функция  $f(x) = e^x$  имеет первообразную  $F(x) = e^x$ . Следовательно,  $\mathcal{I} = e^x \Big|_0^1 = e - 1$ .  $\triangle$

Казалось бы, формула Ньютона — Лейбница является эффективным средством вычисления определённых интегралов (ведь нередко удаётся выразить первообразную в виде композиции элементарных функций<sup>7</sup>). Однако она не даёт общего правила решения рассматриваемой задачи, так как поиск первообразной — математическая проблема, для решения которой в явном виде нет универсальных методов.

ПРИМЕР 6. Например, по формуле (13) невозможно найти значение функции ошибок  $\mathcal{I} = \operatorname{erf} x = 2/\sqrt{\pi} \int_0^x e^{-u^2} du$ , так как первообразная функции  $f(u) = e^{-u^2}$  не является элементарной функцией<sup>8</sup>.  $\triangle$

---

<sup>6</sup>Функция  $F$  называется *первообразной* для функции  $f$  на заданном промежутке, если для всех  $x$  из этого промежутка  $F'(x) = f(x)$ . Обозначение  $F(x) \Big|_a^b$  именуется *двойной подстановкой* от  $a$  до  $b$ .

<sup>7</sup>При этом для нахождения первообразной функции приходится часто применять различные приёмы и ухищрения, например, замену переменной интегрирования и подстановки, интегрирование по частям  $\int u dv = uv - \int v du$ , правило интегрирования суммы функций и даже прибегать к теории функций комплексной переменной.

<sup>8</sup>См. *Е. Янке, Ф. Эмде и Ф. Леш.* Специальные функции. М.: Наука. 1968.; *И. С. Градштейн, И. М. Рыжик.* Таблицы интегралов, сумм, рядов и произведений. М.: Наука. 1971.; *А. П. Прудников, Ю. А. Брычков, О. И. Маричев.* Интегралы и ряды. М.: Наука. 1981.

Кроме того она не работает, если подынтегральная функция определена на фиксированном конечном множестве точек  $x_i$  отрезка  $[a; b]$  (т. е. в случае графического или табличного задания  $f$ ). Итак, формула (13), хотя и называется *основной формулой интегрального исчисления*, не является алгоритмом вычисления определённого интеграла [17].

*Другой точный метод* основан на определении интеграла (см. примеры 4 и 7) как предела интегральных сумм. Однако он не получил распространения, так как даже в простейших случаях требует приложения значительных усилий.

**ПРИМЕР 7.** Найдём интеграл  $\mathcal{I} = \int_a^b \cos x \, dx$ .

*Решение.* Разобьём отрезок  $[a; b]$  на  $n$  равных частей. Обозначим  $\Delta = (b - a)/n$ . Вычислив функцию  $\cos x$  при  $a < b$  для правых концов элементарных отрезков (при  $a > b$  — для левых), получим

$$I_n = \Delta \left( \cos(a + \Delta) + \cos(a + 2\Delta) + \dots + \cos(a + n\Delta) \right).$$

Упростим интегральную сумму  $I_n$ : умножим и разделим её на  $2 \sin(\Delta/2)$ , а затем каждое слагаемое заменим суммой синусов. После приведения подобных членов найдём, что

$$I_n = \frac{\Delta}{2 \sin \frac{\Delta}{2}} \left( \sin \left( a + \left( n + \frac{1}{2} \right) \Delta \right) - \sin \left( a + \frac{1}{2} \Delta \right) \right).$$

Так как  $\Delta \rightarrow 0$  при  $n \rightarrow \infty$ , то получим окончательно:

$$\begin{aligned} \mathcal{I} &= \int_a^b \cos x \, dx = \\ &= \lim_{\Delta \rightarrow 0} \frac{\Delta}{2 \sin \frac{\Delta}{2}} \left( \sin \left( b + \frac{1}{2} \Delta \right) - \sin \left( a + \frac{1}{2} \Delta \right) \right) = \\ &= \sin b - \sin a. \quad \triangle \end{aligned}$$

Если точным методом не удаётся воспользоваться, то применяют *приближённые методы* вычисления интегралов.

## 1.4 Интегрируемость функций

Но прежде чем обсуждать приближённые методы, рассмотрим вопрос об интегрируемости функций. Это непростая проблема [12]. Изучим её на примере монотонной и неубывающей функции  $f(x)$ , заданной на отрезке  $[a; b]$ :

$$f(x_1) \leq f(x_2) \quad \text{при} \quad a \leq x_1 \leq x_2 \leq b. \quad (14)$$

### 1.4.1 Суммы Дарбу и их свойства

По примеру Дарбу применим вспомогательное средство исследования — две дополнительные суммы, схожие с интегральной суммой, но несколько проще её. Для этого выполним произвольное разбиение  $Z(x_i)$  отрезка интегрирования. В силу монотонности функции  $f$  для каждого элементарного отрезка  $[x_{i-1}; x_i]$  ( $i = 1, 2, \dots, n$ ) имеем двойное неравенство

$$f(x_{i-1}) \leq f(x) \leq f(x_i), \quad x \in [x_{i-1}; x_i]. \quad (15)$$

Оно позволяет образовать две специальные интегральные суммы

$$\begin{aligned} s(x_i) &= I(x_i; x_{i-1}) = \\ &= f(x_0)\Delta x_1 + f(x_1)\Delta x_2 + \dots + f(x_{n-1})\Delta x_n, \\ S(x_i) &= I(x_i; x_i) = \\ &= f(x_1)\Delta x_1 + f(x_2)\Delta x_2 + \dots + f(x_n)\Delta x_n, \end{aligned} \quad (16)$$

которые и называются соответственно *нижней* и *верхней интегральными суммами*, или *суммами Дарбу*. В первой из них в роли точек  $\xi_i$  взяты левые концы отрезков  $[x_{i-1}; x_i]$ , а во второй — правые. На основании (15) заключаем, что любая интегральная сумма (10)  $I(x_i; \xi_i)$  для того

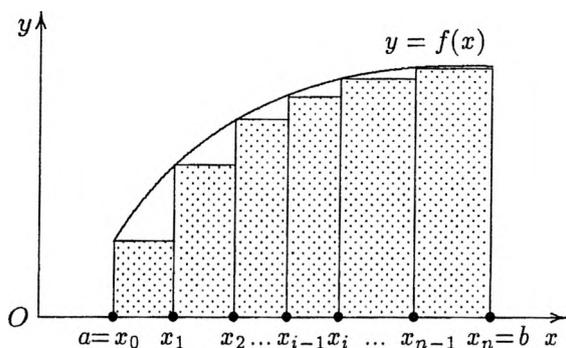


Рис. 5.

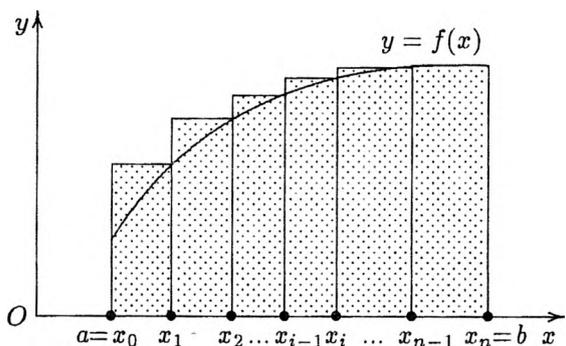


Рис. 6.

же разбиения удовлетворяет неравенству

$$s(x_i) \leq I(x_i; \xi_i) \leq S(x_i). \quad (17)$$

Геометрически нижняя сумма Дарбу равна площади ступенчатого многоугольника, содержащегося в криволинейной трапеции (рис. 5), а верхняя — площади ступенчатого многоугольника, содержащего внутри себя криволинейную трапецию (рис. 6).

Следуя Дарбу, найдём разность между верхней и нижней суммами:

$$S(x_i) - s(x_i) = (f(x_1) - f(x_0))\Delta x_1 + \\ + (f(x_2) - f(x_1))\Delta x_2 + \cdots + (f(x_n) - f(x_{n-1}))\Delta x_n. \quad (18)$$

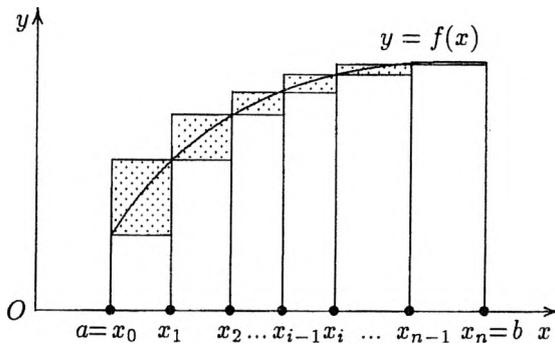


Рис. 7.

Как видим, она равна сумме площадей прямоугольников, получившихся в результате вычитания друг из друга ступенчатых многоугольников (см. рис. 6 и 5). Заметьте, что эти прямоугольники, как бусинки, нанизаны на график функции  $f(x)$  (рис. 7).

Суммы Дарбу обладают двумя замечательными свойствами. Докажем следующие леммы.

**ЛЕММА 1.** Пусть разбиение  $Z_1$  получено из разбиения  $Z$  добавлением нескольких новых точек. Тогда нижняя и верхняя суммы Дарбу для функции  $f(x)$ , соответствующие этим разбиениям, удовлетворяют неравенствам

$$s(x_i) \leq s_1(x_i) \leq S_1(x_i) \leq S(x_i). \quad (19)$$

Эти неравенства означают, что, если к имеющимся точкам деления добавлять новые точки, то в результате нижняя сумма Дарбу станет лишь возрастать, а верхняя — разве что уменьшаться. В итоге при таком измельчении разбиения эти суммы будут сближаться, т. е.

$$S(x_i) - s(x_i) \geq S_1(x_i) - s_1(x_i).$$

*Доказательство.* Сначала изучим поведение верхней суммы. Для доказательства достаточно получить из разбиения  $Z$  разбиение  $Z_1$  присоединением к уже имеющимся точкам деления ещё одной точки  $x^*$ , принадлежащей отрезку

$[x_{i_0-1}; x_{i_0}]$ , т. е. такой, что  $x_{i_0-1} < x^* < x_{i_0}$ . В верхней сумме  $S(x_i)$ , соответствующей разбиению  $Z(x_i)$ , выделим слагаемое с номером  $i_0$ :

$$f(x_{i_0})\Delta x_{i_0} = f(x_{i_0})(x_{i_0} - x_{i_0-1}). \quad (20)$$

Переход от разбиения  $Z$  к разбиению  $Z_1$  приведёт к тому, что вместо одного слагаемого (20) в сумме  $S_1(x_i)$  появятся два новых слагаемых:

$$f(x^*)(x^* - x_{i_0-1}) + f(x_{i_0})(x_{i_0} - x^*). \quad (21)$$

Остальные же слагаемые сумм  $S(x_i)$  и  $S_1(x_i)$  одинаковы.

Сравним (20) и (21). В силу монотонности (14) функции  $f(x)$  выполняется неравенство  $f(x^*) \leq f(x_{i_0})$ . Значит, имеет место неравенство

$$f(x^*)(x^* - x_{i_0-1}) + f(x_{i_0})(x_{i_0} - x^*) \leq f(x_{i_0})(x_{i_0} - x_{i_0-1}),$$

означающее, что замена слагаемого (20) суммы  $S(x_i)$  двумя слагаемыми (21) в сумме  $S_1(x_i)$  не может увеличить этой суммы в сравнении с  $S(x_i)$ .

Для нижней суммы Дарбу доказательство выполняется аналогично. Если желательно, чтобы разбиение  $Z_1$  отличалось от разбиения  $Z$  несколькими точками деления, то следует постепенно добавлять их по одной. Так как каждое новое добавление точки не уменьшает нижнюю сумму и не увеличивает верхнюю, то за конечное число шагов будет совершён переход от разбиения  $Z$  к разбиению  $Z_1$ , а в итоге для сумм Дарбу имеют место справедливые неравенства (19).  $\square$

**ЛЕММА 2.** Пусть  $Z_1$  и  $Z_2$  — произвольные разбиения отрезка  $[a; b]$ , а  $s_1(x_i)$  и  $S_2(x_i)$  — соответственно нижняя сумма Дарбу для разбиения  $Z_1$  и верхняя сумма для разбиения  $Z_2$ , соответствующие функции  $f(x)$ . Тогда справедливо неравенство

$$s_1(x_i) \leq S_2(x_i). \quad (22)$$

Другими словами, неравенство (22) означает, что каждая нижняя сумма Дарбу не превосходит каждую верхнюю сумму, хотя бы отвечающую и другому разбиению отрезка.

*Доказательство.* Для одного и того же разбиения справедливость утверждения (22) следует из двойного неравенства (15). Рассмотрим третье, вспомогательное разбиение  $Z$ , содержащее все точки разбиений  $Z_1$  и  $Z_2$ . Пусть ему отвечают нижняя  $s(x_i)$  и верхняя  $S(x_i)$  суммы Дарбу. Тогда на основании леммы 1 напишем цепочку неравенств

$$s_1(x_i) \leq s(x_i) \leq S(x_i) \leq S_2(x_i).$$

Опустив в ней внутренние неравенства, получим (22).  $\square$

### 1.4.2 Классы интегрируемых функций

Укажем один *простой класс интегрируемых функций*. С помощью сумм Дарбу это сделать нетрудно.

**ТЕОРЕМА 1.** *Если функция  $f(x)$  монотонна на отрезке  $[a; b]$ , то она интегрируема на этом отрезке.*

*Доказательство.* Теорему докажем в три этапа. Сначала исследуем *случай равномерного разбиения отрезка  $[a; b]$* .

1. Итак, построим последовательность разбиений  $Z_n$  следующим образом: первое разбиение  $Z_1$  осуществим делением отрезка  $[a; b]$  пополам, второе  $Z_2$  — делением на четыре равные части, третье  $Z_3$  — на восемь равных частей и т. д. Для длин отрезков разбиения  $Z_n$  имеем в результате

$$\Delta_n = \frac{b-a}{2^n} \rightarrow 0 \quad \text{при} \quad n \rightarrow \infty. \quad (23)$$

В силу равномерности разбиений  $Z_n$  верхние и нижние суммы Дарбу примут простой вид:

$$\begin{aligned} S_n(x_i) &= (f(x_1) + f(x_2) + \cdots + f(x_n)) \Delta_n, \\ s_n(x_i) &= (f(x_0) + f(x_1) + \cdots + f(x_{n-1})) \Delta_n, \end{aligned}$$

а их разность равна

$$S_n(x_i) - s_n(x_i) = (f(x_n) - f(x_0))\Delta_n = (f(b) - f(a))\Delta_n,$$

так что, используя (23), получим

$$\lim_{n \rightarrow \infty} (S_n(x_i) - s_n(x_i)) = 0. \quad (24)$$

Так как каждая новая последовательность разбиений содержит в себе все точки деления предыдущих разбиений, то, учитывая ещё и леммы 1 и 2, приходим к выводу, что последовательности верхних и нижних сумм суть монотонные ограниченные последовательности:

$$\begin{aligned} S_n(x_i) &\geq S_{n+1}(x_i), & S_n(x_i) &\geq s_1, \\ s_n(x_i) &\leq s_{n+1}(x_i), & s_n(x_i) &\leq S_1. \end{aligned}$$

По теореме о монотонных ограниченных числовых последовательностях они должны иметь предел. Согласно (24) их пределы совпадают:

$$\lim_{n \rightarrow \infty} S_n(x_i) = \lim_{n \rightarrow \infty} s_n(x_i) = \mathcal{I}. \quad (25)$$

При этом последовательность  $S_n(x_i)$  стремится к пределу  $\mathcal{I}$  сверху, а последовательность  $s_n(x_i)$  — снизу:

$$S_n(x_i) \geq \mathcal{I} \geq s_n(x_i). \quad (26)$$

Выводом соотношений (24), (25) и (26) закончим первый этап доказательства теоремы.

Проанализируем полученный результат. Он имеет очевидную геометрическую интерпретацию. Выше показано, что при неограниченном измельчении равномерных разбиений отрезка  $[a; b]$  последовательности площадей ступенчатых многоугольников, содержащихся в криволинейной трапеции и окаймляющих её, стремятся к одному и тому же пределу,  $\mathcal{I}$  — к площади самой криволинейной трапеции. К этому пределу стремятся и интегральные суммы, заключённые между нижними и верхними суммами Дарбу (см. (17)).

Обобщим этот результат на *общий случай разбиений* (не обязательно равномерных).

2. На этом этапе убедимся в том, что верхняя и нижняя суммы Дарбу, соответствующие произвольному разбиению  $Z$ , связаны неравенством

$$S(x_i) \geq I \geq s(x_i). \quad (27)$$

Сначала отработаем с верхней суммой. Допустим противное: пусть найдётся такое разбиение  $Z^*$ , для которого  $S^*(x_i) < I$ . Положим  $\varepsilon = I - S^*(x_i) > 0$ . На основании (25) и (26) для данного  $\varepsilon$  можно указать такой номер  $N$ , что член последовательности нижних сумм  $s_N$ , построенной на первом этапе доказательства, удовлетворяет неравенству

$$0 \leq I - s_N(x_i) < \varepsilon = I - S^*(x_i).$$

Используя его, приходим к неравенству  $S^*(x_i) < s_N(x_i)$ , противоречащему лемме 2. Выявленное противоречие доказывает справедливость неравенства (27) для верхней суммы. Для нижней суммы Дарбу доказательство выполняется аналогично.

Неравенство (27) имеет простой геометрический смысл: площадь криволинейной трапеции  $I$  меньше площади любого окаймляющего ступенчатого многоугольника  $S(x_i)$  и больше площади любого содержащегося в ней ступенчатого многоугольника  $s(x_i)$ . Отметим ещё, что этот результат, в отличие от (26), никак не связан с предположением о равномерном разбиении отрезка  $[a; b]$ .

И, наконец, убедимся в том, что число  $I$  — предел интегральных сумм (10)  $I(x_i; \xi_i)$  при  $\Delta \rightarrow 0$ .

3. Выполним произвольное разбиение  $Z(x_i)$ . Для него интегральная сумма  $I(x_i; \xi_i)$ , суммы Дарбу  $S(x_i)$ ,  $s(x_i)$  и число  $I$  удовлетворяют условиям (17) и (27). Вычитая (17) из (27), получим двойное неравенство

$$-\left(S(x_i) - s(x_i)\right) \leq I - I(x_i; \xi_i) \leq S(x_i) - s(x_i).$$

Заменяем его эквивалентным неравенством

$$|\mathcal{I} - I(x_i; \xi_i)| \leq S(x_i) - s(x_i). \quad (28)$$

Последнее соотношение означает следующее: чтобы оценить разность между числом  $\mathcal{I}$  и произвольной интегральной суммой  $I(x_i; \xi_i)$ , достаточно оценить разность (18) между верхней и нижней суммами Дарбу.

Зададимся произвольной точностью  $\varepsilon$  и сопоставим ей число  $\delta$ , положив его равным

$$\delta = \frac{\varepsilon}{f(b) - f(a)}. \quad (29)$$

Считаем разбиение  $Z(x_i)$  достаточно мелким. Пусть для него длина наибольшего элементарного отрезка удовлетворяет условию  $\Delta = \max\{x_i\} < \delta$ . Оценим разность (18), заменив в ней все  $\Delta x_i$  ( $i = 1, 2, \dots, n$ ) на  $\delta$ . Тогда имеем

$$\begin{aligned} S(x_i) - s(x_i) &< (f(x_1) - f(x_0) + f(x_2) - f(x_1) + \\ &+ f(x_3) - f(x_2) + \dots + f(x_n) - f(x_{n-1}))\delta = \\ &= (f(x_n) - f(x_0))\delta = \frac{(f(b) - f(a))\varepsilon}{f(b) - f(a)} = \varepsilon, \end{aligned}$$

т. е. получим

$$S(x_i) - s(x_i) < \varepsilon. \quad (30)$$

Неравенство (30) показывает, что для монотонной функции  $f(x)$  разность между верхней и нижней суммами Дарбу стремится к нулю при  $\Delta \rightarrow 0$ . Этот результат обобщает вывод (24), сделанный на первом этапе доказательства данной теоремы. Сравнивая (30) и (28), получим неравенство

$$|\mathcal{I} - I(x_i; \xi_i)| < \varepsilon. \quad (31)$$

Оно утверждает: для любого заданного  $\varepsilon$  найдено такое  $\delta$  (29), что при произвольном разбиении  $Z(x_i)$ , удовлетворяющем условию  $\Delta < \delta$ , справедливо неравенство (31).

А это, по определению (см. (11)), означает, что число  $\mathcal{I}$  является пределом интегральных сумм  $I(x_i; \xi_i)$  при  $\Delta \rightarrow 0$  и монотонная функция  $f(x)$  на отрезке  $[a; b]$  интегрируема на этом отрезке.  $\square$

В заключение отметим, что монотонность функции — достаточное условие интегрируемости, не являющееся, однако, необходимым. Например, доказанная только что теорема легко обобщается на класс кусочно монотонных функций (они в точках  $c_i$  ( $i = 1, 2, \dots, n - 1$ ) меняют убывание на невозрастание или наоборот; такой функцией, скажем, является синус:  $f(x) = \sin x$ ). Если отрезок  $[a; b]$  разбить на части  $[a; c_1], [c_1; c_2], \dots, [c_{n-1}; b]$ , в каждой из которых функция  $f(x)$  монотонна, а значит, и интегрируема, то она будет интегрируема и на всём промежутке  $[a; b]$ .

И наконец, сформулируем два утверждения, которые намеренно оставим без доказательства: *функция, ограниченная на  $[a; b]$  и имеющая конечное число точек разрыва, или функция, непрерывная на  $[a; b]$ , интегрируемы*. Они называют ещё два класса функций, для которых определённый интеграл всегда существует.

## 2. Приближённые методы вычисления интегралов

Приближённые методы вычисления интегралов можно условно разделить на две группы: *аналитические* и *численные*. Эти методы основаны на аппроксимации подынтегральной функции  $f(x)$  некоторой более простой функцией, например, рядом или полиномом.

### 2.1 Аналитические методы

*Аналитические методы* заключаются в приближённом построении первообразной и в дальнейшем применении формулы (13). Практически обычно пользуются разложением подынтегральной функции  $f(x)$  в степенной ряд, что позволяет свести вычисление исходного интеграла к инте-

грированию полинома. *Погрешность результата* в таком случае складывается:

1) из погрешности замены ряда частичной суммой (это погрешность усечения; она равна остатку ряда);

2) из ошибок округления при вычислении самой частичной суммы (для их уменьшения следует вести расчёты с запасными цифрами).

Заметим, что в практике вычислений нередко имеют дело со знакопередающимися рядами. Если при любом  $x$  из промежутка интегрирования общий член такого ряда стремится к нулю и модули членов ряда монотонно убывают, то по признаку Лейбница [7] ошибка усечения всегда меньше модуля первого отброшенного члена. Этим обстоятельством можно воспользоваться для оценки погрешности. В случае знакопостоянных рядов для оценки остатка ряда привлекают в дополнение такие числовые ряды, остатки которых нетрудно подсчитать.

**ПРИМЕР 8.** Вычислим интеграл из примера 6 с точностью до четырёх десятичных знаков, положив  $x = 1$ .

*Решение.* Представим функцию  $f(u) = e^{-u^2}$  степенным рядом (рядом Тейлора) [7]. Воспользуемся разложением экспоненты в ряд

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \frac{x^6}{6!} + \dots,$$

сходящийся при любом  $x$ . Заменяв в нём  $x$  на  $-u^2$ , получим

$$e^{-u^2} = 1 - u^2 + \frac{u^4}{2!} - \frac{u^6}{3!} + \frac{u^8}{4!} - \frac{u^{10}}{5!} + \frac{u^{12}}{6!} - \dots$$

Так как  $0 \leq u \leq 1$ , то модули членов ряда убывают, и для оценки остатка ряда при любом  $u$  из промежутка интегрирования можно воспользоваться признаком Лейбница. Таким образом, если удержать члены ряда

до  $(-1)^{n+1} \frac{u^{2n-1}}{(2n-1)(n-1)!}$ , то для остатка получим оценку

$$|R_{n+1}| \leq \frac{u^{2n+1}}{(2n+1)n!} \Big|_{u=1} < 0.00005.$$

Это неравенство выполняется уже при  $n = 7$ :

$$|R_7| \leq \frac{u^{15}}{15 \cdot 7!} \Big|_{u=1} = \frac{1}{75600} \approx 1.5 \cdot 10^{-5} < 0.00005.$$

Следовательно, проинтегрируем почленно первые семь членов ряда:

$$\begin{aligned} \mathcal{I} &= \frac{2}{\sqrt{\pi}} \int_0^1 e^{-u^2} du = \\ &= \frac{2}{\sqrt{\pi}} \int_0^1 \left( 1 - u^2 + \frac{u^4}{2!} - \frac{u^6}{3!} + \frac{u^8}{4!} - \frac{u^{10}}{5!} + \frac{u^{12}}{6!} \right) du = \\ &= \frac{2}{\sqrt{\pi}} \left( u - \frac{u^3}{3 \cdot 1!} + \frac{u^5}{5 \cdot 2!} - \frac{u^7}{7 \cdot 3!} + \frac{u^9}{9 \cdot 4!} - \frac{u^{11}}{11 \cdot 5!} + \frac{u^{13}}{13 \cdot 6!} \right) \Big|_0^1 = \\ &= \frac{2}{\sqrt{\pi}} \left( 1 - \frac{1}{3} + \frac{1}{10} - \frac{1}{42} + \frac{1}{216} - \frac{1}{1320} + \frac{1}{9360} \right). \end{aligned}$$

Вычислим сумму ряда с пятью знаками после точки (с одним запасным знаком). Окончательно получим  $\mathcal{I} = 0.8427$  со всеми верными знаками (значение интеграла, найденное с точностью  $10^{-15}$ , равно 0.842700792949715).

Заметим, что практически нет необходимости заранее определять число  $n$ . Проще подсчитывать члены ряда при соответствующем значении  $u$  до тех пор, пока не дойдём до слагаемого, модуль которого меньше заданной погрешности.  $\triangle$

Однако не всякую подынтегральную функцию (или её часть) можно разложить в ряд. Поэтому не будем больше останавливаться на обсуждении аналитических методов.

Универсальными являются *алгоритмы численного интегрирования* (их больше трёх десятков [11]). Они позволяют найти значение интеграла по значениям функции  $f(x)$  на отрезке  $[a; b]$ .

## 2.2 Численные методы

Приведённое выше определение интеграла (см. (12)) даёт одновременно и способ его вычисления путём прямого нахождения предела интегральных сумм. Но такой способ, как мы уже знаем, сложен и его (почти) не применяют в точных вычислениях. Зато в приближённых алгоритмах он часто используется. При этом подынтегральная функция  $f(x)$  заменяется легко интегрируемой функцией, например, такой, как полиномиальная или рациональная функция.

Ниже выводятся формулы прямоугольников, трапеций и парабол. В них  $f(x)$  заменяется соответственно полиномом нулевой, первой и второй степени. Напомним, что сам процесс численного нахождения интеграла называется *квадратурой*, а соответствующие формулы — *квадратурными* (буквально — *формулами вычисления площади*).

### 2.2.1 Формулы прямоугольников

Примем следующий порядок действий: будем исходить из *определения интеграла как предела интегральных сумм*; выполним произвольное и достаточно мелкое разбиение отрезка  $[a; b]$  и, образовав для него интегральную сумму (10), возьмём её значение в роли значения соответствующего интеграла.

Итак, разобьём отрезок  $[a; b]$  точками деления  $x_i$  ( $i = 0, 1, \dots, n$ ;  $x_0 = a$ ,  $x_n = b$ ) на  $n$  *равных частей* длины  $h = (b - a)/n$ . На каждом частичном отрезке возьмём по произвольной точке  $\xi_i \in [x_{i-1}; x_i]$  ( $i = 1, 2, \dots, n$ ). Тогда выражение для интегральной суммы запишется в виде

$$I_n = h(f(\xi_1) + f(\xi_2) + \dots + f(\xi_n)).$$

Если функция  $f(x)$  интегрируема на отрезке  $[a; b]$ , то получим

$$\lim_{n \rightarrow \infty} I_n = \mathcal{I} = \int_a^b f(x) dx,$$

что даёт нам право формулу для интеграла  $\mathcal{I}$  представить в виде суммы

$$\mathcal{I} = I_n + R_n, \quad \text{где} \quad \lim_{n \rightarrow \infty} R_n = 0. \quad (32)$$

Осталось пренебречь величиной остаточного члена  $R_n$ , чтобы написать приближённую формулу для вычисления  $\mathcal{I}$ . Она называется *формулой прямоугольников*:

$$\mathcal{I} \approx I_n = h(f(\xi_1) + f(\xi_2) + \dots + f(\xi_n)). \quad (33)$$

Причина такого названия кроется в геометрическом истолковании формулы (33): интеграл  $\mathcal{I}$  есть площадь криволинейной трапеции (рис. 2), а интегральная сумма — площадь ступенчатого многоугольника, составленного из прямоугольников (рис. 5, 3 или 6) с одинаковыми основаниями длины  $h = (b - a)/n$  и высотами, равными значениям функции  $f(x)$  в произвольных точках  $\xi_i$  отрезков разбиения  $[x_{i-1}; x_i]$  ( $i = 1, 2, \dots, n$ ). Разность между этими площадями  $R_n$  с возрастанием  $n$  стремится к нулю.

Если в роли точек  $\xi_i$  взять левые концы  $\xi_i = x_{i-1}$ , средние точки  $\xi_i = (x_{i-1} + x_i)/2$  или правые концы  $\xi_i = x_i$  отрезков разбиения, то получим соответственно *формулы левых* ( $\xi_i = a + (i-1)h$ ) ( $i = 1, 2, \dots, n$ ), *средних* ( $\xi_i = a + (i-1/2)h$ ) и *правых* ( $\xi_i = a + ih$ ) *прямоугольников*:

$$\mathcal{I} \approx I_n = h \sum_{i=1}^n f(\xi_i). \quad (34)$$

Из трёх формул в практике вычислений используется *формула средних прямоугольников*. Она даёт хорошие результаты при интегрировании непрерывных периодических

функций. Если функция  $f(x)$  имеет на отрезке  $[a; b]$  непрерывную вторую производную, то всегда найдётся такая точка  $\zeta \in [a; b]$ , что погрешность формулы средних прямоугольников [7] будет представима в виде<sup>9</sup>

$$R_n = \frac{(b-a)^3}{24n^2} f''(\zeta) = \frac{b-a}{24} h^2 f''(\zeta). \quad (35)$$

Из равенства (35) следует, во-первых, что формула средних прямоугольников даёт точный результат, если подынтегральная функция является полиномом степени не выше первой, и, во-вторых, если  $f''(x) > 0 \forall x \in [a; b]$ , то по формуле будет найдено значение интеграла с недостатком; если же  $f''(x) < 0$  — с избытком. Так как местоположение точки  $\zeta$  на отрезке  $[a; b]$  не известно, то равенство (35) не позволяет подсчитать погрешность  $R_n$ , но даёт возможность оценить её сверху. Обозначим через  $M_2 = \sup_{x \in [a; b]} |f''(x)|$ , а через  $\varepsilon$  — заданную точность вычислений, получим

$$|R_n| = \frac{(b-a)^3}{24n^2} M_2 \leq \varepsilon. \quad (36)$$

Неравенство (36) показывает, что с возрастанием  $n$  погрешность формулы средних прямоугольников убывает не медленнее, чем  $1/n^2$ . Это неравенство позволяет выбрать заранее начальное число делений отрезка  $[a; b]$  и, таким образом, найти значение интеграла с погрешностью, не превосходящей по модулю  $\varepsilon$ . Вот это число<sup>10</sup>:

$$n = \sqrt{\frac{(b-a)^3}{24\varepsilon} M_2}. \quad (37)$$

<sup>9</sup>Для формул левых и правых прямоугольников имеем [7] соответственно  $R_n \neq \frac{(b-a)^2}{2n} f'(\zeta) = \frac{b-a}{2} h f'(\zeta)$  и  $R_n = -\frac{(b-a)^2}{2n} f'(\zeta)$ .

<sup>10</sup>Для формул левых и правых прямоугольников  $n = \frac{(b-a)^2}{2\varepsilon} M_1$ , где  $M_1 = \sup_{x \in [a; b]} |f'(x)|$ .

**Упражнение 1.1.** Докажите интегрируемость кусочно монотонных функций.

2. Вычислите интегралы с помощью интегральных сумм:

$$1) \int_a^b x^k dx \quad (a, b \in \mathbf{R}, k \in \mathbf{N}); \quad 3) \int_a^b \sin x dx;$$

$$2) \int_a^b x^p dx \quad (a, b \in \mathbf{R}^+, p \in \mathbf{R}); \quad 4) \int_0^\pi \ln(1 - 2r \cos x + r^2) dx.$$

3. Какая связь между суммами Дарбу и формулами прямоугольников?

4. Модифицируйте формулу левых и средних прямоугольников, изменив пределы индекса суммирования  $i$ .

5. Получите формулы прямоугольников, заменив на отрезке  $[x_{i-1}; x_i]$  кривую  $y = f(x)$  прямой  $y = y_{i-1}$ ,  $y = y_{i-\frac{1}{2}}$  или  $y = y_i$ , проходящей соответственно через точки  $(x_{i-1}, y_{i-1})$ ,  $(x_{i-\frac{1}{2}}, y_{i-\frac{1}{2}})$  или  $(x_i, y_i)$ . Здесь  $x_{i-\frac{1}{2}} = \frac{x_{i-1} + x_i}{2}$ ,  $y_{i-\frac{1}{2}} = f(x_{i-\frac{1}{2}})$  ( $i = 1, 2, \dots, n$ ).

6. Когда верны формулы левых и правых прямоугольников?

### 2.2.2 Формула трапеций

Теперь применим другой подход к выводу формул численного интегрирования. Выберем простую вспомогательную функцию, близкую к подынтегральной функции  $f(x)$ , и приближённо заменим интеграл  $I = \int_a^b f(x) dx$  интегралом от вспомогательной функции.

Представим простейшую реализацию этой идеи. Разобьём отрезок  $[a; b]$  на  $n$  равных частей длины  $h = (b-a)/n$  точками  $x_i = a + ih$  ( $i = 0, 1, \dots, n$ ),  $x_0 = a$ ,  $x_n = b$ . Ясно,

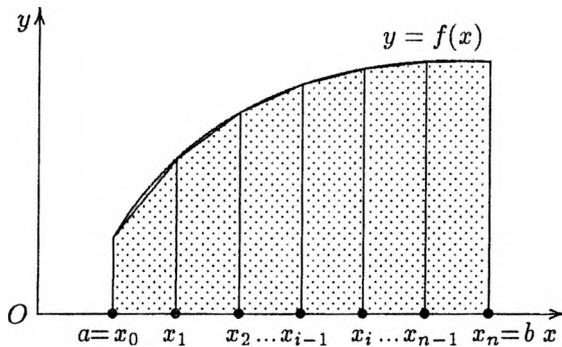


Рис. 8.

что искомый интеграл равен сумме интегралов, т. е.

$$\int_a^b f(x) dx = \int_{x_0}^{x_1} f(x) dx + \int_{x_1}^{x_2} f(x) dx + \dots + \int_{x_{n-1}}^{x_n} f(x) dx. \quad (38)$$

На каждом элементарном отрезке  $[x_{i-1}; x_i]$  заменим подынтегральную функцию линейной функцией вида

$$y = A(x - x_{i-1}) + B(x - x_i), \quad (39)$$

$$x \in [x_{i-1}; x_i] \quad (i = 1, 2, \dots, n).$$

Она в граничных точках  $x_{i-1}, x_i$  принимает те же значения, что и функция  $f(x)$ :  $y_{i-1} = f(x_{i-1})$  и  $y_i = f(x_i)$ . Её график является ломаной линией, начальная, угловые и конечная точки которой принадлежат также графику подынтегральной функции  $f(x)$  (рис. 8). С увеличением  $n$  число общих точек будет расти, а в результате ломаная станет всё больше приближаться к кривой  $y = f(x)$ .

Положив в (39)  $x = x_i$  и  $x = x_{i-1}$ , найдём коэффициенты  $A = \frac{y_i}{h}$  и  $B = \frac{-y_{i-1}}{h}$ . Затем по формуле Ньютона — Лейбница вычислим интеграл от функции (39) по отрезку

ку  $[x_{i-1}; x_i]$ :

$$\begin{aligned}
 & \int_{x_{i-1}}^{x_i} (A(x - x_{i-1}) + B(x - x_i)) dx = \\
 & = \int_{x_{i-1}}^{x_{i-1}+h} A(x - x_{i-1}) dx + \int_{x_{i-1}}^{x_{i-1}+h} B(x - x_{i-1} - h) dx = \\
 & = A \left( \frac{x^2}{2} - x_{i-1}x \right) \Big|_{x_{i-1}}^{x_{i-1}+h} + B \left( \frac{x^2}{2} - x_{i-1}x - hx \right) \Big|_{x_{i-1}}^{x_{i-1}+h} = \\
 & = \frac{(A - B)}{2} h^2 = \frac{f(x_{i-1}) + f(x_i)}{2} h. \quad (40)
 \end{aligned}$$

Полученный результат имеет простой геометрический смысл: интеграл равен площади фигуры, ограниченной графиком функции (39), осью абсцисс и двумя вертикальными линиями  $x = x_{i-1}$ ,  $x = x_i$ . Эта фигура представляет собой трапецию, а, как известно, её площадь равна произведению полусуммы оснований  $(y_{i-1} + y_i)/2$  на высоту  $h$ . Но здесь результат (40) получен прямым интегрированием, а не ссылкой на формулу планиметрии.

Подсчитаем теперь интеграл  $\mathcal{I} = \int_a^b f(x) dx$  на всём промежутке:

$$\begin{aligned}
 & \int_{x_0}^{x_1} f(x) dx \approx \frac{y_0 + y_1}{2} h, \\
 & \int_{x_1}^{x_2} f(x) dx \approx \frac{y_1 + y_2}{2} h, \\
 & \dots \dots \dots \\
 & \int_{x_{n-1}}^{x_n} f(x) dx \approx \frac{y_{n-1} + y_n}{2} h.
 \end{aligned} \quad (41)$$

Складывая почленно равенства (41), получим приближен-

ную формулу для вычисления интеграла

$$\begin{aligned} \mathcal{I} &= \int_a^b f(x) dx \approx I_n = \\ &= h \left( \frac{1}{2}f(a) + f(x_1) + f(x_2) + \cdots + f(x_{n-1}) + \frac{1}{2}f(b) \right) = \\ &= h \left( \frac{f(a) + f(b)}{2} + \sum_{i=1}^{n-1} f(x_i) \right), \quad (42) \end{aligned}$$

называемую *формулой трапеций*. Как и в предыдущем случае, это название связано с геометрической интерпретацией формулы. Разность между площадью криволинейной трапеции, соответствующей интегралу  $\mathcal{I}$ , и суммой площадей обычных трапеций, образованных графиком функции (39), осью абсцисс и вертикалями  $x_i = a + ih$  ( $i = 0, 1, \dots, n$ ), равна [7]

$$R_n = -\frac{(b-a)^3}{12n^2} f''(\zeta) = -\frac{b-a}{12} h^2 f''(\zeta). \quad (43)$$

Она с возрастанием  $n$  стремится к нулю. Заметим, что *выражения для остаточных членов (43) и (35) схожи*. Это позволяет сделать вывод: формулы трапеций и средних прямоугольников характеризуются примерно одинаковой точностью.

Формула (43) написана в предположении, что подынтегральная функция  $f(x)$  имеет на отрезке  $[a; b]$  непрерывную вторую производную. Задавшись точностью  $\varepsilon$  и вычислив  $M_2 = \sup_{x \in [a; b]} |f''(x)|$ , приведём оценку сверху для модуля погрешности:

$$|R_n| = \frac{(b-a)^3}{12n^2} M_2 \leq \varepsilon. \quad (44)$$

С её помощью можно подсчитать *рекомендуемое число де-*

лений отрезка  $[a; b]$ :

$$n = \sqrt{\frac{(b-a)^3}{12\varepsilon}} M_2. \quad (45)$$

**Упражнение 2.** 1. Выведите формулу трапеций:

1) из формул левых и правых прямоугольников;

2) разбейте промежуток интегрирования на  $n$  равных частей и интеграл на каждом элементарном отрезке считайте равным произведению длины отрезка на среднее арифметическое значений подинтегральной функции в начале и в конце отрезка.

2. Получите формулу трапеций, заменив на отрезке  $[x_{i-1}; x_i]$  кривую  $y = f(x)$  прямой:

1)  $y = Ax + B$ , проходящей через точки  $(x_{i-1}, y_{i-1})$  и  $(x_i, y_i)$  ( $i = 1, 2, \dots, n$ ). Коэффициенты  $A$  и  $B$  найдите из условия

$$\begin{cases} y_{i-1} = Ax_{i-1} + B, \\ y_i = Ax_i + B; \end{cases}$$

2)  $y = y_{i-1} + \frac{y_i - y_{i-1}}{h}(x - x_{i-1})$ ,  $h = \frac{b-a}{n}$ ,  $x \in [x_{i-1}; x_i]$  ( $i = 1, 2, \dots, n$ ).

3. Когда верна формула трапеций?

4. Почему погрешность формулы трапеций в два раза (по модулю) больше погрешности формулы средних прямоугольников, хотя аппроксимация подинтегральной функции на каждом элементарном отрезке в методе трапеций проводится полиномом первой, а не нулевой степени?

5. Докажите, что через три точки  $P_0(x_0, y_0)$ ,  $P_1(x_1, y_1)$  и  $P_2(x_2, y_2)$  проходит единственная парабола  $y = ax^2 + bx + c$ .

### 2.2.3 Формула парабол

Будем развивать идею, использованную при выводе формулы трапеций, для получения более точной формулы численного интегрирования. Если на каждом элементарном

отрезке интегрирования *подынтегральную функцию заменить квадратичной функцией*, то получим формулу парабол. Суммируя все интегралы, выведем обобщённую формулу парабол, называемую также *формулой Симпсона*.

Итак, разобьём отрезок  $[a; b]$  на чётное число  $m = 2n$  равных частей длины  $h = (b - a)/(2n)$ . Границы элементарных отрезков суть  $x_0 = a, x_1, x_2, \dots, x_{2n} = b$ , так что  $x_1 = x_0 + h, x_2 = x_1 + h = x_0 + 2h, \dots, x_{2n} = x_{2n-1} + h = x_0 + 2nh$ . Как договаривались, кривую  $y = f(x)$  на отрезке  $[x_{2i-2}; x_{2i}]$  длины  $2h$  аппроксимируем параболой

$$y = A(x - x_{2i-2})(x - x_{2i-1}) + B(x - x_{2i-2})(x - x_{2i}) + C(x - x_{2i-1})(x - x_{2i}), \quad (46)$$

проходящей через точки  $(x_{2i-2}, y_{2i-2}), (x_{2i-1}, y_{2i-1})$  и  $(x_{2i}, y_{2i})$  ( $i = 1, 2, \dots, n$ ) и, таким образом, площадь под кривой приближённо заменим площадью под этой параболой.

Найдём коэффициенты  $A, B$  и  $C$ . Для этого в формуле (46) последовательно положим  $x = x_{2i} = x_{2i-2} + 2h, x = x_{2i-1} = x_{2i-2} + h$  и  $x = x_{2i-2}$ . Получим соответственно

$$A = \frac{y_{2i}}{2h^2}, \quad B = -\frac{y_{2i-1}}{h^2}, \quad C = \frac{y_{2i-2}}{2h^2}. \quad (47)$$

Площадь под параболой есть<sup>11</sup>

$$\int_{x_{2i-2}}^{x_{2i-2}+2h} \left( A(x - x_{2i-2})(x - x_{2i-1}) + B(x - x_{2i-2})(x - x_{2i}) + C(x - x_{2i-1})(x - x_{2i}) \right) dx = A \frac{2}{3} h^3 - B \frac{4}{3} h^3 + C \frac{2}{3} h^3.$$

Поэтому имеем  $\int_{x_{2i-2}}^{x_{2i}} f(x) dx \approx A \frac{2}{3} h^3 - B \frac{4}{3} h^3 + C \frac{2}{3} h^3$ . Исполни-

<sup>11</sup>Здесь удобны замена переменной интегрирования  $x - x_{2i-1} = t$  и учёт формул  $x_{2i-2} = x_{2i-1} - h, x_{2i} = x_{2i-1} + h$ .

зую (47), найдём

$$\int_{x_{2i-2}}^{x_{2i}} f(x) dx \approx \frac{h}{3}(y_{2i-2} + 4y_{2i-1} + y_{2i}). \quad (48)$$

Положив в (48)  $i = 1, 2, \dots, n$ , для интеграла по всему отрезку  $[a; b]$  получим

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

Откуда напомним

$$\mathcal{I} = \int_a^b f(x) dx \approx I_m = \\ = \frac{h}{3} \left( y_0 + y_{2n} + 4(y_1 + y_3 + \dots + y_{2n-1}) + \right. \\ \left. + 2(y_2 + y_4 + \dots + y_{2n-2}) \right). \quad (49)$$

Формула (49) называется *формулой парабол*. Напомним, что в ней шаг интегрирования  $h$  равен  $(b - a)/(2n)$ .

А это *компактная форма* записи формулы парабол:

$$\mathcal{I} = \frac{h}{3} \left( f(a) + f(b) + 4 \sum_{i=1}^n f(x_{2i-1}) + 2 \sum_{i=1}^{n-1} f(x_{2i}) \right) = \quad (50) \\ = \frac{h}{3} \left( f(a) + f(b) + 4f(x_{2n-1}) + 2 \sum_{i=1}^{n-1} (2f(x_{2i-1}) + f(x_{2i})) \right).$$

Именно формулу парабол применяют для приближённого вычисления интегралов, так как она при одинаковых затратах труда даёт более точный результат, чем формулы средних прямоугольников и трапеций. Пусть функция  $f(x) \forall x \in [a; b]$  имеет непрерывную четвёртую производную. Тогда на данном отрезке найдётся такая точка  $\zeta$ , что

остаточный член формул (49)–(50) запишется в виде [7]

$$R_m = -\frac{(b-a)^5}{180m^4} f^{(4)}(\zeta) = -\frac{b-a}{180} h^4 f^{(4)}(\zeta). \quad (51)$$

Оценим его значение сверху. Имеем:

$$|R_m| = \frac{(b-a)^5}{180m^4} M_4 \leq \varepsilon, \quad (52)$$

где  $M_4 = \sup_{x \in [a; b]} |f^{(4)}(x)|$ , т. е. погрешность формулы парабол с возрастанием  $m$  убывает не медленнее, чем  $1/m^4$ . Таким образом, видно, что формула (49) действительно точнее формул (34) и (42). Из неравенства (52) для *рекомендуемого числа делений отрезка*  $[a; b]$  получим

$$m = \sqrt[4]{\frac{(b-a)^5}{180\varepsilon} M_4}. \quad (53)$$

В заключение отметим *любопытный факт*: формула парабол даст точное значение интеграла при интегрировании полиномов до третьего порядка включительно (см. выражение для остаточного члена (51)).

Проиллюстрируем изложенный выше материал примерами.

**ПРИМЕР 9.** Снова вычислим интеграл  $\mathcal{I} = \int_0^1 e^x dx$  (см. пример 5). Значение  $\mathcal{I}$  по формуле Ньютона — Лейбница равно  $e - 1 \approx 1.718282$ . Найдём теперь его по формулам прямоугольников, трапеций и парабол, выбрав произвольное число делений  $n$ , равное, например, 10.

*Решение.* Для функции  $f(x) = e^x$  имеем  $f^{(k)}(x) = (e^x)^{(k)} = e^x$ . Так как экспоненциальная функция  $e^x$  на отрезке интегрирования  $[0; 1]$  возрастает, то получим  $M_1 = M_2 = M_4 = e$ . Сначала по формулам (34), (42) и (49) вычислим  $I_{10}$ . Результаты сведём в табличку. Затем при  $\varepsilon = 10^{-7}$  оценим остаточные члены  $|R_{10}|$  (см. неравенства (36), (44) и (52)), подсчитаем ошибки  $\Delta_{10} = \mathcal{I} - I_{10}$  всех

пяти формул и рекомендуемые начальные числа делений отрезка  $[0; 1]$  (формулы (37), (45) и (53)):

Метод	$I_{10}$	$ R_{10} $	$\Delta_{10}$	$n$
Лев пр	1.633799	$1.4 \cdot 10^{-1}$	$8.4 \cdot 10^{-2}$	13591410
Прав пр	1.805628	$1.4 \cdot 10^{-1}$	$-8.7 \cdot 10^{-2}$	13591410
Трапец	1.719713	$2.3 \cdot 10^{-3}$	$-1.4 \cdot 10^{-3}$	1506
Средн	1.717566	$1.1 \cdot 10^{-3}$	$7.2 \cdot 10^{-4}$	1064
Параб	1.718283	$1.5 \cdot 10^{-6}$	$-9.5 \cdot 10^{-7}$	20

△

Заметим, что в табличке квадратурные формулы перечислены в порядке возрастания точности<sup>12</sup>.

**ПРИМЕР 10.** Теперь, соблюдая условия примера 9, найдём значение функции ошибок  $\mathcal{I} = \operatorname{erf} x = 2/\sqrt{\pi} \int_0^x e^{-u^2} du$ . Его нельзя вычислить по формуле Ньютона — Лейбница, но зато легко подсчитать посредством формул численного интегрирования. Результат при  $x = 1$ , равный 0.8427008, приведён в примере 8.

*Решение.* Для функции  $f(u) = e^{-u^2}$ , определённой на промежутке  $[0; 1]$ , имеем

$$\begin{aligned}
 f'(u) &= -2ue^{-u^2}, & M_1 &= \left| f' \left( \frac{1}{\sqrt{2}} \right) \right| \approx \\
 f''(u) &= 2e^{-u^2}(2u^2 - 1), & &\approx 0.8577639, \\
 f'''(u) &= -4ue^{-u^2}(2u^2 - 3), & M_2 &= |f''(0)| = 2, \\
 f^{(4)}(u) &= 4e^{-u^2}(4u^4 - 12u^2 + 3), & M_4 &= f^{(4)}(0) = 12, \\
 f^{(5)}(u) &= -8ue^{-u^2}(4u^4 - 20u^2 + 15).
 \end{aligned}$$

Все расчётные данные сведём в следующую табличку:

<sup>12</sup>Формулы левых и правых прямоугольников обладают одинаковой точностью.

Метод	$I_{10}$	$ R_{10} $	$\Delta_{10}$	$n$
Лев пр	0.8776723	$4.3 \cdot 10^{-2}$	$-3.5 \cdot 10^{-2}$	4288820
Прав пр	0.8063451	$4.3 \cdot 10^{-2}$	$3.6 \cdot 10^{-2}$	4288820
Трапец	0.8420087	$1.7 \cdot 10^{-3}$	$6.9 \cdot 10^{-4}$	1292
Средн	0.8430469	$8.3 \cdot 10^{-4}$	$-3.5 \cdot 10^{-4}$	914
Параб	0.8427017	$6.7 \cdot 10^{-6}$	$-9.2 \cdot 10^{-7}$	30

△

Отметим, *во-первых*, что формула средних прямоугольников “примерно на одну треть эффективнее формулы трапеций”, и, *во-вторых*, высокую точность формулы парабол.

**Упражнение 3. 1.** Выведите формулу парабол, заменив на отрезке  $[x_{2i-2}; x_{2i}]$  кривую  $y = f(x)$  параболой  $y = Ax^2 + Bx + C$ , проходящей через точки  $(x_{2i-2}, y_{2i-2})$ ,  $(x_{2i-1}, y_{2i-1})$  и  $(x_{2i}, y_{2i})$  ( $i = 1, 2, \dots, n$ ). Коэффициенты  $A$ ,  $B$  и  $C$  определите (см. гл. III) из условия

$$\begin{cases} y_{2i-2} = Ax_{2i-2}^2 + Bx_{2i-2} + C, \\ y_{2i-1} = Ax_{2i-1}^2 + Bx_{2i-1} + C, \\ y_{2i} = Ax_{2i}^2 + Bx_{2i} + C. \end{cases}$$

2. Получите формулу парабол, заменив на отрезке  $[x_{2i-2}; x_{2i}]$  дугу графика  $y = f(x)$  параболой

$$y = y_{2i-2} + \frac{y_{2i-1} - y_{2i-2}}{h}(x - x_{2i-2}) + \\ + \frac{y_{2i} - 2y_{2i-1} + y_{2i-2}}{2h^2}(x - x_{2i-2})(x - x_{2i-1}),$$

проходящей через точки  $(x_{2i-2}, y_{2i-2})$ ,  $(x_{2i-1}, y_{2i-1})$  и  $(x_{2i}, y_{2i})$  ( $i = 1, 2, \dots, n$ );  $h = (b - a)/(2n)$ .

3. Выведите формулу парабол, используя формулы средних прямоугольников и трапеций ( $h = (b - a)/(2n)$ ):

$$I \approx \frac{2}{3}h \sum_{i=1}^{n-1} f\left(\frac{x_{2i-1} + x_{2i+1}}{2}\right) + \frac{1}{3}h \sum_{i=1}^n \frac{f(x_{2i-2}) + f(x_{2i})}{2}.$$

### 2.2.4 Практическая оценка погрешности квадратурных формул

Рассмотрим вопрос о погрешности квадратурных формул. Теоретические оценки погрешности  $R_n$  (см. (35), (43) и (51)) являются *априорными*, так как их можно произвести до нахождения интеграла. На практике они редко используются, так как, *во-первых*, производная, входящая в формулу  $R_n$ , вычисляется в точке  $\zeta$ , положение которой на  $[a; b]$  неизвестно, а оценка сверху (см. (36), (44) и (52)) обычно приводит к чрезмерному её завышению; *во-вторых*, если производная разрывна, то формулу вообще применять нельзя и, *в-третьих*, функция  $f(x)$  может быть задана таблицей. На практике при подсчёте погрешности чаще всего пользуются принципом Рунге. Он заключается в следующем.

Назовём *порядком точности* квадратурной формулы степень  $k$  шага  $h$ , которой пропорциональна величина остаточного члена. Пусть  $I_h$  и  $I_{2h}$  — приближённые значения интеграла  $\mathcal{I} = \int_a^b f(x) dx$ , вычисленные по выбранной квадратурной формуле соответственно с шагами  $h$  и  $2h$  (другими словами, с числами разбиений  $2n$  и  $n$ ). На основании (35), (43) или (51) предположим, что погрешности полученных значений пропорциональны  $k$ -ой степени шагов, причём коэффициент пропорциональности (назовём его  $M$ ) постоянен и не зависит от длины шагов. Тогда напишем два равенства

$$\mathcal{I} = I_h + Mh^k \quad \text{и} \quad \mathcal{I} = I_{2h} + M(2h)^k. \quad (54)$$

Исключив из (54)  $M$ , получим, что

$$\mathcal{I} = I_h + \delta_h, \quad \text{где} \quad \delta_h = \frac{I_h - I_{2h}}{2^k - 1}. \quad (55)$$

Изложенное выше правило практического оценивания погрешностей называют *принципом Рунге*, а формула  $\delta_h$  —

поправкой Ричардсона. Так как оценка погрешности осуществляется после вычисления интеграла, то она является апостериорной.

Применим теперь принцип Рунге к вычислению интеграла с заданной допустимой *относительной ошибкой*  $\varepsilon$ . Выполним какое-нибудь разбиение промежутка интегрирования и подсчитаем интеграл по квадратурной формуле дважды: с заданным  $n$  и удвоенным  $2n$  числом разбиений. Затем по правилу Рунге для числа разбиений  $2n$  найдём относительную погрешность

$$\Delta_{2n} = \frac{I_{2n} - I_n}{(2^k - 1)I_{2n}}. \quad (56)$$

Здесь  $I_{2n}$  и  $I_n$  — приближённые значения интеграла, вычисленные соответственно при  $2n$  и  $n$  разбиениях отрезка  $[a; b]$ . Ошибку (56) сравним с заданной точностью  $\varepsilon$  и, если окажется, что выполняется неравенство  $|\Delta_{2n}| \leq \varepsilon$ , то процесс разбиений закончим. В противном случае удвоим число разбиений, найдём новое значение интеграла, снова проверим неравенство и т. д. Так как значение интеграла может оказаться близким или равным нулю, то при работе на компьютере используем “линейную форму” условия окончания разбиений отрезка  $[a; b]$

$$|I_{2n} - I_n| \leq (2^k - 1)\varepsilon |I_{2n}|. \quad (57)$$

Здесь  $k = 1$  для формул левых и правых прямоугольников,  $k = 2$  — для формул средних прямоугольников и трапеций,  $k = 4$  — для формулы парабол и  $k = 8$  — для формулы Гаусса — Лежандра ( $n = 4$ ).

В случае *абсолютной ошибки* также избавимся от операции деления. Тогда условие окончания разбиений запишется в виде

$$|I_{2n} - I_n| \leq (2^k - 1)\varepsilon. \quad (58)$$

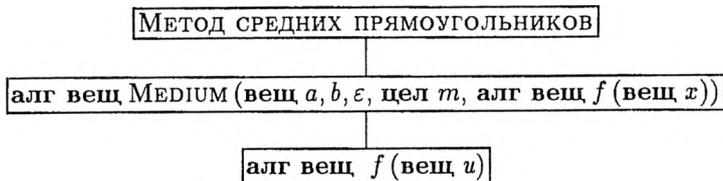


Рис. 9.

ПРИМЕР 11. По формуле средних прямоугольников (34) вычислим значение функции  $\text{erf}(1)$  из примера 10; пусть  $\varepsilon = 10^{-7}$ ,  $n = 10$ .

*Решение.* Построим функциональную схему программы. Так как по условию задачи требуется при заданных пределах интегрирования  $a = 0$ ,  $b = 1$ , допустимой относительной погрешности  $\varepsilon = 10^{-7}$  и начальном числе разбиений промежутка интегрирования  $n = 10$  методом средних прямоугольников ( $\mathcal{I} = \int_a^b f(x) dx \approx h \sum_{i=0}^{n-1} f(a + h/2 + ih)$ ,  $h = (b - a)/n$ ) вычислить значение интеграла  $\text{erf}(1) = 2/\sqrt{\pi} \int_0^1 e^{-u^2} du$ , то три алгоритма в схеме иерархии очевидны (рис. 9): головной алгоритм (назовём его МЕТОД СРЕДНИХ ПРЯМОУГОЛЬНИКОВ), подчинённый — с именем MEDIUM, выполняющий численное интегрирование, и на третьем уровне иерархии алгоритм  $f$ , поставляющий алгоритму MEDIUM значения ординат.

Пояснений требует, наверное, только алгоритм MEDIUM. Он будет содержать кратную команду повторения, причём внешняя (команда пока) станет служить для проверки неравенства Рунге, а внутренняя (команда для) — для суммирования ординат. Каждая из команд повторения потребует подготовки цикла. Начало внешнего цикла будет состоять из копирования числа разбиений ( $n \leftarrow m$ ; вспомогательная переменная  $n$  необходима, так как в результате побочного эффекта ( $n \leftarrow 2n$ ; см. ниже) изменится исходное число разбиений и, таким образом, будет нарушено правило “не ты вычислял, ни тебе и менять”): резуль-

татом работы алгоритма MEDIUM должно быть  $I_{2n}$ , а не изменённое значение  $m$ ), подсчёта начального шага интегрирования ( $h \leftarrow (b - a)/n$ ), задания фиктивного значения интеграла ( $I_n \leftarrow 0$ ; при таком начальном значении  $I_n$  неравенство Рунге не выполнится и будет совершён хотя бы один виток цикла по  $\varepsilon$ , в котором число разбиений промежутка интегрирования увеличится вдвое), подсчёта константы ( $\varepsilon_3 \leftarrow 3\varepsilon$ ; это действие называется *чисткой цикла*) и задания начального значения логической переменной ( $finished \leftarrow ложь$ ; при таком её значении внешний цикл (цикл по  $\varepsilon$ ) выполнится, по крайней мере, один раз).

Начало внутреннего цикла (цикл с параметром  $i$ ) будет содержать команды вычисления вспомогательных величин: половины шага интегрирования ( $h_{half} \leftarrow h/2$ ) и начальных значений абсциссы  $t \leftarrow a + h/2$  и ординаты  $S \leftarrow f(a + h/2)$ . После цикла суммирования ординат ( $S \leftarrow S + f(a + h/2 + ih)$ , ( $i = 1, 2, \dots, n - 1$ )) найдём площадь ступенчатого многоугольника ( $I_{2n} \leftarrow S \cdot h$ ), подсчитаем текущее значение лог переменной  $finished$  ( $finished \leftarrow (|I_{2n} - I_n| \leq 3\varepsilon |I_{2n}|)$ ) и подготовимся к очередному возможному циклу по  $\varepsilon$ : выполним пересылку значения интеграла ( $I_n \leftarrow I_{2n}$ ), уменьшим вдвое шаг интегрирования ( $h \leftarrow h/2$ ) и одновременно увеличим вдвое число разбиений ( $n \leftarrow 2n$ ). Как только неравенство Рунге выполнится ( $finished$  станет равной *истина*), алгоритм MEDIUM закончит свою работу (MEDIUM  $\leftarrow I_{2n}$ ).

Приведём текст проекта программы:

{ \_\_\_\_\_ }

алг МЕТОД СРЕДНИХ ПРЯМОУГОЛЬНИКОВ

вещ  $a, b, \varepsilon, I$

цел  $n$

начало

вывести '  $a, b$  — пределы интегрирования,  $\varepsilon$  — точ-

ность и  $n$  — начальное число разбиений  $\Rightarrow$

ввести  $a, b, \varepsilon, n$

вывести заголовок задачи,  $a, b, \varepsilon, n$

$\mathcal{I} \leftarrow \frac{2}{\sqrt{\pi}} \text{MEDIUM}(a, b, \varepsilon, n, f)$

вывести заголовок выходных данных,  $\mathcal{I}$

конец { головного алгоритма }

{ \_\_\_\_\_ }

алг вещь MEDIUM (вещ  $a, b, \varepsilon$ , цел  $m$ , алг вещь  $f$  (вещ  $x$ ))

знач  $a, b, \varepsilon, m$

вещ  $h, I_n, I_{2n}, \varepsilon_3, h_{half}, t, S$

лог  $finished$

цел  $n, i$

начало

$n \leftarrow m \quad h \leftarrow (b - a)/n \quad I_n \leftarrow 0$

$\varepsilon_3 \leftarrow 3 \cdot \varepsilon \quad finished \leftarrow ложь$

пока не  $finished$  выполнять

$h_{half} \leftarrow 0.5 \cdot h \quad t \leftarrow a + h_{half} \quad S \leftarrow f(t)$

для  $i$  от 1 до  $n - 1$  шаг 1 выполнять

$S \leftarrow S + f(t + i \cdot h)$

конец-цикл { суммирования ординат }

$I_{2n} \leftarrow S \cdot h \quad finished \leftarrow (|I_{2n} - I_n| \leq \varepsilon_3 \cdot |I_{2n}|)$

$I_n \leftarrow I_{2n} \quad h \leftarrow h_{half} \quad n \leftarrow n + n$

конец-цикл { по  $\varepsilon$  }

MEDIUM  $\leftarrow I_{2n}$

возвратиться

конец { алг Medium }

{ \_\_\_\_\_ }

алг вещь  $f$  (вещ  $u$ )

знач  $u$

начало

$f \leftarrow \exp(-u^2)$

возвратиться

конец { алг  $f(u)$  }

{ \_\_\_\_\_ }

Результат выполнения Фортран 90-программы следующий:  $\text{erf}(1) = 0.8427008$ .  $\triangle$

В дополнение к примеру 11 приведём текст алг вещь PARABOLA. Его эффективность усилена (почти вдвое) за счёт уменьшения числа обращений к подынтегральной функции (для равноотстоящих точек деления или, другими словами, узлов, такой экономии можно добиться использованием интегральных сумм или ординат, накопленных при предыдущих разбиениях промежутка интегрирования).

{  
 алг вещь PARABOLA (вещ  $a, b, \varepsilon$ , цел  $m$ ,  
 алг вещь  $f$  (вещ  $x$ ))

знач  $a, b, \varepsilon, m$   
 вещь  $h, I_n, I_{2n}, \varepsilon_{15}, h_2, t, S_0, S_2, S_4$   
 лог  $finished$   
 цел  $n, i$   
 начало  
 $n \leftarrow m$        $h \leftarrow (b - a)/(2 \cdot n)$        $I_n \leftarrow 0$   
 $\varepsilon_{15} \leftarrow 15 \cdot \varepsilon$        $finished \leftarrow ложь$        $h_2 \leftarrow 2 \cdot h$   
 $S_0 \leftarrow f(a) + f(b)$        $S_2 \leftarrow f(a + h_2)$   
 для  $i$  от 2 до  $n - 1$  шаг 1 выполнять  
 $S_2 \leftarrow S_2 + f(a + i \cdot h_2)$   
 конец-цикл { по  $i; S_2$  }  
 пока не  $finished$  выполнять  
 $t \leftarrow a + h$        $S_4 \leftarrow f(t)$   
 для  $i$  от 1 до  $n - 1$  шаг 1 выполнять  
 $S_4 \leftarrow S_4 + f(t + i \cdot h_2)$   
 конец-цикл { по  $i; S_4$  }  
 $I_{2n} \leftarrow h \cdot (S_0 + 2 \cdot (S_2 + 2 \cdot S_4))$   
 $finished \leftarrow (|I_{2n} - I_n| \leq \varepsilon_{15} \cdot |I_{2n}|)$   
 $I_n \leftarrow I_{2n}^{1/2}$        $h_2 \leftarrow h$        $n \leftarrow n + n$   
 $h \leftarrow 0.5 \cdot h$        $S_2 \leftarrow S_2 + S_4$   
 конец-цикл { по  $\varepsilon$  }  
 PARABOLA  $\leftarrow I_{2n}/3$

возвратиться

конец

{ алг Parabola }

{ ————— }

### 2.2.5 Формулы Гаусса — Лежандра

Рассмотренные выше формулы численного интегрирования относятся к группе *формул Ньютона — Котеса*. Они основаны на полиномиальной аппроксимации подынтегральной функции и отличаются друг от друга степенью аппроксимирующего полинома. Степень полинома определяет количество узлов, в которых надо знать значения функции  $f(x)$ . При этом расстояния между узлами постоянны и равны величине шага интегрирования. Повышая степень аппроксимирующего полинома, с одной стороны, и увеличивая число разбиений промежутка интегрирования, с другой, можно добиться улучшения точности результата интегрирования. Но при этом резко возрастает объём вычислительной работы.

Желая достигнуть высокой точности вычислений с возможно малым числом  $n$ , где  $n$  — число узлов интегрирования, Гаусс получил формулы, дающие точный результат в случае, когда подынтегральная функция является произвольным полиномом степени не выше  $2n - 1$ . В них используются неравноотстоящие узлы — корни полиномов Лежандра  $n$ -й степени. Для  $n = 2, 3, \dots, 16$  узлы и соответствующие им коэффициенты затабулированы (см., например, [3], [7] и [13]).

Для интеграла (12) *формула Гаусса — Лежандра* имеет вид

$$I = \int_a^b f(x) dx \approx \frac{b-a}{2} \sum_{i=1}^n C_i^{(n)} f\left(\frac{b+a}{2} + \frac{b-a}{2} t_i^{(n)}\right), \quad (59)$$

$$R_n = \frac{(b-a)^{2n+1} (n!)^4}{(2n+1)((2n)!)^3} f^{(2n)}(\zeta), \quad \zeta \in [a; b].$$

При  $n = 4$  узлы  $t_i^{(4)}$  ( $i = 1, 2, 3, 4$ ) и коэффициенты (веса́)

$C_i^{(4)}$  формулы таковы:

$$\begin{aligned} t_3^{(4)} &= -t_2^{(4)} = 0.33998\ 10435\ 84856\ 265, \\ t_4^{(4)} &= -t_1^{(4)} = 0.86113\ 63115\ 94052\ 575, \\ C_3^{(4)} &= C_2^{(4)} = 0.65214\ 51548\ 62546\ 143, \\ C_4^{(4)} &= C_1^{(4)} = 0.34785\ 48451\ 37453\ 858. \end{aligned} \quad (60)$$

А это формулы для оценки погрешностей:

$$\begin{aligned} |I_{\text{последующий}} - I_{\text{предыдущий}}| &\leq 255\varepsilon |I_{\text{последующий}}|, \\ |I_{\text{последующий}} - I_{\text{предыдущий}}| &\leq 255\varepsilon. \end{aligned} \quad (61)$$

И, наконец, приведём текст алгоритма GAUSS\_4 (формулы (59)–(61)):

{ \_\_\_\_\_ }

алг длиндвточн GAUSS\_4 (длиндвточн  $a, b, \varepsilon,$

алг длиндвточн  $f$  (длиндвточн  $x$ ))

знач  $a, b, \varepsilon$

длиндвточн  $\varepsilon_{255}, h, I_n, I_{2n}, b_1, a_1, k_1, k_2, k_{2t_3}, k_{2t_4}$

цел  $n, i$  лог  $finished$

начало

$\varepsilon_{255} \leftarrow 255\varepsilon$        $h \leftarrow 0.25(b - a)$

$I_n \leftarrow 0$        $n \leftarrow 4$        $finished \leftarrow ложь$

пока не  $finished$  выполнять

$I_{2n} \leftarrow 0$

для  $i$  от 1 до  $n$  шаг 1 выполнять

$b_1 \leftarrow a + i \cdot h$        $a_1 \leftarrow b_1 - h$

$k_1 \leftarrow 0.5(b_1 + a_1)$        $k_2 \leftarrow 0.5(b_1 - a_1)$

$k_{2t_3} \leftarrow 0.33998\ 10435\ 84856\ 265k_2$

$k_{2t_4} \leftarrow 0.86113\ 63115\ 94052\ 575k_2$

$I_{2n} \leftarrow I_{2n} + k_2 \cdot$

$(0.34785\ 48451\ 37453\ 858(f(k_1 - k_{2t_4}) + f(k_1 + k_{2t_4}))) +$

$0.65214\ 51548\ 62546\ 143(f(k_1 - k_{2t_3}) + f(k_1 + k_{2t_3}))$

**конец-цикл**{ по  $i$  } $finished \leftarrow (|I_{2n} - I_n| \leq \varepsilon_{255} \cdot |I_{2n}|)$  $n \leftarrow n + n \quad h \leftarrow 0.5h \quad I_n \leftarrow I_{2n}$ **конец-цикл**{ по  $\varepsilon$  }**GAUSS\_4**  $\leftarrow I_{2n}$ **возвратиться****конец**

{ алг Gauss\_4 }

---

**Обсуждение.** С помощью любой формулы численного интегрирования при достаточно малом шаге  $h$  можно получить результат с произвольной наперёд заданной точностью  $\varepsilon$  [11]. Так как при численном интегрировании много ресурсов уходит на подсчёт значений подынтегральной функции  $f(x)$ , то сначала следует изучить её поведение на промежутке интегрирования. С целью сокращения количества операций можно попытаться преобразовать выражение для  $f(x)$ .

**ПРИМЕР 12.** Длину эллипса  $\ell$  можно найти по формуле  $\ell = \int_0^{2\pi} \sqrt{a^2 \sin^2 x + b^2 \cos^2 x} dx$ , где  $a$  и  $b$  — полуоси эллипса; уменьшив число операций вчетверо, для чего, учитывая симметричность геометрической фигуры относительно обеих осей координат, будем вычислять длину по формуле  $\ell = 4 \int_0^{\pi/2} \sqrt{a^2 \sin^2 x + b^2 \cos^2 x} dx$ ; затем введём в рассмотрение квадрат эксцентриситета  $e^2 = 1 - (b/a)^2$  и, используя тригонометрическую единицу в форме  $\sin^2 x = 1 - \cos^2 x$ , получим формулу  $\ell = 4a \int_0^{\pi/2} \sqrt{1 - (e \cos x)^2} dx$  и, таким образом, сэкономим примерно ещё одну треть операций.  $\triangle$

Затем следует найти точки, в которых  $f'(x)$  обращается в бесконечность. Наличие их сильно увеличивает время вычислений. Можно попытаться подыскать такую замену переменной интегрирования, которая устранила бы особенность в первой производной. Обращение в бесконечность производных более высоких порядков также ухудшает временной параметр, но не в такой степени, как в случае

первой производной.

Если функция  $f(x)$  имеет непрерывные производные достаточно высокого порядка, то при одинаковом количестве узлов формулы численного интегрирования *в порядке возрастания точности* можно расположить так:

- 1) формулы левых и правых прямоугольников,
- 2) формула трапеций,
- 3) формула средних прямоугольников (формула средних точек),
- 4) формула парабол,
- 5) формулы Гаусса — Лежандра.

Перечисленные соображения следует рассматривать как ориентировочные рекомендации [11].

**Упражнение 4.** 1. Напишите условие окончания разбиений отрезка интегрирования для шагов  $h$  и  $ph$  ( $p=3, 4, \dots$ ).

2\*. Выведите квадратурную формулу с двумя узлами, по которой точно интегрируются полиномы до третьей степени включительно.

3. При заданном числе разбиений  $n$  промежутка интегрирования ( $n = 2^k$ ,  $k = 1, 2, \dots, 7$ ) найдите длину  $\ell = 4a \int_0^{\pi/2} \sqrt{1 - (e \cos x)^2} dx$  эллипса  $\left\{ (x, y) \left| \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \right. \right\}$  ( $a, b > 0$ ) и соответствующую относительную ошибку. Для каждого  $n$  определите время вычисления интеграла. Как отtestировать результат интегрирования?

4. С заданной точностью  $\varepsilon$ , меняя начальное число разбиений  $n$  ( $n = 2^k$ ,  $k = 1, 2, \dots, 7$ ), вычислите длину эллипса. При каком  $n$  за наименьшее время был подсчитан интеграл?

5. Возьмите из задачи 4 “оптимальное”  $n$  и, меняя точность  $\varepsilon$  ( $\varepsilon = 10^{-q}$ ;  $q = 1, 2, \dots, 15$ ), снова найдите длину эллипса. Для каждого значения  $\varepsilon$  подсчитайте время выполнения выбранного алгоритма интегрирования.

6. При заданных точности  $\varepsilon$  и начальном числе разбиений  $n$  по выбранной квадратурной формуле подсчитайте

значение интеграла (используйте ординаты, вычисленные на предыдущих этапах разбиения промежутка интегрирования):

- |   |  |
|---|--|
| 1) $\int_0^1 \frac{dx}{1+x}$ ;              | 18) $\int_0^{\pi/2} \frac{\cos x}{1+x} dx$ ;   |
| 2) $\int_1^3 \frac{dx}{1+x^3}$ ;            | 19) $\int_0^1 \frac{dx}{\sqrt{1+x^2}}$ ;   |
| 3) $\int_0^1 \frac{dx}{1+x^4}$ ;            | 20) $\int_0^1 \frac{\operatorname{arctg} x}{x+1} dx$ ;   |
| 4) $\int_0^1 \frac{x}{x^2+1} dx$ ;          | 21) $\int_0^1 \cos \frac{\pi}{2} x^2 dx$ ;   |
| 5) $\int_4^{5.2} \ln x dx$ ;                | 22) $\int_0^1 \sqrt{1+x} dx$ ;   |
| 6) $\int_{0.1}^{1.6} \frac{dx}{\sqrt{x}}$ ; | 23) $\Phi(3) = \sqrt{\frac{2}{\pi}} \int_0^3 e^{-x^2/2} dx$ ;                                    |
| 7) $\int_0^{\pi/2} \cos x dx$ ;             | 24) $\Phi(3\sqrt{2}) = \operatorname{erf} 3 = \frac{2}{\sqrt{\pi}} \int_0^3 e^{-x^2} dx$ ;       |
| 8) $\int_1^{2.2} \frac{dx}{x}$ ;            | 25) $T = 4\sqrt{\frac{l}{g}} \int_0^{\pi/2} \frac{dx}{\sqrt{1-0.1 \sin^2 x}}$ ;                  |
| 9) $\int_1^2 x^p dx$ ;                      | 26) $\pi = (-6 \int_0^{\pi} x^2 \sin 6x dx)^{1/2}$ ;   |
| 10) $\int_1^2 e^x dx$ ;                     | 27) $\pi = \frac{e}{2} \int_0^{2\pi} \frac{e + \cos x}{1 + 2e \cos x + e^2} dx$ ;                |
| 11) $\int_3^4 2^x dx$ ;                     | 28) $L = \int_0^{\pi/2} \sqrt{b^2 + (a^2 - b^2) \sin^2 x} dx$ ;                                  |
| 12) $\int_{0.1}^1 \operatorname{tg} x dx$ ; | 29) $L = a \int_0^{2\pi} \sqrt{1 + \lambda^2 - 2\lambda \cos x} dx$ ;                            |
| 13) $\int_1^2 \operatorname{ctg} x dx$ ;    | 30) $J = ab \frac{\delta}{2} \int_0^{\pi/2} \sqrt{a^2 \sin^2 x + b^2 \cos^2 x} \sin 2x dx$ ;     |
| 14) $\int_{0.3}^1 \frac{dx}{\cos^2 x}$ ;    | 31) $\frac{1}{16} \int_0^1 (231x^6 - 315x^4 + 105x^2 - 5) dx$ ;                                  |
| 15) $\int_1^2 \frac{dx}{\sin^2 x}$ ;        | 32) $P_6(x) = \frac{1}{\pi} \int_0^{\pi} \frac{d\varphi}{(x + \sqrt{x^2 - 1} \cos \varphi)^7}$ ; |

$$16) \int_1^2 \operatorname{ch} x \, dx; \quad 33) J_p(x) = \frac{1}{\pi} \int_0^\pi \cos(p\varphi - x \sin \varphi) \, d\varphi.$$

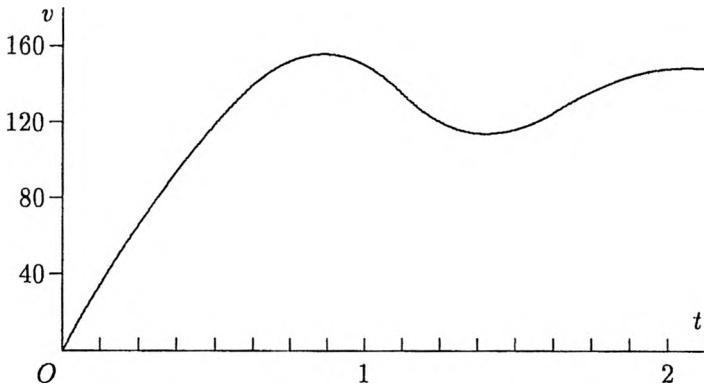
$$17) \int_1^2 \operatorname{sh} x \, dx;$$

7. Пользуясь следующей таблицей значений функ-

0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
1.000	1.063	1.266	1.647	2.280	3.290	4.881	7.378	11.30

ции  $f(x)$ , вычислите интеграл  $\int_a^b f(x) \, dx$  по формуле парабол.

8. Пользуясь графиком скорости  $v$  (км/ч) автомобиля,



определите пройденный им путь за время от 0 ч до 2 ч.

9. Вычислите интеграл  $I_n = \frac{1}{e} \int_0^1 x^n e^x \, dx$  ( $n \in \mathbb{N}$ ) по рекуррентным формулам (неустойчивый и устойчивый алгоритмы [4]):

$$I_n = 1 - nI_{n-1}, \quad I_0 = 1 - 1/e = 0.632\,120\,558\,828\,557\,678$$

$(n = 1, 2, \dots, 15);$

$$I_{n-1} = 1/n \cdot (1 - I_n), \quad I_{15} = 0.059\,017\,540\,879\,297\,775$$

$(n = 15, 14, \dots, 1).$

## ГЛАВА III. Решение систем линейных алгебраических уравнений

### 1. Элементарная алгебра матриц

В школьных курсах математики, физики и химии встречаются задачи, сводящиеся к решению систем линейных алгебраических уравнений (короче, СЛАУ). Эти системы состоят обычно из двух или трёх уравнений соответственно с двумя и тремя неизвестными. Практические задачи, решаемые в математике, физике, механике, геофизике, астрономии, метеорологии, экономике и т. д., нередко приводят к системам из сотен и даже тысяч уравнений с соответствующим числом неизвестных. Более того, Валях Е. утверждает [4], что 75% всех расчётных задач приходится на решение СЛАУ. Это не удивительно, так как математические модели тех или иных процессов либо сразу строятся как линейные алгебраические, либо сводятся к таковым посредством дискретизации и (или) линеаризации. Но прежде чем излагать компьютерные методы решения СЛАУ, приведём некоторые сведения из линейной алгебры.

#### 1.1 Матрицы и векторы

*Прямоугольной матрицей* размера  $m \times n$  называется совокупность элементов любой природы (например, чисел (действительных или комплексных), переменных, полиномов, функций, векторов, матриц и т. д.), расположенных в виде таблицы, состоящей из  $m$  строк и  $n$  столбцов. Матрицы обозначаются прописными буквами  $A, B, C, \dots$ , а их элементы — соответствующими строчными буквами с двойными индексами (нумерация строк производится сверху вниз, а столбцов — слева направо). Например, исходя

из определения некоторую матрицу  $A$  представим в виде

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Кроме того допускается запись этой матрицы в строку в виде  $A = (a_{ij})$  ( $i = 1, 2, \dots, m; j = 1, 2, \dots, n$ ), или  $A = (a_{ij})_{m \times n}$ , или ещё короче:  $A_{m \times n}$ . Здесь первый индекс  $i$  элемента  $a_{ij}$  есть номер строки, а второй  $j$  — номер столбца. На их пересечении находится этот элемент. Вместо круглых скобок в обозначении матрицы можно использовать квадратные скобки и даже двойные вертикальные черты.

Матрица, состоящая из одной строки (при этом  $m = 1$ ), называется *матрицей-строкой* (например,  $A = (a_{11} \ a_{12} \ \dots \ a_{1n})$ ), а матрица, имеющая только один столбец ( $n = 1$ ),

именуется *матрицей-столбцом* (например,  $A = \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix}$ ).

Если матрица размера  $1 \times n$  есть матрица-строка, матрица размера  $m \times 1$  — матрица-столбец, то число (скаляр) можно рассматривать как матрицу размера  $1 \times 1$ . В дальнейшем матрицу-строку или матрицу-столбец будем называть кратко *вектором* (например, некоторый вектор-столбец  $\mathbf{x}$

обозначим так:  $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{pmatrix}$ ; при этом элементы  $x_1, x_2, \dots, x_m$  есть *координаты* вектора  $\mathbf{x}$ ).

Если число строк матрицы  $A$  равно числу столбцов ( $m = n$ ), то матрица  $A = (a_{ij})_{n \times n} = (a_{ij})_n = A_{n \times n} = A_n$  называется *квадратной матрицей* порядка  $n$ :

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

Такая матрица имеет две диагонали, причём *главной* является диагональ, проходящая через верхний левый и нижний правый углы. Принадлежащие ей элементы  $a_{ii}$  ( $i = 1, 2, \dots, n$ ) называются *главными диагональными* или просто *диагональными элементами*. Вторая диагональ именуется *побочной*, а элементы  $a_{i, n-i+1}$  ( $i = 1, 2, \dots, n$ ) — *побочными диагональными элементами*.

Квадратная матрица, в которой элементы, не лежащие на главной диагонали, суть нули, называется *диагональной матрицей* ( $a_{ij} = 0$  для всех  $i \neq j$ ;  $i, j = 1, 2, \dots, n$ ).

Среди диагональных матриц особую роль играет матрица  $E$  вида

$$E_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

Речь идёт об *единичной матрице* порядка  $n$ . В матричном исчислении она играет роль единицы.

Матрица, состоящая из нулей, именуется *нулевой*. Она обозначается буквой  $O = O_{m \times n}$ .

Квадратная матрица  $A$ , ненулевые элементы которой составляют “ленту”, параллельную диагонали, называется *ленточной*. Частный её случай — *трёхдиагональная матрица* вида

$$A = \begin{pmatrix} a_{11} & a_{12} & 0 & \dots & 0 \\ a_{21} & a_{22} & a_{23} & \dots & 0 \\ 0 & a_{32} & a_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{pmatrix}. \quad (1)$$

В ней отличны от нуля только элементы диагонали и двух соседних линий ( $a_{ij} = 0$ , если  $|i - j| > 1$ ;  $i, j = 1, 2, \dots, n$ ).

Квадратную матрицу

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

называют *верхней треугольной матрицей* порядка  $n$ . В ней элементы  $a_{ij} = 0$  для всех  $i > j$  ( $i, j = 1, 2, \dots, n$ ).

Квадратная матрица  $A$  называется *симметрической*, если для всех  $i \neq j$  выполняется равенство  $a_{ij} = a_{ji}$  ( $i, j = 1, 2, \dots, n$ ), и *кососимметрической*, если  $a_{ij} = -a_{ji}$ .

И, наконец, переставляя в матрице  $A = (a_{ij})_{m \times n}$  строки со столбцами, получим так называемую *транспонированную матрицу*  $A^T = (a_{ij}^T)_{n \times m}$ . При этом  $a_{ij}^T = a_{ji}$  ( $i = 1, 2, \dots, n$ ;  $j = 1, 2, \dots, m$ ).

## 1.2 Основные действия над матрицами

Рассмотрим теперь, как определяются действия над матрицами.

*Равенство матриц.* Две матрицы  $A = (a_{ij})$  и  $B = (b_{ij})$  считаются *равными*,  $A = B$ , если они имеют одинаковые размеры и равные соответствующие элементы, т. е.  $a_{ij} = b_{ij}$  ( $i = 1, 2, \dots, m$ ;  $j = 1, 2, \dots, n$ ).

*Сложение матриц.* Суммой (разностью) двух матриц одного и того же размера  $A \pm B = (a_{ij}) \pm (b_{ij})$  ( $i = 1, 2, \dots, m$ ;  $j = 1, 2, \dots, n$ ) называется матрица  $C = (c_{ij})$  того же размера, элементы которой  $c_{ij}$  равны суммам (разностям) соответствующих элементов  $a_{ij}$  и  $b_{ij}$  матриц  $A$  и  $B$ , т. е.  $c_{ij} = a_{ij} \pm b_{ij}$  ( $i = 1, 2, \dots, m$ ;  $j = 1, 2, \dots, n$ ).

Операция сложения матриц обладает следующими свойствами:

- 1)  $A + (B + C) = (A + B) + C$ ;    3)  $A + O = A$ ;
- 2)  $A + B = B + A$ ;    4)  $(A + B)^T = A^T + B^T$ .

*Умножение матрицы на число* (или умножение числа на матрицу). Произведением матрицы  $A = (a_{ij})$  на число  $\alpha$  называется матрица  $A\alpha = \alpha A = (\alpha a_{ij})$ .

Это произведение подчиняется таким законам:

- 1)  $1A = A$ ;
- 2)  $0A = O$ ;
- 3)  $\alpha(\beta A) = (\alpha\beta)A$ ;
- 4)  $(\alpha + \beta)A = \alpha A + \beta A$ ;
- 5)  $\alpha(A + B) = \alpha A + \alpha B$ .

*Умножение (сцепленных) матриц.* Матрицы  $A = (a_{ik})$  размера  $m \times n$  и  $B = (b_{kj})$  размера  $n \times l$  называются *сцепленными*, если число столбцов первой матрицы равно числу строк второй. Такие матрицы можно перемножать. *Произведением*  $C = AB$  является матрица размера  $m \times l$ , элементы которой  $(c_{ij})$  вычисляются по правилу

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^n a_{ik}b_{kj},$$

$$(i = 1, 2, \dots, m; j = 1, 2, \dots, l),$$

т. е. элемент, стоящий в  $i$ -й строке и  $j$ -м столбце матрицы произведения, получается в виде скалярного произведения  $i$ -й вектор-строки матрицы  $A$  на  $j$ -й вектор-столбец матрицы  $B$ . В частности, можно перемножать квадратные матрицы лишь одинакового порядка.

**ПРИМЕР 1.** Найдём следующее произведение:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 \cdot 1 + 2 \cdot 1 + 3 \cdot 1 & 1 \cdot 2 + 2 \cdot 2 + 3 \cdot 2 & 1 \cdot 3 + 2 \cdot 3 + 3 \cdot 3 \\ 4 \cdot 1 + 5 \cdot 1 + 6 \cdot 1 & 4 \cdot 2 + 5 \cdot 2 + 6 \cdot 2 & 4 \cdot 3 + 5 \cdot 3 + 6 \cdot 3 \\ 7 \cdot 1 + 8 \cdot 1 + 9 \cdot 1 & 7 \cdot 2 + 8 \cdot 2 + 9 \cdot 2 & 7 \cdot 3 + 8 \cdot 3 + 9 \cdot 3 \end{pmatrix} =$$

$$= \begin{pmatrix} 6 & 12 & 18 \\ 15 & 30 & 45 \\ 24 & 48 & 72 \end{pmatrix}.$$

Теперь умножим матрицу на вектор-столбец:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 1 + 3 \cdot 1 \\ 4 \cdot 1 + 5 \cdot 1 + 6 \cdot 1 \\ 7 \cdot 1 + 8 \cdot 1 + 9 \cdot 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 15 \\ 24 \end{pmatrix}.$$

Здесь  $AB = (a_{ik})_{3 \times 3} \cdot (b_{kj})_{3 \times 1} = C = (c_{ij})_{3 \times 1}$ .

И, наконец:

$$(1 \ 2 \ 3) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = (1 \cdot 1 + 2 \cdot 1 + 3 \cdot 1) = 6.$$

А здесь  $AB = (a_{ik})_{1 \times 3} \cdot (b_{kj})_{3 \times 1} = C = (c_{ij})_{1 \times 1}$ .  $\Delta$

Матричное произведение обладает следующими свойствами ( $A$ ,  $B$  и  $C$  — матрицы;  $\alpha$  — число):

- 1)  $A(BC) = (AB)C$ ;      4)  $(A + B)C = AC + BC$ ;
- 2)  $\alpha(AB) = (\alpha A)B = A(\alpha B)$ ;    5)  $A(B + C) = AB + AC$ ;
- 3)  $\alpha(A + B) = \alpha A + \alpha B$ ;      6)  $(AB)^T = B^T A^T$ .

При этом, если  $\alpha = -1$ , то матрица  $-A = (-1)A$  называется *противоположной* матрице  $A$ . Равенства 1)–6) понимаются в том смысле, что если одна из их частей существует, то другая часть также существует и они равны между собой.

Произведение двух произвольных матриц не всегда определено.

ПРИМЕР 2. Если  $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ ,  $B = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$ , то  $AB = \begin{pmatrix} 6 & 12 & 18 \\ 15 & 30 & 45 \end{pmatrix}$ , а  $BA$  не существует.  $\Delta$

Матрицы  $AB$  и  $BA$  имеют смысл одновременно в случае равенства числа строк  $A$  и столбцов  $B$ , а также — числа строк  $B$  и столбцов  $A$ . При выполнении этих условий матрицы  $AB$  и  $BA$  будут квадратными. Но даже для квадратных матриц  $A$  и  $B$  матрица  $AB$  не всегда равна матрице  $BA$ .

ПРИМЕР 3. Если  $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ ,  $B = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ , то  $AB = \begin{pmatrix} 3 & 6 \\ 7 & 14 \end{pmatrix}$ , а  $BA = \begin{pmatrix} 7 & 10 \\ 7 & 10 \end{pmatrix}$ .  $\Delta$

Поэтому принято различать произведение  $AB$  матрицы  $A$  на матрицу  $B$  *справа* и произведение  $BA$  матрицы  $A$  на  $B$  *слева*. Только в некоторых частных случаях операция умножения матриц перестановочна (коммутативна). Тогда,

если  $AB = BA$ , матрицы  $A$  и  $B$  называются *перестановочными* (к примеру, таковыми являются две диагональные матрицы; в частности, единичная матрица  $E$  перестановочна с любой квадратной матрицей  $A$  того же порядка:  $AE = EA = A$ ).

Действия над матрицами-строками и матрицами-столбцами (т. е. векторами) аналогичны соответствующим действиям над квадратными матрицами (одинаковых порядков).

**Упражнение 1.** 1. Приведите лаконичное определение (в виде варианта) единичной матрицы  $E_n = (e_{ij})$  ( $i, j = 1, 2, \dots, n$ ).

2. Сформулируйте определение нижней треугольной матрицы.

3. Для одномерной матрицы укажите транспонированную матрицу.

4. Исходя из определения транспонированной матрицы и понятия равенства матриц, дайте определение симметрической матрицы.

5. Что можно сказать о виде матрицы  $B$ , если:

1)  $B = (A^T)^T$ ;    2)  $B = AA^T$ ;    3)  $B = A^T A$ .

6. Сформулируйте правило перемножения одномерных матриц.

7. Верно ли, что для любой матрицы  $A$  размера  $m \times n$  справедливо равенство  $AE_n = E_m A = A$ ?

8. Найдите произведение матриц  $\begin{pmatrix} 5 & 2 & -2 & 3 \\ 9 & 2 & -3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 2 & 2 & 2 \\ -1 & 3 & -5 \\ 16 & 8 & 24 \\ 8 & 0 & 16 \end{pmatrix}$ .

Можно ли утверждать, что  $B=0$ , если  $AB=0$  и  $A \neq 0$ ?

9. Можно ли считать, что  $B=C$ , если  $AB=AC$  и  $A \neq 0$ ?

10. Дайте определение

1) произведения нескольких матриц;

2)  $n$ -й степени матрицы.

11. Приведите примеры перестановочных матриц.

### 1.3 Определители

Любой квадратной матрице  $A = (a_{ij})$  порядка  $n$  с действительными или комплексными элементами можно однозначно сопоставить действительное или комплексное число  $D$ , которое называется *определителем* (*детерминантом*) матрицы  $A$ , так что

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

Это число обозначается также  $\det A$ ,  $\Delta$  или  $|a_{ij}|$ . Оно подсчитывается по определённым правилам, к рассмотрению которых мы и переходим.

Порядок определителя соответствует порядку матрицы  $A$ , определителем которой он является. Просто вычисляются определители до третьего порядка включительно. Так, если порядок матрицы равен единице (матрица  $A$  состоит из одного элемента  $a_{11}$ ), то *определителем первого порядка*, соответствующим такой матрице, является число, равное этому элементу:  $\det A = a_{11}$ .

Возьмём теперь матрицу второго порядка  $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ . *Определителем второго порядка*, соответствующим матрице  $A$ , называется следующее число:

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}, \quad (3)$$

т. е. *определитель второго порядка* (3) равен произведению главных диагональных элементов минус произведение побочных диагональных элементов.

**ПРИМЕР 4.** Вычислите определитель матрицы  $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ .

*Решение.*  $\det A = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 1 \cdot 4 - 2 \cdot 3 = -2. \quad \Delta$

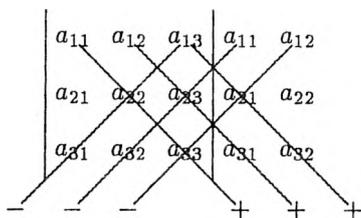
Определителем третьего порядка называется такое число

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} -$$

$$- a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \quad (4)$$

Для его подсчёта укажем следующее мнемоническое правило, носящее имя *Саррюса*<sup>1</sup>: чтобы получить (4), к обозначению определителя припишите справа два первых столбца, не меняя их порядка, и составьте сумму произведения элементов главной диагонали и произведений элементов, параллельных ей, из которой затем вычтите сумму произведения элементов побочной диагонали и произведений элементов, параллельных побочной диагонали, т. е.



ПРИМЕР 5. Вычислите определитель матрицы  $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ .

Решение.  $\begin{vmatrix} 1 & 2 & 3 & 1 & 2 \\ 4 & 5 & 6 & 4 & 5 \\ 7 & 8 & 9 & 7 & 8 \end{vmatrix} =$   
 $= 1 \cdot 5 \cdot 9 + 2 \cdot 6 \cdot 7 + 3 \cdot 4 \cdot 8 - 3 \cdot 5 \cdot 7 - 1 \cdot 6 \cdot 8 - 2 \cdot 4 \cdot 9 = 0.$   
 Отметим, что данная матрица  $A$  является особенной<sup>2</sup>.  $\triangle$

Теперь рассмотрим определитель  $n$ -го порядка. Для вычисления такого определителя введём понятия минора<sup>3</sup> и

<sup>1</sup> Другое мнемоническое правило сформулировано в Приложении.

<sup>2</sup> Если определитель матрицы  $A$  равен нулю, то она называется *особенной* (*вырожденной*, *сингулярной*). В противном случае ( $|A| \neq 0$ ) матрица  $A$  — *неособенная* (*невыврожденная*).

<sup>3</sup> Лат. *minor* — меньший.

алгебраического дополнения. Если  $D = |a_{ij}|$  — определитель порядка  $n$ , то *минором*  $M_{ij}$  элемента  $a_{ij}$  называется определитель порядка  $n - 1$ , получающийся из  $D$  вычёркиванием  $i$ -й строки и  $j$ -го столбца. Под *алгебраическим дополнением*  $A_{ij}$  элемента  $a_{ij}$  понимают минор  $M_{ij}$ , домноженный на  $(-1)^{i+j}$ . Таким образом, если сумма  $i + j$  номеров строки и столбца, на пересечении которых стоит элемент, чётная, то алгебраическое дополнение имеет тот же знак, что и минор; если же сумма  $i + j$  нечётная, то знак минора изменится на противоположный.

**ПРИМЕР 6.** В определителе третьего порядка  $D = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$  минорами элементов  $a_{31}$ ,  $a_{32}$  и  $a_{33}$  являются определители второго порядка соответственно  $M_{31} = \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = -3$ ,  $M_{32} = \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} = -6$  и  $M_{33} = \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = -3$ , а алгебраическими дополнениями —  $A_{31} = (-1)^{3+1}M_{31} = M_{31} = -3$ ,  $A_{32} = (-1)^{3+2}M_{32} = -M_{32} = 6$  и  $A_{33} = (-1)^{3+3}M_{33} = M_{33} = -3$ .  
 $\triangle$

Перечислим теперь *свойства определителей*. Они будут полезны при вычислении определителей больших порядков.

1) *Определитель не меняется при транспонировании:*  $\det A = \det A^T$ . Это свойство означает, что строки и столбцы определителя равноправны, поэтому все свойства, которые будут сформулированы ниже для строк, верны и для столбцов.

2) *При перестановке двух строк (столбцов) модуль определителя остаётся прежним, а знак меняется на обратный.* Итак, произвольная перестановка строк (столбцов) определителя может поменять только его знак.

3) *Если некоторую строку (столбец) определителя умножить на число  $\lambda \neq 0$ , то сам определитель умножится на это число.* Следовательно, общий множитель всех элементов некоторой строки (столбца) определителя можно вынести за его знак.

4) Определитель не изменится, если одну из строк (столбцов) сложить с другой строкой (столбцом), умноженной на число  $\lambda \neq 0$ .

ПРИМЕР 7. Умножим на 2 первую строку определителя  $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = -2$  и прибавим результат к второй строке, получим  $\begin{vmatrix} 1 & 2 \\ 5 & 8 \end{vmatrix} = -2$ .  $\triangle$

5) Если каждый элемент какой-либо строки (столбца) есть сумма двух слагаемых, то определитель равен сумме двух определителей: в одном вместо каждой суммы стоит только первое слагаемое, в другом — только второе; остальные элементы в обоих определителях те же, что в исходном, т. е. в случае строки имеем

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} = b_{21} + c_{21} & a_{22} = b_{22} + c_{22} & \dots & a_{2n} = b_{2n} + c_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} =$$

$$= \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Другими словами, при сложении двух определителей, различающихся только одной строкой (столбцом), соответствующие элементы этой строки (столбца) складываются.

ПРИМЕР 8.  $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 3 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 3 \\ 3 & 3 \end{vmatrix} = 1 + 3 + 3 - 9 = -2$ .  $\triangle$

6) Определитель равен нулю, во-первых, если одна из строк (один из столбцов) состоит из нулей или, во-вторых, если он содержит две пропорциональные строки (два пропорциональных столбца), в частности, если он имеет две одинаковые строки (два одинаковых столбца).

ПРИМЕР 9. Подсчитаем значение  $D = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$ .

*Решение.* Имеем  $D = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 7 & 8 & 9 \end{vmatrix} + \begin{vmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \\ 7 & 8 & 9 \end{vmatrix} = 1 \cdot 2 \cdot 9 + 2 \cdot 3 \cdot 7 + 3 \cdot 1 \cdot 8 - 3 \cdot 2 \cdot 7 - 1 \cdot 3 \cdot 8 - 2 \cdot 1 \cdot 9 + 1 \cdot 3 \cdot 9 + 2 \cdot 3 \cdot 7 + 3 \cdot 3 \cdot 8 - 3 \cdot 3 \cdot 7 - 1 \cdot 3 \cdot 8 - 2 \cdot 3 \cdot 9 = 0 + 0 = 0. \quad \triangle$

Теперь сформулируем теорему, которая облегчит вычисление определителей.

**ТЕОРЕМА 1. (Теорема разложения.)** *Определитель  $D = |a_{ij}|$   $n$ -го порядка равен сумме произведений элементов какой-либо строки (столбца) на их алгебраические дополнения:*

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

$$= a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in} \quad (i = 1, 2, \dots, n) \quad (5)$$

или

$$\det A = a_{1j}A_{1j} + a_{2j}A_{2j} + \dots + a_{nj}A_{nj} \quad (j = 1, 2, \dots, n). \quad (6)$$

Формула (5) называется *разложением определителя (2) по элементам  $i$ -й строки*, а формула (6) — *разложением определителя по элементам  $j$ -го столбца*.

При разложении определителя второго порядка по элементам любой строки (столбца) получится приведённая выше формула (3), а для третьего порядка — формула (4).

Определитель высокого порядка можно вычислить, исходя из определения, однако это крайне нерационально. Подсчёты существенно упростятся, если придерживаться правила “постепенного понижения порядка”. В основе его лежит теорема 1: определитель  $n$ -го порядка представляется в виде суммы определителей  $(n - 1)$ -го порядка (“раскладывается по элементам  $i$ -й строки (или  $j$ -го столбца”); желательно, чтобы в этой строке (столбце) было как

можно больше нулей (используя свойства определителей, этого можно просто добиться)); затем к каждому определителю  $(n - 1)$ -го порядка вновь применяют теорему разложения и действуют подобным образом до тех пор, пока не получат определители третьего (или даже второго) порядка.

**ПРИМЕР 10.** Вычислим определитель  $D = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$ .

*Решение.* Разложим его по элементам третьей строки. Создадим на месте элемента  $a_{33} = 9$  нуль. Для этого к третьей строке добавим элементы первой строки, предварительно домноженные на  $-3$ . Получим  $D = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 4 & 2 & 0 \end{vmatrix}$ . Применив к последнему определителю теорему 1, найдём, что  $D = 4 \cdot (-1)^{3+1} \cdot \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} + 2 \cdot (-1)^{3+2} \cdot \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} + 0 = 4 \cdot (12 - 15) - 2 \cdot (6 - 12) = -12 + 12 = 0$ .  $\triangle$

В общем случае вычисление определителя (без применения специальных приёмов) является крайне трудоёмкой операцией. Действительно, определитель  $n$ -го порядка представляет собой алгебраическую сумму  $n!$  слагаемых, каждое из которых есть произведение  $n$  элементов; значит, для подсчёта определителя потребуется  $(n - 1)n!$  умножений и  $n! - 1$  сложений, а всего

$$N = n \cdot n! - 1 \quad (7)$$

операций.

**Упражнение 2.** 1. Вычислите следующие определители:

$$\begin{aligned} & 1) \begin{vmatrix} 1 & 2 & 3 \\ 3 & 6 & 9 \\ 7 & 8 & 9 \end{vmatrix}; \quad 2) \begin{vmatrix} 2 & 4 & 6 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}; \quad 3) \begin{vmatrix} 2 & 5 & -2 \\ 3 & 8 & 0 \\ 1 & 3 & 5 \end{vmatrix}; \quad 4) \begin{vmatrix} 4 & -3 & 2 \\ 6 & 11 & 1 \\ 0 & 3 & 0 \end{vmatrix}; \quad 5) \begin{vmatrix} 2 & 5 & 3 \\ 0 & 6 & 2 \\ 7 & 3 & -1 \end{vmatrix}; \\ & 6) \begin{vmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{vmatrix}; \quad 7) \begin{vmatrix} 16 & 3 & 2 & 13 \\ 5 & 10 & 11 & 8 \\ 9 & 6 & 7 & 12 \\ 4 & 15 & 14 & 1 \end{vmatrix}; \quad 8) \begin{vmatrix} 6 & 3 & 0 & 3 \\ 4 & 4 & 2 & 1 \\ 0 & 4 & 4 & 2 \\ 7 & 7 & 8 & 5 \end{vmatrix}; \quad 9) \begin{vmatrix} -1 & -2 & 4 & 1 \\ 2 & 3 & 0 & 6 \\ 2 & -2 & 1 & 4 \\ 3 & 1 & -2 & -1 \end{vmatrix}. \end{aligned}$$

2. Найдите площадь треугольника с вершинами  $P_1(x_1, y_1)$ ,  $P_2(x_2, y_2)$  и  $P_3(x_3, y_3)$ :  $S = \frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$  (площадь  $S$  неотрицательна, если обход вершин в порядке нумерации проис-

ходит против часовой стрелки). Проверьте (с точностью  $\varepsilon = 10^{-3}$ ), лежат ли точки  $P_1(x_1, y_1)$ ,  $P_2(x_2, y_2)$  и  $P_3(x_3, y_3)$  на одной прямой?

3. Найдите координаты  $(x_0, y_0)$  точки пересечения прямых, заданных уравнениями  $a_1x + b_1y + c_1 = 0$ ,  $a_2x + b_2y + c_2 = 0$ :  $x_0 = \left| \begin{matrix} b_1 & c_1 \\ b_2 & c_2 \end{matrix} \right| / \left| \begin{matrix} a_1 & b_1 \\ a_2 & b_2 \end{matrix} \right|$ ,  $y_0 = \left| \begin{matrix} c_1 & a_1 \\ c_2 & a_2 \end{matrix} \right| / \left| \begin{matrix} a_1 & b_1 \\ a_2 & b_2 \end{matrix} \right|$ . Проходит ли прямая  $a_3x + b_3y + c_3 = 0$  через ту же точку, т. е.  $\left| \begin{matrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{matrix} \right| = 0$ ? ( $\varepsilon = 10^{-3}$ ).

4. Проверьте (с точностью  $\varepsilon = 10^{-3}$ ), лежит ли точка  $P(x, y)$  на окружности, проходящей через точки  $P_1(x_1, y_1)$ ,  $P_2(x_2, y_2)$  и  $P_3(x_3, y_3)$ :  $\left| \begin{matrix} x^2+y^2 & x & y & 1 \\ x_1^2+y_1^2 & x_1 & y_1 & 1 \\ x_2^2+y_2^2 & x_2 & y_2 & 1 \\ x_3^2+y_3^2 & x_3 & y_3 & 1 \end{matrix} \right| = 0$ ?

5. Напишите уравнение плоскости, проходящей через точки  $P_1(x_1, y_1, z_1)$ ,  $P_2(x_2, y_2, z_2)$  и  $P_3(x_3, y_3, z_3)$ :

$$\left| \begin{matrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{matrix} \right| = 0.$$

6. Вычислите объём треугольной пирамиды с вершинами  $P(x, y, z)$ ,  $P_1(x_1, y_1, z_1)$ ,  $P_2(x_2, y_2, z_2)$  и  $P_3(x_3, y_3, z_3)$ :

$V = \frac{1}{6} \left| \begin{matrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{matrix} \right|$ . Когда четыре точки  $P$ ,  $P_1$ ,  $P_2$  и  $P_3$  принадлежат одной плоскости? ( $\varepsilon = 10^{-3}$ ).

7. Проходят ли четыре плоскости  $a_i x + b_i y + c_i z + d_i = 0$  ( $i = 1, 2, 3, 4$ ) через одну точку, т. е.  $\left| \begin{matrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{matrix} \right| = 0$ ? ( $\varepsilon = 10^{-3}$ ).

8. Оцените значение  $N$  (см. формулу (7)) в зависимости от порядка  $n$  определителя ( $n = 3, 10, 20, \dots, 100$ ). Подсчитайте время вычисления таких определителей на компьютере со средним быстродействием  $10^{10}$  операций/сек.

## 2. Прямые методы

### Постановка задачи

Перейдём к решению систем линейных алгебраических уравнений. Напомним, что *линейным уравнением* с неизвестными  $x_1, x_2, \dots, x_n$  называется уравнение вида

$$a_1x_1 + a_2x_2 + \dots + a_nx_n = b,$$

где числа  $a_1, a_2, \dots, a_n$  — *коэффициенты уравнения*, а  $b$  — его *правая часть*. В общей форме система  $n$  линейных алгебраических уравнений с  $n$  неизвестными записывается так:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2, \\ \dots & \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n, \end{aligned} \quad (8)$$

или, короче,  $\sum_{j=1}^n a_{ij}x_j = b_i \quad (i = 1, 2, \dots, n)$ .

Уравнения системы считаются пронумерованными: первое, второе,  $\dots$ ,  $n$ -е. Числа  $x_1, x_2, \dots, x_n$  называются *неизвестными* системы, числа  $a_{11}, a_{12}, \dots, a_{nn}$  — *коэффициентами* при неизвестных (коэффициент  $a_{ij}$  находится в  $i$ -м уравнении при  $j$ -м неизвестном), а числа  $b_1, b_2, \dots, b_n$  — *свободными членами*. Если все  $b_i = 0 \quad (i = 1, 2, \dots, n)$ , то имеем *однородную* систему; в противном случае говорят о *неоднородной* системе.

Если ввести в рассмотрение матрицу  $A$  коэффициентов системы, вектор-столбец из неизвестных (*искомый вектор  $\mathbf{x}$* ) и  $\mathbf{b}$  — вектор-столбец её свободных членов

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix},$$

то систему (8) можно представить в виде векторно-матричного уравнения  $Ax = b$ .

Система (8) называется *совместной*, если она имеет решение. *Решение* системы есть любая совокупность чисел  $x_1, x_2, \dots, x_n$ , которая, будучи подставленной в систему на место неизвестных, обращает все уравнения системы в тождества. При отсутствии специальных оговорок область допустимых значений неизвестных — множество действительных чисел. Если система не имеет решений, она *несовместна* (*противоречива*). Совместная система называется *определённой*, если она имеет единственное решение; в противном случае система *неопределённая*. Совокупность всех решений системы называется *множеством решений*. Две системы линейных уравнений *эквивалентны*, если они имеют одинаковые множества решений (например, пустые множества).

*Методы решения систем линейных уравнений* (их больше тридцати [14]) делятся на две группы — прямые и итеративные. *Прямые методы* для поиска неизвестных предполагают конечные соотношения (формулы). Они дают решение в результате выполнения заранее известного числа арифметических операций. Если при этом точно заданы коэффициенты и свободные члены системы (это рациональные числа, в частности, целые данные) и точно выполняются вычисления (например, по правилам действий над обыкновенными дробями), то получим точное решение системы (8). По этой причине прямые методы нередко называют *точными методами*. Но в практике компьютерных вычислений они не дают точных решений, ибо значения неизвестных содержат неизбежные погрешности. И это понятно: операции при поиске вектора  $x$  ведутся на конечной разрядной сетке, т. е. с округлениями. Прямыми методами, к примеру, являются, правило Крамера и алгоритм Гаусса.

*Итеративные методы* дают приближённое решение задачи (они позволяют найти неизвестные лишь с заданной точностью  $\varepsilon$  даже в предположении, что операции произво-

дятся без округлений). К итеративным методам относятся, например, методы простой итерации, Гаусса — Зейделя, Якоби, релаксации, Рундсона и т. д. [4, 7, 14].

## 2.1 Правило Крамера

Рассмотрим правило Крамера — прямой метод, имеющий большое теоретическое значение. По этому правилу каждое неизвестное записывается в виде отношения определителей<sup>4</sup>. Применим его к “школьным” системам с двумя и тремя неизвестными. Для большей наглядности укажем связь решений с геометрическими фактами [6].

### 2.1.1 Два уравнения с двумя неизвестными

Рассмотрим систему уравнений

$$a_1x + b_1y = c_1, \quad (9)$$

$$a_2x + b_2y = c_2 \quad (10)$$

(каждое из уравнений представляет прямую на плоскости  $xOy$ ).

Введём обозначения

$$D = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad D_x = \begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}, \quad D_y = \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}.$$

Здесь определитель  $D_x$  получается из определителя системы  $D$  заменой элементов первого столбца свободными членами системы; аналогично формируется  $D_y$ .

Возможны три случая.

**Случай 1.** Определитель системы не равен нулю:  $D \neq 0$ .

Система имеет единственное решение

$$x = \frac{D_x}{D}, \quad y = \frac{D_y}{D} \quad (11)$$

---

<sup>4</sup>Определители впервые были введены именно для решения систем линейных алгебраических уравнений.

(т. е. коэффициенты системы не пропорциональны:  $\frac{a_1}{a_2} \neq \frac{b_1}{b_2}$ ; прямые (9) и (10) пересекаются; формулы (11) дают координаты точки пересечения прямых).

**Случай 2.** Определитель системы равен нулю:  $D = 0$  (т. е. коэффициенты при неизвестных пропорциональны). Пусть при этом один из определителей  $D_x$  или  $D_y$  отличен от нуля (т. е. свободные члены не пропорциональны коэффициентам при неизвестных:  $\frac{a_1}{a_2} = \frac{b_1}{b_2} \neq \frac{c_1}{c_2}$ ).

Система не имеет решения (прямые (9) и (10) параллельны, но не совпадают).

Отметим, что равенство  $D = 0$  в вычислениях на компьютере не всегда получается точным. При условии  $D \approx 0$  прямые могут быть (почти) параллельными, а решение системы — весьма чувствительным к небольшому изменению коэффициентов системы. Если малые ошибки в исходных данных или погрешности вычислений приводят к существенным ошибкам в решении, то линейная система считается *плохо обусловленной*. Отметим, что примерное равенство  $D \approx 0$  есть необходимое, но не достаточное условие плохой обусловленности (например, система с диагональной матрицей седьмого порядка и коэффициентами  $a_{ii} = 0.1$  ( $i = 1, 2, \dots, 7$ ) не является плохо обусловленной, хотя её определитель ( $D = 10^{-7}$ ) мал).

**Случай 3.**  $D = 0$ ,  $D_x = 0$ ,  $D_y = 0$  (т. е. и коэффициенты, и свободные члены пропорциональны:  $\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}$ ).

Тогда одно из уравнений (9), (10) является следствием другого; система сводится к одному уравнению с двумя неизвестными и имеет бесконечное множество решений, со-

державшихся в формуле  $y = -\frac{a_2}{b_2}x + \frac{c_2}{b_2}$ , где  $x$  — какое-нибудь число (прямые (9) и (10) совпадают).

### 2.1.2 Два уравнения с тремя неизвестными

Рассмотрим теперь систему уравнений

$$a_1x + b_1y + c_1z = d_1, \quad (12)$$

$$a_2x + b_2y + c_2z = d_2 \quad (13)$$

(каждое из них представляет плоскость в пространстве).

Здесь возможны три случая.

**Случай 1.** Из трёх определителей

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}, \quad \begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix} \quad (14)$$

хотя бы один не равен нулю (т. е. коэффициенты при неизвестных не пропорциональны). Система имеет бесконечное множество решений, причём *одному* из неизвестных можно дать любое значение (например, если  $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0$ , то произвольное значение можно назначить неизвестному  $z$ , а  $x$  и  $y$  найти из системы

$$a_1x + b_1y = d_1 - c_1z,$$

$$a_2x + b_2y = d_2 - c_2z$$

(плоскости (12) и (13) не параллельны; система представляет прямую)).

**ПРИМЕР 11.** Решим систему  $x + 2y - 3z = 4$ ,  $2x + 4y + 6z = 7$ .

*Решение.* Имеем  $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0$ ,  $\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} = \begin{vmatrix} 2 & -3 \\ 4 & 6 \end{vmatrix} = 24$ ,  $\begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix} = \begin{vmatrix} -3 & 1 \\ 6 & 2 \end{vmatrix} = -12$ . Среди определителей есть не равные нулю. Тогда множество решений бесконечно. В силу неравенств  $\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \neq 0$  и  $\begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix} \neq 0$  произвольным может быть либо  $x$ , либо  $y$ . *Неизвестному  $z$  любого значения назначить нельзя.*

Выберем произвольное  $x$ . Найдём  $y$  и  $z$ . Имеем  $2y - 3z = 4 - x$ ,  $4y + 6z = 7 - 2x$ . Откуда  $y = \frac{\begin{vmatrix} 4-x & -3 \\ 7-2x & 6 \end{vmatrix}}{24} = 1\frac{7}{8} - \frac{1}{2}x$ ,  $z = \frac{\begin{vmatrix} 2 & 4-x \\ 4 & 7-2x \end{vmatrix}}{24} = -\frac{1}{12}$  (система представляет прямую, перпендикулярную оси  $Oz$ ).  $\triangle$

**Случай 2.** Все определители (14) равны нулю, но один из определителей

$$\begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix}, \quad \begin{vmatrix} b_1 & d_1 \\ b_2 & d_2 \end{vmatrix}, \quad \begin{vmatrix} c_1 & d_1 \\ c_2 & d_2 \end{vmatrix} \quad (15)$$

отличен от нуля (т. е. коэффициенты при неизвестных пропорциональны, а свободные члены таковыми не являются). Система не имеет решения (плоскости (12) и (13) параллельны, но не совпадают).

**ПРИМЕР 12.** Множество решений системы  $x + 2y - 3z = 4$ ,  $2x + 4y - 6z = 7$  пустое, так как все определители (14) равны нулю, а определитель  $\begin{vmatrix} a_1 & d_1 \\ a_2 & d_2 \end{vmatrix} = \begin{vmatrix} 1 & 4 \\ 2 & 7 \end{vmatrix}$  отличен от нуля.  $\triangle$

**Случай 3.** Все определители (14) и (15) равны нулю (т. е. и коэффициенты, и свободные члены пропорциональны). Система сводится к одному уравнению и имеет бесконечное множество решений, причём произвольные значения можно дать любым двум неизвестным (к примеру, если  $c_1 \neq 0$ , то ими могут быть  $x$  и  $y$  (плоскости (12) и (13) совпадают)).

**ПРИМЕР 13.** Множество решений системы  $x + 2y - 3z = 4$ ,  $2x + 4y - 6z = 8$  бесконечное (здесь и коэффициенты, и свободные члены пропорциональны). Система сводится к одному уравнению (плоскости совпали). Для любой пары неизвестных (скажем,  $y$  и  $z$ ) можно выбрать произвольные значения (тогда  $x = 4 - 2y + 3z$ ).  $\triangle$

### 2.1.3 Три уравнения с тремя неизвестными

Решим теперь следующую систему уравнений:

$$a_1x + b_1y + c_1z = d_1, \quad (16)$$

$$a_2x + b_2y + c_2z = d_2, \quad (17)$$

$$a_3x + b_3y + c_3z = d_3. \quad (18)$$

Введём обозначения

$$D = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, \quad (19)$$

$$D_x = \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}, \quad D_y = \begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix}, \quad D_z = \begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix}. \quad (20)$$

Здесь определитель  $D_x$  получается из определителя системы  $D$  заменой элементов первого столбца столбцом свободных членов системы; аналогично образуются  $D_y$  и  $D_z$ .

Если в определителе  $D$  соответственные элементы двух строк, скажем второй и третьей, пропорциональны, то уравнения (17) и (18) либо несовместны (система не имеет решений; см. п. 2.1.2, случай 2), либо сводятся к одному уравнению (п. 2.1.2, случай 3) и тогда будем иметь дело с системой двух уравнений (16) и (17), которая в свою очередь может свестись к одному уравнению. Так как все эти случаи исследованы в п. 2.1.2, то тогда положим, что в определителе  $D$  нет ни одной пары строк с пропорциональными элементами (среди плоскостей (16), (17), (18) нет ни одной пары параллельных).

В таком предположении возможны три случая.

**Случай 1.** Определитель системы не равен нулю:  $D \neq 0$ .

Система имеет единственное решение

$$x = \frac{D_x}{D}, \quad y = \frac{D_y}{D}, \quad z = \frac{D_z}{D} \quad (21)$$

(все плоскости пересекаются в одной точке).

**Случай 2.** Определитель системы равен нулю:  $D = 0$  и один из определителей (20) не равен нулю. Тогда и два других определителя отличны от нуля:<sup>5</sup>  $D_x \neq 0$ ,  $D_y \neq 0$ ,  $D_z \neq 0$ . Система не имеет решения (равенство  $D = 0$  означает, что плоскости (16), (17), (18) параллельны одной прямой и образуют призматическую поверхность).

**Случай 3.**  $D = 0$ ,  $D_x = 0$ ,  $D_y = 0$ ,  $D_z = 0$ . В этом случае любое уравнение (16), (17), (18) является следствием двух других. Данная система сводится к двум уравнениям с тремя неизвестными и множество её решений бесконечно (см. п. 2.1.2, случай 1; случаи 2 и 3 не могут иметь места ввиду сделанного выше предположения; плоскости (16), (17), (18) параллельны одной прямой, но теперь они, в отличие от случая 2, образуют пучок).

**ПРИМЕР 14.** Решим по формулам Крамера систему линейных уравнений

$$\begin{aligned}x_1 + 3x_2 + x_3 &= 0.6, \\2x_1 + x_2 - x_3 &= -0.7, \\x_1 - x_2 + 2x_3 &= 2.6.\end{aligned}\tag{22}$$

*Решение.* Матрица коэффициентов при неизвестных  $A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 1 & -1 \\ 1 & -1 & 2 \end{pmatrix}$  неособенная, так как её определитель  $D = \begin{vmatrix} 1 & 3 & 1 \\ 2 & 1 & -1 \\ 1 & -1 & 2 \end{vmatrix} = -17 \neq 0$ . Значит, система (22) — совместная и определённая. Множество её решений конечно (см.

---

<sup>5</sup>Если соответственные элементы двух строк определителя  $D$  пропорциональны (эту возможность мы исключили из рассмотрения), то может статься, что из трёх определителей  $D_x$ ,  $D_y$ ,  $D_z$  только один или только два равны нулю.

формулы (21) ):

$$x_1 = \frac{\begin{vmatrix} 0.6 & 3 & 1 \\ -0.7 & 1 & -1 \\ 2.6 & -1 & 2 \end{vmatrix}}{-17} = \frac{-4.9}{-17} = 0.288235294117647059,$$

$$x_2 = \frac{\begin{vmatrix} 1 & 0.6 & 1 \\ 2 & -0.7 & -1 \\ 1 & 2.6 & 2 \end{vmatrix}}{-17} = \frac{4.1}{-17} = -0.241176470588235294,$$

$$x_3 = \frac{\begin{vmatrix} 1 & 3 & 0.6 \\ 2 & 1 & -0.7 \\ 1 & -1 & 2.6 \end{vmatrix}}{-17} = \frac{-17.6}{-17} = 1.035294117647058820.$$

Здесь значения неизвестных приведены с 18-ю знаками после точки.  $\triangle$

Можно попытаться использовать правило Крамера для решения систем с матрицей коэффициентов  $n$ -го порядка. Но лучше этого не делать, так как для вычисления  $n$  неизвестных потребуется найти  $(n + 1)$  определитель. Предполагая, что определитель подсчитывается непосредственно (без применения эффективных алгоритмов), оценим число арифметических операций (см. формулу (7)):

$$N = n \cdot (n + 1)! - 1. \quad (23)$$

Как видим, при большом  $n$  число операций слишком велико (даже при  $n = 3$  потребуется 71 арифметическая операция; для сравнения, в схеме единственного деления решение системы (22) будет найдено за 46 действий (см. формулу (43))). Таким образом, порядок матрицы  $A$  является главным фактором, заставляющим отказаться от весьма привлекательных в теоретическом плане формул Крамера.

**Упражнение 3.** 1. Решите однородную систему

- 1) двух уравнений с двумя неизвестными;
- 2) двух уравнений с тремя неизвестными;

3) трёх уравнений с тремя неизвестными.

2. В зависимости от порядка  $n$  матрицы коэффициентов системы ( $n = 3, 10, 20, \dots, 100$ ) оцените значение  $N$  (см. формулу (23)). Подсчитайте время решения таких систем методом Крамера на компьютере, среднее быстродействие которого равно  $10^{10}$  операций/сек.

## 2.2 Метод Гаусса

Снова решим систему линейных уравнений. Применим теперь наиболее популярный прямой метод — метод Гаусса. Этот метод, основанный на последовательном исключении неизвестных, имеет много вариантов. Общее в них — переход от системы (8) к эквивалентной системе, которая просто решается; различия — в порядке исключения неизвестных, в способе хранения промежуточных результатов и в способе предупреждения нарастания погрешностей округления. Напомним, что *эквивалентными преобразованиями* любой линейной системы являются:

- 1) перестановка двух уравнений системы;
- 2) умножение (деление) обеих частей уравнения системы на произвольное число  $\lambda \neq 0$ ;
- 3) прибавление (вычитание) к обеим частям одного уравнения соответствующих частей другого уравнения, умноженных на любое число  $\lambda \neq 0$ .

Выберем простейший вариант метода Гаусса — *метод единственного деления*. Следуя Гауссу, сначала преобразуем квадратную матрицу коэффициентов при неизвестных к верхней треугольной матрице, из которой затем найдём решение системы. Проиллюстрируем сказанное на примере системы трёх уравнений с тремя неизвестными

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3. \end{aligned}$$

Более того, решим конкретную систему.

ПРИМЕР 15. Пусть это будет система (22):

$$x_1 + 3x_2 + x_3 = 0.6, \quad (24)$$

$$2x_1 + x_2 - x_3 = -0.7, \quad (25)$$

$$x_1 - x_2 + 2x_3 = 2.6. \quad (26)$$

*Решение.* Схема единственного деления состоит в следующем.

*Первый шаг.* Исключим неизвестное  $x_1$  из всех уравнений системы, кроме первого. Назовём  $x_1$  *ведущим неизвестным*, а коэффициент  $a_{11}$  — *ведущим элементом*. Разделив уравнение (24) на  $a_{11}$  (это возможно, так как  $a_{11} \neq 0$ )<sup>6</sup>, имеем

$$x_1 + 3 \cdot x_2 + 1 \cdot x_3 = 0.6. \quad (27)$$

Далее, вычитая уравнение (27), умноженное соответственно на  $a_{21} = 2$  и  $a_{31} = 1$ , из уравнений (25) и (26), получим систему двух уравнений с двумя неизвестными  $x_2$  и  $x_3$ :

$$-5 \cdot x_2 - 3 \cdot x_3 = -1.9, \quad (28)$$

$$-4 \cdot x_2 + 1 \cdot x_3 = 2. \quad (29)$$

Как видим, матрица коэффициентов этой системы имеет порядок на единицу меньший, чем матрица исходной системы.

*Второй шаг.* С последней системой проделаем аналогичные преобразования. Исключим теперь  $x_2$ . Тогда  $x_2$  — *ведущее неизвестное*, а коэффициент  $-5$  при нём — *ведущий элемент*. Поделив уравнение (28) на  $-5$ , получим

$$x_2 + 0.6 \cdot x_3 = 0.38. \quad (30)$$

Умножив (30) на  $-4$  и вычтя результат из уравнения (29), придём к одному уравнению

$$3.4 \cdot x_3 = 3.52. \quad (31)$$

---

<sup>6</sup>Здесь мы поступаем формально, несмотря на то, что  $a_{11} = 1$  и  $a_{31} = 1$ .

Таким образом, порядок матрицы коэффициентов исходной системы понизился ещё на единицу.

*Третий шаг.* Действуем по установленному выше образцу:  $x_3$  — ведущее неизвестное, а коэффициент 3.4 — ведущий элемент. Делим (31) на 3.4. Имеем

$$x_3 = \frac{3.52}{3.4} = 1.035294117647058820. \quad (32)$$

Объединяя уравнения (27), (30) и (32), придём к эквивалентной системе с треугольной матрицей коэффициентов

$$x_1 + 3 \cdot x_2 + 1 \cdot x_3 = 0.6, \quad (33)$$

$$x_2 + 0.6 \cdot x_3 = 0.38, \quad (34)$$

$$x_3 = 1.035294117647058820. \quad (35)$$

Решить её не составляет труда. Действительно, последнее уравнение (35) даёт  $x_3$ . Зная  $x_3$ , из уравнения (34) найдём  $x_2$ , а затем —  $x_1$ :

$$\begin{aligned} x_2 &= 0.38 - 0.6 \cdot x_3 = 0.38 - 0.6 \cdot 1.035294117647058820 = \\ &= -0.241176470588235292, \end{aligned}$$

$$\begin{aligned} x_1 &= 0.6 - 3 \cdot x_2 - x_3 = 0.6 - 3 \cdot (-0.241176470588235292) - \\ &- 1.035294117647058820 = 0.288235294117647056. \end{aligned}$$

Итак, решение системы линейных алгебраических уравнений методом Гаусса выполнено за два этапа:

*прямой ход* — приведение системы (22) к трапециевидной форме;

*обратный ход* — определение неизвестных системы.

Как уже отмечалось, этот метод даст точное решение, если все исходные данные точны и все вычисления проводятся точно. Но в расчётах на компьютере округления неизбежны. Они вносят погрешность в гауссовы методы (см. значения неизвестных и пример 14).

Отметим, что схему единственного деления можно использовать для вычисления определителя: определитель







системы (эта схема называется *методом Гаусса с постолбцовым выбором главного элемента*). Более того, целесообразно в роли ведущего элемента брать наибольший по модулю элемент матрицы  $A$  (такая модификация именуется *методом Гаусса с выбором главного элемента по всей матрице*). Это исключает вычитание близких чисел, что позволяет избежать пропадания значащих цифр.<sup>7</sup>

Подсчитаем число  $N$  арифметических операций, необходимых для решения линейной системы с  $n$  неизвестными по схеме единственного деления. Для прямого хода необходимо выполнить

$$n(n+1) + (n-1)n + \dots + 1 \cdot 2 = \\ = (1^2 + 2^2 + \dots + n^2) + (1 + 2 + \dots + n) = \frac{n(n+1)(n+2)}{3}$$

умножений и делений, и столько же вычитаний, а для обратного хода —  $\frac{n(n-1)}{2}$  тех же действий. Следовательно, общее число арифметических действий таково:

$$N = \frac{2n(n+1)(n+2)}{3} + n(n-1) = \frac{n(2n^2 + 9n + 1)}{3}. \quad (43)$$

Заметим, что при  $n > 7$  значение  $N$  удовлетворяет неравенству  $N < n^3$ . Таким образом, время, необходимое для решения линейной системы методом Гаусса, примерно пропорционально кубу числа неизвестных.

**ПРИМЕР 16.** Теперь систему (22) (см. пример 15) решим по схеме единственного деления с помощью компьютера.

*Решение.* Построим схему иерархии. На её первом уровне разместим головной алгоритм с именем МЕТОД ГАУССА. Подчинёнными будут алгоритмы ввода/вывода векторов (дадим им соответственно имена GET\_VECT и

---

<sup>7</sup>Можно уменьшить ошибки округления, если выполнять вычисления с некоторым числом запасных знаков; рабочим типом данных при реализации вычислительных алгоритмов на компьютере должен быть тип **double**, **extended** или **long double** (см. [17]).

PUT\_VECT) и матриц (GET\_MATR и PUT\_MATR), решения системы (GAUSS), а также контроля и вывода результатов (RES\_AX\_EQ\_V). Все они не зависят друг от друга и поэтому их место на одном (втором) уровне подчинения. Формальными параметрами алгоритмов будут:  $A$  — матрица коэффициентов системы,  $x$  — вектор-столбец решения,  $b$  — вектор-столбец свободных членов,  $lim$  — предельная верхняя граница индекса (нижняя граница равна (по умолчанию) единице) и  $n$  — конкретная верхняя граница индекса.

В головном алгоритме определим управляющий символ  $nl$  — “новая линия” и две строковые константы:  $one$  — “одинарная линия” (это отрезок прямой; он разделит исходные и результирующие данные) и  $two$  — “двойная линия” (она окаймлит сверху и снизу все выводные данные).

Из подчинённых алгоритмов требует пояснений, очевидно, только алг GAUSS. Цикл, описывающий прямой ход, троекратный. Во внешнем цикле (цикл по  $k$ ) сначала подсчитаем обратную величину ведущего элемента (тем самым во вспомогательном цикле (цикл по  $j$ ) деление заменим умножением и, кроме того, вместо индексированной переменной используем простую переменную  $p$  — всё это экономит время вычислений). Во вспомогательном цикле первую строку текущей системы умножим на  $p$  (см. формулы (40)), и позже, в двойном цикле (циклы по  $i$  и по  $j$ ), запрограммируем формулы (41). Результатом троекратного цикла будет система трапецевидной формы. Цикл обратной подстановки (см. (42)) двукратный. Отметим, что в нём, во-первых, с целью экономии времени снова используется простая переменная  $p$ , и, во-вторых, внутренних цикл (цикл по  $j$ ) при значении  $j$ , равном  $n + 1$ , никогда не будет выполняться<sup>8</sup>.

<sup>8</sup>Команды **для... выполнять** и **пока... выполнять** — команды повторения с предусловием (см. [17]).







цел  $i, j, k, k_1$   
 двточн  $p$   
 начало {————— Прямой ход:}  
 для  $k$  от 1 до  $n$  шаг 1 выполнять  
 $p \leftarrow 1/a_{kk}$      $k_1 \leftarrow k + 1$   
 для  $j$  от  $k_1$  до  $n$  шаг 1 выполнять  
 $a_{kj} \leftarrow a_{kj} \cdot p$   
 конец-цикл  
 $b_k \leftarrow b_k \cdot p$   
 для  $i$  от  $k_1$  до  $n$  шаг 1 выполнять  
 для  $j$  от  $k_1$  до  $n$  шаг 1 выполнять  
 $a_{ij} \leftarrow a_{ij} - a_{ik}a_{kj}$   
 конец-цикл  
 $b_i \leftarrow b_i - a_{ik}b_k$   
 конец-цикл  
 конец-цикл {————— Обратный ход:}  
 для  $k$  от  $n$  до 1 шаг -1 выполнять  
 $p \leftarrow b_k$   
 для  $j$  от  $k + 1$  до  $n$  шаг 1 выполнять  
 $p \leftarrow p - a_{kj}x_j$   
 конец-цикл  
 $x_k \leftarrow p$   
 конец-цикл  
 возвратиться  
 конец {алг Gauss}  
 {—————}

Приведём результаты решения системы (22) (язык программирования ТР 7.0):

Ответ:	$Ax-b=0$
$x[1] = 2.88235294117647E-0001$	$0.0E+0000$
$x[2] = -2.41176470588235E-0001$	$2.2E-0016$
$x[3] = 1.03529411764706E+0000$	$-4.4E-0016$

Отметим, что при кодировании проекта C++-ом, Фортраном или Бейсиком объявление массивов должно иметь вид:

**двточн массив**  $A[lim, lim], b[lim], x[lim]$   
 $A_{copy}[lim, lim], b_{copy}[lim]$

обращение к алг GAUSS — вид:

**вызвать** COPY\_MATR( $A, lim, n, n, A_{copy}$ )

**вызвать** COPY\_VECT( $b, lim, n, b_{copy}$ )

**вызвать** GAUSS( $A_{copy}, b_{copy}, lim, n, x$ )

а группу подчинённых алгоритмов надо пополнить алгоритмами копирования матриц и векторов<sup>9</sup>:

{  
 \_\_\_\_\_  
 алг COPY\_MATR(двточн массив  $A[lim, lim]$ , цел  $lim, m, n$ ,  
 $A_{copy}[lim, lim]$ )

**знач**  $A, lim, m, n$       **рез**  $A_{copy}$

**цел**  $i, j$

**начало**

**для**  $i$  **от** 1 **до**  $m$  **шаг** 1 **выполнять**

**для**  $j$  **от** 1 **до**  $n$  **шаг** 1 **выполнять**

$a_{copy_{ij}} \leftarrow a_{ij}$

**конец-цикл**

**конец-цикл**

**возвратиться**

**конец** {алг Copy\_matr}

{  
 \_\_\_\_\_  
 алг COPY\_VECT(двточн массив  $x[lim]$ , цел  $lim, n$ ,

$x_{copy}[lim]$ )

**знач**  $x, lim, n$       **рез**  $x_{copy}$

**цел**  $i$

**начало**

**для**  $i$  **от** 1 **до**  $n$  **шаг** 1 **выполнять**

<sup>9</sup>Фортран 90, 95 и 2000 допускает массивовое присваивание:  $A_{copy} = A$  и  $b_{copy} = b$ .

$$x_{\text{copy}_i} \leftarrow x_i$$

конец-цикл

возвратиться

конец

{алг Copy\_vect}

{\_\_\_\_\_}

Они необходимы, так как алгоритм Гаусса является алгоритмом с *побочным* или, другими словами, с *волновым эффектом*: он, кроме поиска неизвестных, изменяет, “портит” исходную систему (на её месте получается система трапециевидной формы). А так как матрица  $A$  и вектор  $\mathbf{b}$  должны ещё участвовать в оценке качества решения (в вычислении вектора-невязки  $\mathbf{r} = A\mathbf{x} - \mathbf{b}$ ), то надо принять меры для их сохранения. Язык программирования Паскаль имеет средства копирования параметров-значений (см., например, фразу `знач A, b, lim, n`) и средства указания адресов памяти параметров-результатов (см. фразу `рез x`). Языки C++ (в случае массивов), Фортран и Бейсик такими средствами не обладают: значит, пользователь перед вызовом алгоритма GAUSS должен сам позаботиться о создании копий матрицы коэффициентов системы и вектора-столбца свободных членов.  $\triangle$

Подмена арифметики действительных чисел компьютерной арифметикой, как правило, отрицательно отражается на результате решения системы  $A\mathbf{x} = \mathbf{b}$ . Если матрица  $A$  неособенная (невырожденная), то  $\mathbf{x} = A^{-1}\mathbf{b}$ . Вычисленное методом Гаусса (или любым другим методом) решение  $\tilde{\mathbf{x}}$  отличается от точного  $\mathbf{x}$  из-за ошибок округления<sup>10</sup>. Степень отклонения гауссова решения от точного можно характеризовать разностью  $\Delta = \mathbf{x} - \tilde{\mathbf{x}}$ . Но точное решение, как правило, неизвестно. Проверка путём подстановки  $\tilde{\mathbf{x}}$  в систему даёт невязку  $\mathbf{r} = A\tilde{\mathbf{x}} - \mathbf{b}$ . Её мы подсчитали в алгоритме МЕТОД ГАУССА. Можно показать, что если  $\Delta = 0$ , то и  $\mathbf{r} = 0$  [10]. Более того, если  $\Delta \approx 0$ , то, как правило,  $\mathbf{r} \approx 0$ .

<sup>10</sup>Напомним, что ошибки округления — результат конечности разрядной сетки компьютера.

Обратное же утверждение не всегда верно (например, для плохо обусловленной системы при  $\mathbf{r} \approx 0$  погрешность решения может быть большой<sup>11</sup>). Следовательно, *посредством невязки можно выполнять контроль качества решения, если система  $A\mathbf{x} = \mathbf{b}$  не является плохо обусловленной*. В таком случае метод Гаусса (с выбором главного элемента) даёт малые невязки  $r_i = \sum_{j=1}^n a_{ij}\tilde{x}_j - b_i$  ( $i = 1, 2, \dots, n$ ).

**Обсуждение.** Схема единственного деления является простой, эффективной (число операций, необходимых для решения системы, лишь немногим превышает число операций, требуемых для подсчёта одного определителя<sup>12</sup>) и удобной для программирования. *Необходимое и достаточное условие её применимости — неравенство нулю всех ведущих элементов*. К недостаткам схемы относятся отсутствие критерия, по которому ещё до начала вычислений можно было бы установить выполнимость условия  $a_{kk}^{(k-1)} \neq 0$ , и потеря точности при малом по модулю  $a_{kk}^{(k-1)}$  (в сравнении с другими коэффициентами в строке).

Вычислительные погрешности будут меньшими, если в роли ведущего элемента брать главный элемент. Соответствующий алгоритм является универсальным и рекомендуется для решения систем невысокого порядка ( $n \leq 200$ ). К настоящему времени не найдено алгоритмов, лучших по эффективности (времени решения, памяти), чем методы исключения Гаусса (ПОСТУЛАТ. Гаусс всегда прав.).

**Упражнение 4.** 1. Решите систему (22) методом Гаусса

<sup>11</sup>Оцените решение системы  $\begin{pmatrix} 0.780 & 0.563 \\ 0.913 & 0.659 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.217 \\ 0.254 \end{pmatrix}$ , полученное на трёхзначной и семизначной разрядной сетке (здесь  $x_1 = 1$ ,  $x_2 = -1$ ,  $\det A = 10^{-8}$ ; решение с двучн таково:  $\tilde{x}_1 = 0.999999999914390$ ,  $\tilde{x}_2 = -0.9999999999881394$ ,  $r_1 = -2.8 \cdot 10^{-17}$ ,  $r_2 = 0$ ).

<sup>12</sup>Эффективность методов решения системы (8) зависит от структуры и свойств матрицы  $A$  (порядка, обусловленности, симметричности, заполненности, специфики расположения ненулевых элементов), параметров компьютера, способа представления данных и специфики реализации алгоритмов (см. [4]).

с выбором главного элемента

1) по столбцам матрицы (чтобы частичное упорядочивание было более эффективным, сначала производите масштабирование (уравновешивание) системы делением элементов каждой строки на наибольший по модулю элемент строки);

2) по всей матрице.

2. Решите системы  $\begin{matrix} 0.0001x_1+x_2=1, \\ x_1+x_2=2 \end{matrix}$  и  $\begin{matrix} x_1+x_2=2, \\ 0.0001x_1+x_2=1 \end{matrix}$  методом Гаусса, удерживая в вычислениях только три десятичных значащих цифры.

3. Решите систему  $\begin{matrix} x_1 + 592x_2 = 437, \\ 592x_1 + 4308x_2 = 2251 \end{matrix}$  методом Гаусса и методом Гаусса с выбором главного элемента, выполняя подсчёты с точностью до четырёх значащих цифр.

4. То же, что в задаче 3, но для системы

$$\begin{aligned} 2x_1 + 6x_2 - x_3 &= -12, \\ 5x_1 - x_2 + 2x_3 &= 29, \\ -3x_1 - 4x_2 + x_3 &= 5. \end{aligned}$$

5. То же, что в задаче 3, но для систем

$$\begin{aligned} 10x_1 - 7x_2 &= 7, & 10x_1 - 7x_2 &= 7, \\ -3x_1 + 2.099x_2 + 6x_3 &= 3.901, & \text{и} & -3x_1 + 2x_2 + 6x_3 = 4, \\ 5x_1 - x_2 + 5x_3 &= 6 & 5x_1 - x_2 + 5x_3 &= 6. \end{aligned}$$

7. Разработайте алгоритм решения системы  $n$  уравнений с  $n$  неизвестными методом Гаусса с выбором главного элемента

1) по столбцам матрицы;

2) по всей матрице.

8. Разработайте алгоритм вычисления определителя  $n$ -го порядка методом Гаусса с выбором главного элемента

1) по столбцам матрицы;

2) по всей матрице.

9. Модифицируйте алгоритм Гаусса, заменив матрицу коэффициентов при неизвестных расширенной матрицей,

1) последний столбец которой есть столбец свободных членов;

2) последние  $m$  столбцов которой есть  $m$  столбцов свободных членов.

10\*. Разработайте алгоритм решения системы линейных алгебраических уравнений с трёхдиагональной матрицей (1). *Замечание.* Так как каждое уравнение этой ленточной системы связывает три “соседних” неизвестных, то, очевидно, не следует хранить в памяти компьютера все  $n^2$  коэффициентов матрицы  $A$ ; алгоритм нужно построить так, чтобы в вычислениях участвовали только элементы трёх диагоналей [17].

11. Предложите свой “точный” метод решения систем линейных алгебраических уравнений (8).

### 3. Итеративные методы

Итеративные методы (их около десятка [14]) дают приближённое решение системы (8). Несмотря на это они обладают качествами, которые делают их выгодными с вычислительной точки зрения. Решение системы получается как предел некоторой единым образом построенной последовательности приближений.

#### 3.1 Метод Гаусса — Зейделя

Дадим краткое описание метода Гаусса — Зейделя<sup>13</sup>. Сущность метода изложим сначала на примере системы двух уравнений с двумя неизвестными.

ПРИМЕР 17. Итак, возьмём следующую систему:

$$2x_1 + x_2 = 3,$$

$$x_1 + 2x_2 = 3,$$

( $\det A = 3$ , точное решение есть  $x_1 = 1$ ,  $x_2 = 1$ ). Решим её методом Зейделя и геометрически проиллюстрируем процесс итераций (рис. 1).

<sup>13</sup>Будем его называть просто: *метод Зейделя* [3].

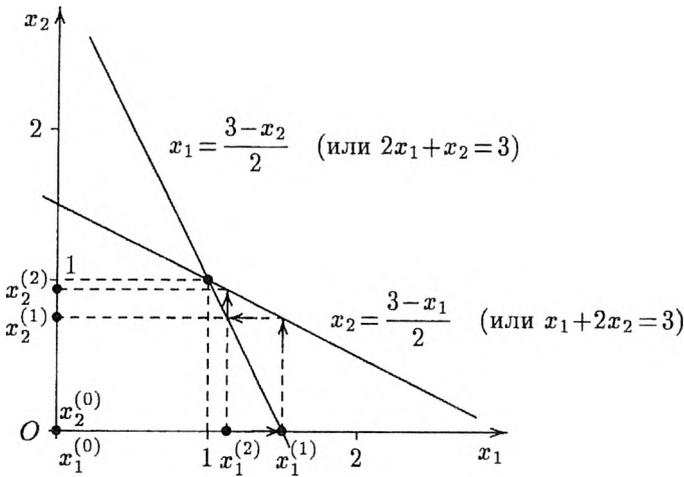


Рис. 1.

*Решение.* Разрешим первое уравнение системы относительно  $x_1$ , а второе — относительно  $x_2$ . Получим эквивалентную систему вида

$$x_1 = \frac{3 - x_2}{2}, \quad (44)$$

$$x_2 = \frac{3 - x_1}{2}. \quad (45)$$

Зададимся некоторым начальным приближением значений неизвестных  $x_1 = x_1^{(0)}$  и  $x_2 = x_2^{(0)}$ : двигаясь, например, от начала координат, положим  $x_1^{(0)} = x_2^{(0)} = 0$ . Значение  $x_2^{(0)} = 0$  подставим в уравнение (44), найдём первое приближение первого корня  $x_1^{(1)} = 3/2$  (при вычислении  $x_1^{(1)}$  значение  $x_2^{(0)}$  остаётся неизменным, геометрически это соответствует движению по горизонтали до пересечения с прямой  $2x_1 + x_2 = 3$ ; точка пересечения имеет координаты  $(1.5, 0)$ ). Затем, подставляя  $x_1^{(1)} = 3/2$  в уравнение (45), вычислим первое приближение второго корня  $x_2^{(1)} = 3/4$  (при постоянном значении  $x_1^{(1)} = 1.5$  идём по вертикали, пока не пересечём прямую  $x_1 + 2x_2 = 3$  в точке с координатами  $(1.5, 0.75)$ ).

тами  $(1.5, 0.75)$ ). На этом заканчивается первая итерация решения системы (44)–(45). Дальнейшие шаги производим подобным образом (они показаны горизонтальными и вертикальными отрезками со стрелками): значение  $x_2^{(1)} = 3/4$  подставим в уравнение (44), определим второе приближение первого корня  $x_1^{(2)} = 9/8 = 1.125$ ; потом это число используем при нахождении второго приближения второго корня  $x_2^{(2)} = 15/16 = 0.9375$ , ...,  $x_1^{(6)} = 2049/2048 \approx 1.0004882$ ,  $x_2^{(6)} = 4095/4096 \approx 0.9997558$  и т. д. Процесс итераций сходится. Последовательные приближения стремятся к единице. Они подсчитывались по формулам

$$\begin{aligned} x_1^{(k+1)} &= \frac{3 - x_2^{(k)}}{2}, \\ x_2^{(k+1)} &= \frac{3 - x_1^{(k+1)}}{2} \quad (k = 0, 1, \dots). \end{aligned}$$

Здесь верхний индекс в скобках  $(k)$ ,  $(k+1)$  означает порядковый номер приближения. Заметим, что если вторую формулу записать в виде  $x_2^{(k+1)} = (3 - x_1^{(k)})/2$ , то новая система

$$\begin{aligned} x_1^{(k+1)} &= \frac{3 - x_2^{(k)}}{2}, \\ x_2^{(k+1)} &= \frac{3 - x_1^{(k)}}{2} \end{aligned}$$

представляет собой *метод (простых) итераций*, или *метод Якоби* [3] (скорость сходимости процесса при этом может замедлиться).

Начальное приближение произвольно: за нулевое приближение можно принять, например, и столбец свободных членов системы.  $\triangle$

Решим теперь *систему  $n$  уравнений с  $n$  неизвестными* (8). В этом случае зейделевы итерации будут выпол-

няться по формулам

$$\begin{aligned}
 x_1^{(k+1)} &= \left( b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \dots - a_{1n}x_n^{(k)} \right) / a_{11}, \\
 x_2^{(k+1)} &= \left( b_2 - a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} - \dots - a_{2n}x_n^{(k)} \right) / a_{22}, \\
 &\dots\dots\dots (46) \\
 x_n^{(k+1)} &= \left( b_n - a_{n1}x_1^{(k+1)} - a_{n2}x_2^{(k+1)} - \dots - a_{n,n-1}x_{n-1}^{(k+1)} \right) / a_{nn}.
 \end{aligned}$$

Формулы (46) запишем короче в виде

$$x_i^{(k+1)} = \left( b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right) / a_{ii}, \quad (47)$$

$x_i^{(0)}$  произвольно ( $i = 1, 2, \dots, n$ ;  $k = 0, 1, \dots$ ).

Процесс итераций продолжаем до тех пор, пока значения  $x_i^{(k+1)}$  ( $i = 1, 2, \dots, n$ ) не станут близкими к значениям  $x_i^{(k)}$  с заданной погрешностью  $\varepsilon$ , домноженной на некоторое число (см. п. 3.3).

### 3.2 Сходимость итеративного процесса

Сформулируем достаточное условие сходимости процесса последовательных приближений [7].

**ТЕОРЕМА 2. (О сходимости процесса итераций.)** *Процесс итераций (47) сходится к единственному решению  $\mathbf{x} = \lim_{k \rightarrow \infty} \mathbf{x}^{(k+1)}$  системы  $\sum_{j=1}^n a_{ij}x_j = b_i$  ( $i = 1, 2, \dots, n$ ) при любом выборе начального приближения  $\mathbf{x}^{(0)}$ , если для коэффициентов при неизвестных выполняется следующее неравенство:*

$$\sum_{\substack{j=1 \\ j \neq i}}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1 \quad (i = 1, 2, \dots, n). \quad (48)$$

Другими словами, процесс итераций сходится, если в матрице  $A = (a_{ij})_{i,j=1}^n$  имеет место диагональное преобладание:

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \quad (i = 1, 2, \dots, n), \quad (49)$$

т. е., когда модули диагональных элементов матрицы для каждой строки превышают сумму модулей недиагональных элементов этой строки.

**ПРИМЕР 18.** Заметим, что признак сходимости (49) легко проверяется. Возьмём систему (22). Она не удовлетворяет условию сходимости:

$$\begin{aligned} |a_{11}| &> |a_{12}| + |a_{13}|, \\ |a_{22}| &> |a_{21}| + |a_{23}|, \\ |a_{33}| &> |a_{31}| + |a_{32}|. \quad \triangle \end{aligned} \quad (50)$$

Условие (49) является достаточным для сходимости метода, но оно не является необходимым, т. е. для некоторых систем приближения сходятся и при невыполнении указанного условия.

Начальное приближение  $x_i^{(0)}$  ( $i = 1, 2, \dots, n$ ) произвольное потому, что сходимость итеративного процесса определяется лишь природой матрицы  $A$ . Быстрота сходимости зависит не только от свойств матрицы коэффициентов (процесс итераций хорошо сходится, когда модули диагональных коэффициентов системы велики, т. е. когда выполняется неравенство  $|a_{ii}| \gg \sum_{j=1}^n |a_{ij}|$  ( $i = 1, 2, \dots, n; j \neq i$ )), но и от удачного выбора начального приближения. Целесообразно в качестве компонентов вектора  $\mathbf{x}^{(0)}$  взять приближённые значения неизвестных (они получены, наверное, методом Гаусса, или найдены из физических соображений). Иногда берут  $\mathbf{x}^{(0)} = \mathbf{b}$ , и это бывает удобно.

Число арифметических действий, необходимых для выполнения одной итерации, пропорционально  $n^2$ . Всего же потребуется произвести примерно  $kn^2$  действий, где  $k$  — общее число итераций.

### 3.3 Оценка погрешности метода Зейделя

Итак, точное решение системы  $Ax = b$  есть результат бесконечного процесса и всякий вектор  $x^{(k+1)}$  из последовательности полученных векторов является приближённым решением. Если выполнено неравенство (49), то оценка погрешности приближённого решения  $x^{(k+1)}$  даётся формулой [7]

$$\max_i |x_i^{(k+1)} - x_i^{(k)}| \leq \frac{1-q}{q} \varepsilon, \quad (51)$$

где  $0 < q = \max_i \sum_{\substack{j=1 \\ j \neq i}}^n \left| \frac{a_{ij}}{a_{ii}} \right| < 1$  ( $i = 1, 2, \dots, n$ ), а  $\varepsilon$  — заданная допустимая погрешность.

ПРИМЕР 19. Условие сходимости (50) для следующей системы

$$\begin{aligned} 7x_1 + 2x_2 - x_3 &= 0.5, \\ x_1 + 3x_2 + x_3 &= 0.6, \\ -2x_1 + x_2 + 9x_3 &= 8.5, \end{aligned} \quad (52)$$

выполнено. Тогда

$$q = \max \left( \left| \frac{2}{7} \right| + \left| \frac{-1}{7} \right|, \left| \frac{1}{3} \right| + \left| \frac{1}{3} \right|, \left| \frac{-2}{9} \right| + \left| \frac{1}{9} \right| \right) = \frac{2}{3}.$$

△

Формула (51) — это критерий по абсолютным отклонениям. При  $|x_i| \gg 1$  его можно заменить критерием по

относительным разностям, т. е. условие окончания итеративного процесса записать в виде

$$\max_i |x_i^{(k+1)} - x_i^{(k)}| \leq \frac{1-q}{q} \varepsilon |x_i^{(k+1)}|. \quad (53)$$

При выполнении условия (51) или (53) итеративный процесс Гаусса — Зейделя называется *сходящимся*. В этом случае максимальные разности значений неизвестных в двух последовательных итерациях убывают, а сами значения стремятся к решению системы уравнений.

Сделаем замечание о *точности расчётов*. Технический подход к учёту погрешностей действий (*принцип А. Н. Крылова* [4]) состоит в следующем. Если коэффициенты и свободные члены системы  $Ax = b$  являются *точными числами*, то итеративным методом её решение можно получить с любым заранее заданным количеством  $m$  верных десятичных знаков. При этом в выкладках будем удерживать  $m + 1$  десятичных знаков, после чего конечные результаты округлим на один знак. Если же исходные данные — *приближённые числа* с  $\ell$  знаками, то, как в случае точных чисел, найдём решение с точностью до  $m = \ell$  знаков.

### 3.4 Приведение линейной системы к виду, пригодному для итераций

Теорема сходимости 2 предъявляет жёсткие требования к коэффициентам системы  $Ax = b$ . Но, если  $\det A \neq 0$ , то эквивалентными преобразованиями (см. с. 153) исходную систему всегда можно заменить системой, для которой условия (49) будут выполнены. Сделать это можно бесконечным числом способов. Поступим, например, так. Из заданной системы выделим уравнения с коэффициентами, модули которых больше суммы модулей остальных коэффициентов при неизвестных. Каждое выделенное уравнение запишем в такую строку новой системы, чтобы наибольший по модулю коэффициент оказался на диагонали. Затем из оставшихся неиспользованных (и, возможно, вы-

деленных уравнений) составим линейно независимые комбинации с таким расчётом, чтобы был выдержан указанный выше принцип формирования новой системы, и все свободные строки в ней оказались заполненными. При этом позаботимся, чтобы каждое неиспользованное ранее уравнение попало хотя бы в одну линейную комбинацию. Поясним сказанное на примере.

**ПРИМЕР 20.** Снова решим систему (22), но теперь — методом Зейделя.

*Решение.* Эта система

$$x_1 + 3x_2 + x_3 = 0.6, \quad (54)$$

$$2x_1 + x_2 - x_3 = -0.7, \quad (55)$$

$$x_1 - x_2 + 2x_3 = 2.6 \quad (56)$$

не удовлетворяет условию сходимости (50). Но в уравнении (54) коэффициент при  $x_2$  превышает сумму остальных коэффициентов строки (не считая свободного члена):  $3 > 1 + 1$ . Поэтому (54) можно взять в роли второго уравнения новой системы

$$\dots\dots\dots (*)$$

$$x_1 + 3x_2 + x_3 = 0.6, \quad (**)$$

$$\dots\dots\dots (***)$$

Уравнение (\*) с преобладающим по модулю коэффициентом при  $x_1$  можно получить, умножив уравнение (55) на 3 и сложив произведение с уравнением (56):

$$3(2x_1 + x_2 - x_3) = -0.7 \cdot 3$$

$$x_1 - x_2 + 2x_3 = 2.6$$

---


$$7x_1 + 2x_2 - x_3 = 0.5 \quad (*)$$

И, наконец, умножим уравнения (54) и (56) на 2, сложим полученные произведения с уравнением (55), предварительно умноженным на  $-3$ , получим третье уравнение

СИСТЕМЫ:

$$\begin{aligned} 2(x_1 + 3x_2 + x_3) &= 0.6 \cdot 2 \\ -3(2x_1 + x_2 - x_3) &= -0.7 \cdot (-3) \\ 2(x_1 - x_2 + 2x_3) &= 2.6 \cdot 2 \end{aligned}$$

---


$$-2x_1 + x_2 + 9x_3 = 8.5 \quad (***)$$

Таким образом, получена система (52), эквивалентная исходной (54)–(56) и удовлетворяющая условию (50) (см. пример 19; там же найдено число  $q = 2/3$ ).

Теперь решим эту систему на компьютере. Головному алгоритму дадим имя МЕТОД ГАУССА\_ЗЕЙДЕЛЯ, а алг GAUSS заменим на алг G\_SEIDEL. В схему иерархии добавим алг двточн Q, вычисляющий  $q$  (см. (51)). Остальные алгоритмы в схеме те же, что в примере 16.

Поясним алг G\_SEIDEL. Его формальные параметры таковы:  $A$  — матрица размера  $n \times n$  из коэффициентов при неизвестных,  $b$  — столбец свободных членов,  $x^{(0)}$  — столбец начальных приближений,  $lim$  и  $n$  — максимальный и фактический размеры,  $\varepsilon$  — абсолютная погрешность,  $k_{max}$  — максимальное (запланированное) число итеративных шагов,  $x$  — столбец из неизвестных и  $k_{end}$  — фактическое число шагов (одновременно — фиксатор ошибок: если  $k_{end}=0$ , то неравенство (51) ни разу не выполнилось (причины этого: нарушены условия сходимости, слишком малые  $\varepsilon = \varepsilon_{corr}$  и (или)  $k_{max}$ ); в противном случае получено решение системы).

Тело алг G\_SEIDEL состоит из двух последовательных команд повторения: первая играет вспомогательную роль (назначает начальные значения элементам столбца из неизвестных, вычисляет обратные величины диагональных элементов), вторая реализует процесс Зейделя. Цикл приближений по Зейделю тоекратный. Во внешнем цикле

(цикл “по  $\varepsilon$ ”) меняется значение счётчика  $k$  от единицы с шагом, равным единице, до заданного предельного числа итераций  $k_{\max}$ . В конце его при истинном неравенстве (51) (тогда  $k_{\text{end}} \neq 0$ ) выполняется возврат к точке вызова данного алгоритма. В цикле “по строкам” (цикл “по  $i$ ”) анализируется условие (51) и подсчитываются текущие приближения неизвестных. В цикле “по столбцам” (цикл “по  $j$ ”) из правых частей уравнений вычитаются левые части.

Приведём проект программы:

{ \_\_\_\_\_ }

алг МЕТОД ГАУССА\_ЗЕЙДЕЛЯ

конст

$nl \leftarrow \#13\#10$

$lim \leftarrow 80$

$Ed \leftarrow 2.23e-16$       {  $\varepsilon_{\text{компьютерный double}} = 2.23 \cdot 10^{-16}$  }

$one \leftarrow \text{_____}$

$two \leftarrow \text{_____}$

цел  $n, k_{\max}, k$

двточн  $q, \varepsilon, E$

двточн массив  $A[lim, lim], b[lim], x^{(0)}[lim], x[lim]$

начало

очистить экран

вывести  $nl, nl, nl, \text{ } \sqcup n, k_{\max} \implies \sqcup$

вести  $n, k_{\max}$

вывести  $nl, \text{ } \sqcup A[n, n] \implies \text{ } \sqcup, nl$

вызвать GET\_MATR( $A, lim, n, n$ )

вывести  $nl, \text{ } \sqcup b[n] \implies \text{ } \sqcup, nl$

вызвать GET\_VECT( $b, lim, n$ )

вывести  $nl, \text{ } \sqcup x^{(0)}[n] \implies \text{ } \sqcup, nl$

вызвать GET\_VECT( $x^{(0)}, lim, n$ )

$q \leftarrow Q(A, lim, n)$

пока истина выполнять

вывести  $nl, \text{ } \sqcup \varepsilon > 2.23e-16 \implies \text{ } \sqcup$

вести  $\varepsilon$

$E \leftarrow ((1 - q)/q)\varepsilon$

если  $E > Ed$

то

прервать

{break или exit}

иначе

вывести  $nl$ , '\*\*\* error: увеличьте  $\epsilon$ .  $E=$ ',  $E$ ,  $nl$

конец-если

конец-цикл

{ввода  $\epsilon$ }

очистить экран

вывести  $nl$ ,  $nl$ ,  $two$ ,

$nl$ ,  $nl$ , 'Task № 3 var 28.  $Ax = b$ .'

'Метод Гаусса - Зейделя.'

$nl$ , 'Coded by Незайцев Лев гр. 1057/1 СПб ГПУ.'

$nl$ ,  $nl$ , 'Исходные данные.'

$nl$ ,  $nl$ , ' $n =$ ',  $n$ , ' $\epsilon =$ ',  $\epsilon$ , ' $\epsilon_{\text{сост}} =$ ',  $E$ ,

$nl$ , ' $q =$ ',  $q$ , ' $k_{\text{max}} =$ ',  $k_{\text{max}}$ ,

$nl$ ,  $nl$ , 'Матрица  $A$ :',  $nl$

вызвать PUT\_MATR( $A$ ,  $lim$ ,  $n$ ,  $n$ )

вывести  $nl$ , 'Вектор  $b$ :',  $nl$

вызвать PUT\_VECT( $b$ ,  $lim$ ,  $n$ )

вывести  $nl$ , 'Вектор  $x^{(0)}$ :',  $nl$

вызвать PUT\_VECT( $x^{(0)}$ ,  $lim$ ,  $n$ )

вызвать G\_SEIDEL( $A$ ,  $b$ ,  $x^{(0)}$ ,  $lim$ ,  $n$ ,  $E$ ,  $k_{\text{max}}$ ,  $x$ ,  $k$ )

вывести  $nl$ ,  $one$ ,

$nl$ ,  $nl$ , 'Ответ:  $Ax - b = 0$ ',  $nl$ ,  $nl$

если  $k = 0$

то

вывести  $nl$ , '\*\*\* error: за  $k_{\text{max}}$ ,

'шагов процесс не сошёлся. G-Seidel.'

$nl$ , 'В этом случае',  $nl$ ,  $nl$

конец-если

вызвать RES\_AX\_EQ\_V( $A$ ,  $x$ ,  $b$ ,  $lim$ ,  $n$ )

вывести  $nl$ , 'Число итераций  $k =$ ',  $k$ ,

$nl$ ,  $nl$ ,  $two$ ,

$nl$ ,  $nl$ , '23.08.2000 Всё ...',  $nl$ ,  $nl$

конец

{головного алгоритма}

{\_\_\_\_\_}

алг двточн Q (двточн массив  $A[lim, lim]$ , цел  $lim, n$ )

знач  $A, lim, n$

двточн  $q, s$       цел  $i, j$

начало

$q \leftarrow 0$

для  $i$  от 1 до  $n$  шаг 1 выполнять

$s \leftarrow 0$

для  $j$  от 1 до  $n$  шаг 1 выполнять

если  $j \neq i$  то  $s \leftarrow s + |a_{ij}/a_{ii}|$  конец-если

конец-цикл

{ по  $j$  }

если  $q < s$  то  $q \leftarrow s$  конец-если

конец-цикл

{ по  $i$  }

$Q \leftarrow q$

возвратиться

конец

{ алг Q }

{ \_\_\_\_\_ }

алг GET\_VECT (...

текст этой части проекта см. на сс. 162 – 163

возвратиться

конец

{ алг Res\_Ax\_eq\_b }

{ \_\_\_\_\_ }

алг G\_SEIDEL (двточн массив  $A[lim, lim]$ ,  $b[lim]$ ,  $x^{(0)}[lim]$ ,

цел  $lim, n$ , двточн  $\varepsilon$ , цел  $k_{\max}$ ,

двточн массив  $x[lim]$ , цел  $k_{\text{end}}$ )

знач  $A, b, x^{(0)}, lim, n, \varepsilon, k_{\max}$

рез  $x, k_{\text{end}}$

{ Если  $k_{\text{end}}=0$ , то  $|\Delta| > \varepsilon$  }

цел  $i, k, j$

{ Если  $k_{\text{end}} \neq 0$ , то  $|\Delta| \leq \varepsilon$  }

двточн  $t, x_{\text{new}}$

двточн массив  $p[lim]$

начало

для  $i$  от 1 до  $n$  шаг 1 выполнять

$x_i \leftarrow x_i^{(0)}$        $p_i \leftarrow 1/a_{ii}$

конец-цикл

{ подготовки }

для  $k$  от 1 до  $k_{\max}$  шаг 1 выполнять

$k_{\text{end}} \leftarrow k$

```

для  $i$  от 1 до  $n$  шаг 1 выполнять
     $t \leftarrow b_i$ 
    для  $j$  от 1 до  $n$  шаг 1 выполнять
        если  $j \neq i$  то  $t \leftarrow t - a_{ij}x_j$  конец-если
    конец-цикл {по  $j$ }
     $x_{\text{new}} \leftarrow t \cdot p_i$ 
    если  $|x_{\text{new}} - x_i| > \varepsilon$  то  $k_{\text{end}} \leftarrow 0$  конец-если
     $x_i \leftarrow x_{\text{new}}$ 
конец-цикл {по  $i$ }
если  $k_{\text{end}} \neq 0$  то возвратиться конец-если
конец-цикл {по  $k, \varepsilon$ }
возвратиться
конец {алг G_Seidel}
{_____}

```

Приведём результаты решения системы (52) ( $\varepsilon = 10^{-15}$ ,  $E = \varepsilon_{\text{сог}} = 5 \cdot 10^{-16}$ ,  $q = 0.(6)$ ,  $k_{\text{max}} = 100$ ,  $x^{(0)} = b$ ; язык программирования BC++):

Ответ:	$Ax-b=0$
$x[1] = 2.88235294117647e-01$	$4.4e-16$
$x[2] = -2.41176470588235e-01$	$-1.1e-16$
$x[3] = 1.03529411764706e+00$	$-1.8e-15$

Число итераций  $k=29$

△

**Обсуждение.** Метод Зейделя прост и поэтому сравнительно легко программируется. Он имеет преимущество перед прямыми методами, так как в нём ошибки округления сказываются меньше и в процессе решения матрица  $A$  и вектор  $b$  не портятся<sup>14</sup>. Его можно применять для систем

<sup>14</sup>Об этих достоинствах надо говорить в условиях сходимости (49). Недостаток: при приведении линейной системы к виду, пригодному для итераций, надо проделать немалую работу (см. пример 20).

с матрицами высокого порядка.

Решая систему методом Зейделя, мы снова встречаемся с замечательным свойством самоисправляемости сходящегося итеративного процесса: ошибка, допущенная в вычислениях, не отражается на окончательном результате, так как ошибочное приближение можно рассматривать как новый начальный вектор (это свойство усиливает надёжность любого итеративного алгоритма; см. гл. I).

Итак, в условиях сходимости итеративные методы предпочтительнее прямых. Они просты, позволяют найти решение с заданной степенью точности<sup>15</sup>  $\varepsilon$  и удобны при уточнении решений, полученных с помощью прямых методов.

**Упражнение 5.** 1. Решите системы методами Зейделя и Якоби (в (46) справа от знака = только  $k$ -е приближения). Приведите геометрическую интерпретацию процесса итераций;  $\varepsilon = 10^{-4}$ :

$$1) 2x_1 + x_2 = 2,$$

$$x_1 - 2x_2 = -2;$$

$$а) x_1^{(k+1)} = 1 - x_2^{(k)}/2, \quad б) x_1^{(k+1)} = -2 + 2x_2^{(k)},$$

$$x_2^{(k+1)} = 1 + x_1^{(k+1)}/2; \quad x_2^{(k+1)} = 2 - 2x_1^{(k+1)};$$

$$2) x_1 + 2x_2 = 3, \quad x_1^{(k+1)} = 3 - 2x_2^{(k)},$$

$$x_1 - 4x_2 = -3; \quad x_2^{(k+1)} = 3/4 + x_1^{(k+1)}/4.$$

2. Решите системы методом Зейделя;  $\varepsilon = 10^{-4}$ :

$$10x_1 + x_2 = 11, \quad 10x_1 + x_2 = 11.01,$$

$$100x_1 + 1001x_2 = 1101; \quad \text{и} \quad 100x_1 + 1001x_2 = 1101.$$

3. Решите системы методом простых итераций;  $\varepsilon = 10^{-4}$ :

$$5x_1 + 7x_2 = 12, \quad 5x_1 + 7x_2 = 12.075,$$

$$7x_1 + 10x_2 = 17; \quad \text{и} \quad 7x_1 + 10x_2 = 16.905.$$

---

<sup>15</sup>Нетрудно понять, что всё сказанное справедливо лишь до тех пор, пока на погрешность метода (остаточную погрешность, погрешность аппроксимации) не наложится вычислительная погрешность (погрешность округлений), неизбежная в компьютерных расчётах.

4. Решите систему  $Ax = b$ : а) простым методом Гаусса по схеме единственного деления, используя *delta*-флажок (см. [4]); б) улучшенным методом Гаусса (с постролбцовым выбором главного элемента).

$$1) A = \begin{pmatrix} 3 & 1 & 2 & 4 \\ -5 & 2 & 6 & 7 \\ 0 & 3 & 4 & 1 \\ 0 & 0 & -5 & -6 \end{pmatrix}, \quad b = \begin{pmatrix} 4 \\ -13 \\ 1 \\ -5 \end{pmatrix};$$

$$2) A = \begin{pmatrix} 0.2 & 0 & 0.2 & 0 & 0 \\ 0 & 0.2 & 0 & 0.2 & 0 \\ 0.2 & 0 & 0.2 & 0 & 0.2 \\ 0 & 0.2 & 0 & 0.2 & 0 \\ 0 & 0 & 0.2 & 0 & 0.2 \end{pmatrix}, \quad b = \begin{pmatrix} 4.2 \\ 4.2 \\ 4.2 \\ 4.2 \\ 4.2 \end{pmatrix};$$

$$3) A = \begin{pmatrix} 1.342 & 0.432 & -0.599 & 0.202 \\ 0.202 & 1.342 & 0.432 & -0.599 \\ -0.599 & 0.202 & 1.342 & 0.432 \\ 0.432 & -0.599 & 0.202 & 1.342 \end{pmatrix}, \quad b = \begin{pmatrix} 1.941 \\ -0.230 \\ -1.941 \\ 0.230 \end{pmatrix};$$

$$4) A = \begin{pmatrix} 5.1 & 7.1 & 6.1 & 5.1 \\ 7.2 & 10.2 & 8.2 & 7.2 \\ 6.1 & 8.1 & 10.1 & 9.1 \\ 5.2 & 7.2 & 9.2 & 10.2 \end{pmatrix}, \quad b = \begin{pmatrix} 23 \\ 32 \\ 33 \\ 31 \end{pmatrix};$$

$$5) A = \begin{pmatrix} 6.22 & 1.42 & -1.72 & 1.91 \\ 1.42 & 5.33 & 1.11 & -1.82 \\ -1.72 & 1.11 & 5.24 & 1.42 \\ 1.91 & -1.82 & 1.42 & 6.55 \end{pmatrix}, \quad b = \begin{pmatrix} 7.53 \\ 6.06 \\ 8.05 \\ 8.06 \end{pmatrix};$$

$$6) A = \begin{pmatrix} 6.2 & 0.1 & 0.3 & 0.4 & -0.5 \\ 0.5 & 5.4 & 0.1 & 0.2 & 0.4 \\ 0 & 0.4 & 6.3 & 0.6 & 0.2 \\ 0 & 0 & 0.4 & 5.3 & 0.5 \\ 0 & 0 & 0 & 0.1 & 6.2 \end{pmatrix}, \quad b = \begin{pmatrix} 0.1 \\ 0.4 \\ -0.3 \\ 1.1 \\ -0.1 \end{pmatrix};$$

$$7) A = \begin{pmatrix} 2.1 & 2.1 & -1.1 & 1.1 \\ 4 & 3.05 & -1.05 & 2.12 \\ 8.1 & 4.99 & -3.1 & 4.08 \\ 3.01 & 3.02 & -1.98 & 2.03 \end{pmatrix}, \quad b = \begin{pmatrix} 4.1 \\ 6.05 \\ 12.034 \\ 6 \end{pmatrix};$$

$$8) A = \begin{pmatrix} 2 & 3.2 & 11.1 & 5.03 \\ 1.04 & 0.98 & 5.03 & 2.12 \\ 2.1 & 1 & 3.01 & 2.08 \\ 1.1 & 0.95 & 2.95 & 4.06 \end{pmatrix}, \quad b = \begin{pmatrix} 2.1 \\ 0.97 \\ -3.02 \\ -3.01 \end{pmatrix};$$

$$9) A = \begin{pmatrix} 2.1 & 5.05 & 4 & 0.98 \\ 1.03 & 3 & 2.09 & 0.99 \\ 2.05 & 10.01 & 9.08 & 7.1 \\ 2.98 & 7.98 & 8.99 & 2.01 \end{pmatrix}, \quad b = \begin{pmatrix} 20.1 \\ 11.03 \\ 40.1 \\ 36.99 \end{pmatrix};$$

$$10) A = \begin{pmatrix} 2.89 & 3.95 & 0.96 & 1.98 \\ 2.97 & 4.98 & 2.99 & 4.92 \\ 5.98 & 7.91 & 0.91 & 4.91 \\ 2.90 & 5.00 & 2.97 & 6.99 \end{pmatrix}, \quad b = \begin{pmatrix} -2.9 \\ -5.989 \\ -7.92 \\ -7.99 \end{pmatrix};$$

$$11) A = \begin{pmatrix} 7.11 & 9.02 & 4.01 & 2.02 \\ 2.04 & -2.05 & 1.01 & 1.09 \\ 5.11 & 6.08 & 3.12 & 2.07 \\ 2.04 & 3.03 & 1.03 & 1.03 \end{pmatrix}, \quad b = \begin{pmatrix} 2.12 \\ 6.13 \\ 3.1 \\ 0 \end{pmatrix};$$

$$12) A = \begin{pmatrix} 6 & 4.98 & -1.97 & 3.96 \\ 9.07 & -0.99 & 3.96 & -0.99 \\ 2.99 & 3.91 & 2.01 & -1.96 \\ 2.94 & -8.93 & 0 & 1.92 \end{pmatrix}, \quad b = \begin{pmatrix} -3.95 \\ 12.94 \\ 0.93 \\ 10.96 \end{pmatrix};$$

$$13) A = \begin{pmatrix} 2.10 & -1.11 & -6.01 & 3.02 \\ 7 & -4 & 2 & -15 \\ 1.01 & -2.02 & -4.03 & 9.04 \\ 1.04 & -1.03 & 2.02 & -6.01 \end{pmatrix}, \quad b = \begin{pmatrix} -1.09 \\ -32 \\ 5.07 \\ -8.06 \end{pmatrix};$$

$$14) A = \begin{pmatrix} 2 & -1 & 2 & 0 \\ 2.01 & 1.02 & 4.02 & 8.01 \\ 2.99 & -1.98 & 1.97 & -1.96 \\ 1.04 & 3.03 & -6.02 & 2.01 \end{pmatrix}, \quad b = \begin{pmatrix} -1.1 \\ 3.05 \\ 8.05 \\ 4.1 \end{pmatrix};$$

$$15) A = \begin{pmatrix} 1.91 & -4.92 & 2.93 & 1.94 \\ 2.95 & -6.96 & 2.97 & -0.98 \\ 4.99 & -9 & 6.01 & 2.02 \\ 4.03 & -6.04 & 3.05 & 1.06 \end{pmatrix}, \quad b = \begin{pmatrix} 5.07 \\ -0.98 \\ 6.96 \\ 7.99 \end{pmatrix};$$

$$16) A = \begin{pmatrix} 1.01 & -1.07 & -4.08 & 9.07 \\ 1.05 & 2.02 & 0.0001 & -4.06 \\ 2.04 & -3.06 & 1.03 & 5.05 \\ 3.03 & -2.05 & -5.04 & 1.04 \end{pmatrix}, \quad b = \begin{pmatrix} 3.1 \\ -3.1 \\ -3.1 \\ 22.1 \end{pmatrix};$$

$$17) A = \begin{pmatrix} 3.99 & -3.02 & 0.91 & 4.89 \\ 1 & -1.96 & -1.97 & -2.94 \\ 2.98 & -0.94 & 1.91 & 0 \\ 1.91 & 2.93 & 1.94 & -7.89 \end{pmatrix}, \quad b = \begin{pmatrix} 7 \\ 2.99 \\ -0.97 \\ -6.98 \end{pmatrix};$$

$$18) A = \begin{pmatrix} 1.01 & 3.01 & 1.03 & -1.04 \\ 3.02 & 4.02 & -1.02 & 2.03 \\ 2.03 & 3.01 & 1.01 & -3.02 \\ 2.04 & -2.02 & 0 & 1.01 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 0 \\ -6 \\ -3 \end{pmatrix};$$

$$19) A = \begin{pmatrix} 1.1 & 1.03 & -6.06 & -4.07 \\ 12.99 & -0.91 & -5.90 & -3.92 \\ 1.94 & 2.95 & 8.97 & 1.97 \\ 2.99 & 2.01 & 2.99 & 7.98 \end{pmatrix}, \quad b = \begin{pmatrix} 6 \\ 2.1 \\ 6.01 \\ -7.01 \end{pmatrix};$$

$$20) A = \begin{pmatrix} 10 & 3 & -3 & -2 \\ 1.99 & -2.99 & 2.97 & 1.99 \\ 6.01 & 8.99 & -1.99 & -0.98 \\ 7.99 & 6.03 & 0.97 & 2.94 \end{pmatrix}, \quad b = \begin{pmatrix} 3 \\ -4 \\ 2.99 \\ -6.98 \end{pmatrix};$$

$$21) A = \begin{pmatrix} 1 & 2.001 & 5.02 & 9.0003 \\ 2.999 & 13 & 18.0001 & 30.00001 \\ 1.9999 & 3.998 & 10.99999 & 15.98765 \\ 0.9988 & 8.9988 & 8.999888 & 8.888999 \end{pmatrix}, \quad b = \begin{pmatrix} 78.999 \\ 262.999 \\ 145.987 \\ 91.969 \end{pmatrix};$$

$$22) A = \begin{pmatrix} 5 & 18.001 & 4.1 & 5.002 \\ 2.001 & 7 & 3.2 & 1.012 \\ 1.0001 & 3.0123 & 5.0123 & -2.022 \\ 1.0002 & 5.00112 & -9.0123 & 8.011 \end{pmatrix}, \quad b = \begin{pmatrix} 12.03 \\ 5 \\ 3 \\ 1.01 \end{pmatrix};$$

$$23) A = \begin{pmatrix} 2 & 2.99 & -0.998 & 1 \\ 7.99 & 11.98 & -8.995 & 7.99 \\ 3.98 & 5.9 & 2.987 & -1.98 \\ 1.98 & 2.94 & 8.987 & -6.99 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2.9 \\ 2.99 \\ 2.999 \end{pmatrix};$$

$$24) A = \begin{pmatrix} 4 & -2.987 & 2.01 & -1 \\ 3 & -1.987 & 1.012345 & -3 \\ 2 & -0.999 & 0 & -5 \\ 5 & -2.888 & 0.9876543 & -8 \end{pmatrix}, \quad b = \begin{pmatrix} 7.95 \\ 6.95 \\ 5.95 \\ 0.95 \end{pmatrix};$$

$$25) A = \begin{pmatrix} 1.99999 & -1 & 1 & -1 \\ 3.98765 & -2 & -2.00001 & 2.9999 \\ 1.998877 & -1 & 5.00001 & -5.98765 \\ 1.798798 & -1 & -3.000001 & 3.95678 \end{pmatrix}, \quad b = \begin{pmatrix} 3.001 \\ 2.013 \\ 0.98 \\ 5 \end{pmatrix};$$

$$26) A = \begin{pmatrix} 3 & -1 & -1 & -2 \\ 0.9 & 0.44 & 1.98 & 2.987 \\ 2 & 3.0001 & -1 & -1.001 \\ 1 & 2 & 3 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} -4 \\ 0.999 \\ -6.001 \\ -4 \end{pmatrix};$$

$$27) A = \begin{pmatrix} 2 & -3 & 2 & 1 \\ 3.001 & 2 & -0.9999 & 2.0123 \\ 2.012 & -0.999 & -1.9988 & -2.9999 \\ 0.999 & 1.987 & 2.9876 & -1.9998 \end{pmatrix}, \quad b = \begin{pmatrix} -8 \\ 3.99 \\ 7.99 \\ 0.99 \end{pmatrix};$$

$$28) A = \begin{pmatrix} 2.0123 & 1.000001 & 2.0001 & 3.0123 \\ 4.0111 & 3.0123 & 2.0123 & 1 \\ 3.0333 & 2.0123 & 1.0123 & 2 \\ 1.0123 & 20.444 & 3.0445 & 4.0123 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ -5 \\ 1 \\ 5 \end{pmatrix};$$

$$29) A = \begin{pmatrix} 0 & 0.999 & -2.99 & 3.98 \\ 0.99 & 0 & -1.999 & 2.999 \\ 3 & 2 & 0 & -4.999 \\ 4.12 & 3.002 & -5.001 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} -4.987 \\ -3.987 \\ 12.012 \\ 5.012 \end{pmatrix};$$

$$30) A = \begin{pmatrix} 0 & 0 & 0.99 & 1.999 & 3.12 \\ 0 & 1 & 2.001 & 3.123 & -0.01 \\ 1 & 1.999 & 3.001 & 0.002 & 0 \\ 0 & 1 & 1.002 & 1 & 1 \\ 1 & 1.003 & 1 & 1.0001 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 2.01 \\ -2.02 \\ 2.03 \\ 0.01 \\ 0 \end{pmatrix};$$

$$31) A = \begin{pmatrix} 3.0001 & 3.0012 & 3.01 & 2.0001 \\ 3.0002 & -1.0013 & -1.02 & 2.0002 \\ 3.0033 & -1.0145 & 3.03 & -1.003 \\ 2.0001 & -1.00002 & 3.00003 & 2.0004 \end{pmatrix}, \quad b = \begin{pmatrix} 6.001 \\ 6.002 \\ 6.003 \\ 4.004 \end{pmatrix};$$

$$32) A = \begin{pmatrix} 1.1 & 1.2 & 1.3 & 1.4 & 1.5 & 1.6 \\ 2.1 & 2.2 & 2.3 & 2.4 & 2.5 & 2.6 \\ 3.5 & 3.4 & 3.3 & 3.2 & 3.1 & 3.0 \\ 4.1 & 4.2 & 4.3 & 4.4 & 4.3 & 4.6 \\ 5.5 & 5.5 & 5.5 & 5.5 & 5.5 & 5.5 \\ 6.1 & 6.2 & 6.3 & 6.4 & 6.5 & 6.6 \end{pmatrix}, \quad b = \begin{pmatrix} 6.1 \\ 7.2 \\ 8.3 \\ 9.4 \\ 10.5 \\ 11.6 \end{pmatrix};$$

$$33) A = \begin{pmatrix} 1.351 & 0.201 & -0.623 & 0.440 \\ 0.201 & 1.351 & 0.123 & 0.985 \\ -0.623 & 2.111 & 1.351 & -1.002 \\ 0.440 & 0.334 & -0.154 & 1.351 \end{pmatrix}, \quad b = \begin{pmatrix} 1.009 \\ -1.008 \\ -1.607 \\ 2.006 \end{pmatrix}.$$

5. Предложите свой метод решения систем.

## ПРИЛОЖЕНИЕ

### Дополнения к главе I

#### Оценка погрешности в методе Ньютона

Выведем ещё одно неравенство для оценки погрешности приближения  $x_{n+1}$  (см. п. 4.3). Воспользуемся тождеством  $x_{n+1} = x_n + (x_{n+1} - x_n)$ . Применяя формулу Тейлора ([7], [8]), получим

$$f(x_{n+1}) = f(x_n + (x_{n+1} - x_n)) = f(x_n) + f'(x_n)(x_{n+1} - x_n) + \frac{1}{2}f''(\eta)(x_{n+1} - x_n)^2, \quad (*)$$

где  $\eta \in (x_n; x_{n+1})$ .

В силу определения приближения  $x_{n+1}$  (см. (43)) имеем

$$f(x_n) + f'(x_n)(x_{n+1} - x_n) = 0,$$

поэтому из (\*) найдём

$$|f(x_{n+1})| = \frac{1}{2} |f''(\eta)| (x_{n+1} - x_n)^2.$$

Обозначим  $M_2 = \sup_{x \in [a; b]} |f''(x)|$ , тогда имеем

$$|f(x_{n+1})| \leq \frac{M_2}{2} (x_{n+1} - x_n)^2.$$

На основании неравенства (46) получим

$$|\xi - x_{n+1}| \leq \frac{M_2}{2m_1} (x_{n+1} - x_n)^2. \quad (**)$$

Если процесс Ньютона сходится, то  $x_{n+1} - x_n \rightarrow 0$  при  $n \rightarrow \infty$ . Поэтому при  $n \geq N$  имеем

$$|\xi - x_{n+1}| \leq |x_{n+1} - x_n|,$$

т. е. совпавшие старшие разряды приближений  $x_n$  и  $x_{n+1}$  верны. Из (\*\*\*) получаем, что для достижения заданной точности  $\varepsilon$  процесс итераций надо продолжать до тех пор, пока для двух последовательных приближений не будет обеспечено выполнение неравенства

$$|x_{n+1} - x_n| \leq \sqrt{\frac{2m_1}{M_2}} \varepsilon. \quad (47)$$

### Решение кубического уравнения

**Первый способ.** *Применение формулы Кардано* (см. [3]). Формула Кардано для уравнения  $a_0x^3 + a_1x^2 + a_2x + a_3 = 0$ , где  $a_0, a_1, a_2$  и  $a_3$  — действительные числа ( $a_0 \neq 0$ ), относится к его приведённому виду  $t^3 + pt + q = 0$  ( $x = t - a_1/(3a_0)$ ). В этом случае имеем

$$\begin{aligned} t_1 &= u + v, \\ t_2 &= -\frac{u+v}{2} + \frac{u-v}{2}i\sqrt{3} = \zeta_1 u + \zeta_2 v, \\ t_3 &= -\frac{u+v}{2} - \frac{u-v}{2}i\sqrt{3} = \zeta_2 u + \zeta_1 v, \end{aligned}$$

где

$$\begin{aligned} u &= \sqrt[3]{-\frac{q}{2} + \sqrt{D}}, & v &= \sqrt[3]{-\frac{q}{2} - \sqrt{D}}, \\ D &= \left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2, & \zeta_{1,2} &= \frac{-1 \pm i\sqrt{3}}{2}. \end{aligned}$$

Решения  $x_j$  исходного кубического уравнения даются формулой  $x_j = t_j - a_1/(3a_0)$  ( $j = 1, 2, 3$ ).

В случае  $D < 0$  кубическое уравнение имеет три действительных решения. Если использовать приведённые выше формулы, то корни будут выражаться через комплексные

величины. Избежать этого можно так (см. также 2-й и 3-й способы): положим  $\varrho = \sqrt{-p^3/27}$ ,  $\cos \varphi = -q/(2\varrho)$ ; тогда решениями приведённого уравнения  $t^3 + pt + q = 0$  будут числа

$$\begin{aligned} t_1 &= 2\sqrt[3]{\varrho} \cos \frac{\varphi}{3}, \\ t_2 &= 2\sqrt[3]{\varrho} \cos \left( \frac{\varphi}{3} + \frac{2\pi}{3} \right), \\ t_3 &= 2\sqrt[3]{\varrho} \cos \left( \frac{\varphi}{3} + \frac{4\pi}{3} \right), \end{aligned}$$

от которых посредством замены  $x_j = t_j - a_1/(3a_0)$  ( $j = 1, 2, 3$ ) перейдём к корням заданного кубического уравнения.

**Второй способ.** *Тригонометрическое решение.*

Если  $D < 0$ , то  $p < 0$  и

$$t_1 = 2\sqrt{-\frac{p}{3}} \cos \frac{\varphi}{3}, \quad t_{2,3} = -2\sqrt{-\frac{p}{3}} \cos \left( \frac{\varphi}{3} \pm \frac{\pi}{3} \right),$$

где значения тригонометрических функций вычисляются по значению  $\cos \varphi = -q/(2\sqrt{-p^3/27})$ .

Если  $D \geq 0$  и  $p < 0$ , то

$$t_1 = -2\sqrt{-\frac{p}{3}} \frac{1}{\sin 2\varphi}, \quad t_{2,3} = \sqrt{-\frac{p}{3}} \left( \frac{1}{\sin 2\varphi} \pm i\sqrt{3} \operatorname{ctg} 2\varphi \right),$$

где значения тригонометрических функций вычисляются по значению  $\operatorname{tg} \varphi = \sqrt[3]{\operatorname{tg}(\psi/2)}$ ,  $\sin \psi = 2\sqrt{-p^3/27}/q$ ,  $|\varphi| \leq \pi/4$ ,  $|\psi| \leq \pi/2$ .

Если  $D \geq 0$  и  $p > 0$ , то

$$t_1 = -2\sqrt{\frac{p}{3}} \operatorname{ctg} 2\varphi, \quad t_{2,3} = \sqrt{\frac{p}{3}} \left( \operatorname{ctg} 2\varphi \pm i\sqrt{3} \frac{1}{\sin 2\varphi} \right),$$

где значения тригонометрических функций вычисляются по значению  $\operatorname{tg} \varphi = \sqrt[3]{\operatorname{tg}(\psi/2)}$ ,  $\operatorname{tg} \psi = 2\sqrt{p^3/27}/q$ ,  $|\varphi| \leq \pi/4$ ,  $|\psi| \leq \pi/2$ .

Во всех трёх случаях берётся действительное значение кубического корня.

**Третий способ.** *Применение гиперболических функций.* В приведённом уравнении  $t^3 + pt + q = 0$  положим<sup>1</sup>  $r = \operatorname{sgn} q \cdot \sqrt{|p|/3}$ . Тогда вспомогательная величина  $\varphi$  и при её помощи корни  $t_1, t_2$  и  $t_3$  определяются в зависимости от знаков  $p$  и  $D = (p/3)^3 + (q/2)^2$  из следующей ниже таблицы:

	$p < 0, D \leq 0$	$p < 0, D > 0$	$p > 0, D > 0$
$\varphi$	$\cos \varphi = \frac{q}{2r^3}$	$\operatorname{ch} \varphi = \frac{q}{2r^3}$	$\operatorname{sh} \varphi = \frac{q}{2r^3}$
$t_1$	$-2r \cos \frac{\varphi}{3}$	$-2r \operatorname{ch} \frac{\varphi}{3}$	$-2r \operatorname{sh} \frac{\varphi}{3}$
$t_2$	$-2r \cos \left( \frac{\varphi}{3} + \frac{2\pi}{3} \right)$	$r \operatorname{ch} \frac{\varphi}{3} + i\sqrt{3}r \operatorname{sh} \frac{\varphi}{3}$	$r \operatorname{sh} \frac{\varphi}{3} + i\sqrt{3}r \operatorname{ch} \frac{\varphi}{3}$
$t_3$	$-2r \cos \left( \frac{\varphi}{3} + \frac{4\pi}{3} \right)$	$r \operatorname{ch} \frac{\varphi}{3} - i\sqrt{3}r \operatorname{sh} \frac{\varphi}{3}$	$r \operatorname{sh} \frac{\varphi}{3} - i\sqrt{3}r \operatorname{ch} \frac{\varphi}{3}$

Для гиперболических функций  $\operatorname{sh} x = (e^x - e^{-x})/2$  и  $\operatorname{ch} x = (e^x + e^{-x})/2$  существуют обратные функции — *арекосинус* и *арекосинус* [3]:

$$y = \operatorname{Arsh} x = \ln(x + \sqrt{x^2 + 1}),$$

$$y = \operatorname{Arch} x = \ln(x - \sqrt{x^2 - 1}) \quad \text{для } x \geq 1 \text{ и } -\infty < y \leq 0,$$

$$y = \operatorname{Arch} x = \ln(x + \sqrt{x^2 - 1}) \quad \text{для } x \geq 1 \text{ и } 0 \leq y < +\infty.$$

### Решение уравнения четвёртой степени

Уравнение  $a_0x^4 + a_1x^3 + a_2x^2 + a_3x + a_4 = 0$ , где  $a_j \in \mathbf{R}$  ( $j = 0, 1, 2, 3, 4$ ;  $a_0 \neq 0$ ) посредством замены  $x = t - a_1/(4a_0)$

<sup>1</sup>Напомним определение функции сигнум:

$$\operatorname{sgn} x = \begin{cases} 1, & \text{если } x \in \mathbf{R}_+; \\ 0, & \text{если } x = 0; \\ -1, & \text{если } x \in \mathbf{R}_-. \end{cases}$$

Обозначается также  $\operatorname{sign} x$ .

преобразуем в уравнение  $t^4 + pt^2 + qt + r = 0$ . Здесь  $p, q$  и  $r$  — рациональные функции коэффициентов  $a_j$  ( $j = 0, 1, 2, 3, 4$ ).

Вид решения этого уравнения определяется видом решения его кубической резольвенты  $z^3 + 2pz^2 + (p^2 - 4r)z - q^2 = 0$ . Заметим, что произведение корней  $z_1, z_2$  и  $z_3$ , по теореме Виета равное  $q^2$ , должно быть всегда положительным ( $q \neq 0$ ).

Если область допустимых значений неизвестного есть множество  $C$  комплексных чисел, то имеет место следующее:

Уравнение четвёртой степени	Кубическая резольвента
Четыре действительных корня	Все корни положительные
Две пары комплексно сопряжённых корней	Один из корней положительный, два отрицательных
Два действительных корня и два комплексно сопряжённых	Один действительный корень и два комплексно сопряжённых

Если  $z_1, z_2, z_3$  — корни кубической резольвенты, то

$$t_1 = \frac{1}{2}(\sqrt{z_1} + \sqrt{z_2} + \sqrt{z_3}), \quad t_3 = \frac{1}{2}(-\sqrt{z_1} + \sqrt{z_2} - \sqrt{z_3}),$$

$$t_2 = \frac{1}{2}(\sqrt{z_1} - \sqrt{z_2} - \sqrt{z_3}), \quad t_4 = \frac{1}{2}(-\sqrt{z_1} - \sqrt{z_2} + \sqrt{z_3})$$

суть решения уравнения  $t^4 + pt^2 + qt + r = 0$  (при этом знаки перед радикалами  $\sqrt{z_1}, \sqrt{z_2}, \sqrt{z_3}$  выбирают так, чтобы выполнялось равенство  $\sqrt{z_1}\sqrt{z_2}\sqrt{z_3} = -q$ ). Далее посредством формулы  $x_j = t_j - a_1/(4a_0)$  ( $j = 1, 2, 3, 4$ ) находят решения исходного уравнения четвёртой степени.

В частных случаях, уравнение  $a_0x^4 + a_2x^2 + a_4 = 0$  является *биквадратным* (подстановка  $x^2 = t$ ), а уравнение  $a_0x^4 + a_1x^3 + a_2x^2 + \lambda a_1x + \lambda^2 a_0 = 0$  — *возвратным* ( $x + \lambda/x = t$ ). Если  $a_0 = 1$  и коэффициенты (приведённого) уравнения  $x^4 + px^3 + qx^2 + rx + s = 0$  удовлетворяют равенству  $p^3 + 8r = 4pq$ , то решениями этого уравнения будут корни уравнения  $(x^2 + px/2)^2 + (q - p^2/4)(x^2 + px/2) + s = 0$ . Иногда приведённое уравнение удаётся решить *методом неопределённых коэффициентов*:  $x^4 + px^3 + qx^2 + rx + s = (x^2 + \alpha x + \beta)(x^2 + \gamma x + \delta)$ , а общее — *разложить на мно-*

жители посредством теоремы Виета:

$$a_0x^4 + a_1x^3 + a_2x^2 + a_3x + a_4 = a_0(x - x_1)(x - x_2)(x^2 + \alpha x + \beta).$$

## Дополнения к главе II

### Решение задачи 2 из упр. 4

*Решение.* Искомая формула должна иметь вид

$$\int_a^b f(x) dx = c_1 f(x_1) + c_2 f(x_2) + R.$$

Так как по условию задачи остаточный член  $R$  для функции  $f(x) = a_0x^3 + a_1x^2 + a_2x + a_3$  равен нулю при любых  $a_i$  ( $i = 0, 1, 2, 3$ ), то, применив к интегралу формулу Ньютона — Лейбница, получим следующее соотношение:

$$\begin{aligned} & \frac{b^4 - a^4}{4} a_0 + \frac{b^3 - a^3}{3} a_1 + \frac{b^2 - a^2}{2} a_2 + (b - a) a_3 = \\ & = c_1 (a_0 x_1^3 + a_1 x_1^2 + a_2 x_1 + a_3) + c_2 (a_0 x_2^3 + a_1 x_2^2 + a_2 x_2 + a_3) = \\ & = (c_1 x_1^3 + c_2 x_2^3) a_0 + (c_1 x_1^2 + c_2 x_2^2) a_1 + (c_1 x_1 + c_2 x_2) a_2 + (c_1 + c_2) a_3. \end{aligned}$$

Отсюда, приравнявая коэффициенты при одинаковых  $a_i$  ( $i = 0, 1, 2, 3$ ) слева и справа от знака равенства, придём к системе

$$\begin{aligned} \frac{b^4 - a^4}{4} &= c_1 x_1^3 + c_2 x_2^3, \\ \frac{b^3 - a^3}{3} &= c_1 x_1^2 + c_2 x_2^2, \\ \frac{b^2 - a^2}{2} &= c_1 x_1 + c_2 x_2, \\ b - a &= c_1 + c_2, \end{aligned}$$

решениями которой являются следующие числа:

$$c_1 = \frac{b-a}{2}, \quad x_1 = \frac{b+a}{2} - \frac{\sqrt{3}}{3} \cdot \frac{b-a}{2},$$

$$c_2 = \frac{b-a}{2}, \quad x_2 = \frac{b+a}{2} + \frac{\sqrt{3}}{3} \cdot \frac{b-a}{2}.$$

Тогда искомая формула примет вид

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

$$= \frac{b-a}{2} \left( f\left(\frac{b+a}{2} - \frac{\sqrt{3}}{3} \cdot \frac{b-a}{2}\right) + f\left(\frac{b+a}{2} + \frac{\sqrt{3}}{3} \cdot \frac{b-a}{2}\right) \right).$$

Итак, при  $n = 2$  узлы  $t_i^{(2)}$  ( $i = 1, 2$ ) и коэффициенты (веса)  $C_i^{(2)}$  таковы:

$$t_2^{(2)} = -t_1^{(2)} = 0.57735\ 02691\ 89625\ 764, \quad C_1^{(2)} = C_2^{(2)} = 1.$$

## Дополнения к главе III

### Правило треугольника

Найдём *определитель третьего порядка*

$$\det A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} -$$

$$- a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

по правилу треугольника. Вот его формулировка: составьте сумму произведения элементов главной диагонали и произведений элементов, расположенных в вершинах равнобедренных треугольников с основаниями, параллельными главной диагонали (рис. А); из этой суммы вычтите

сумму произведения элементов побочной диагонали и произведений элементов, расположенных в вершинах равнобедренных треугольников с основаниями, параллельными побочной диагонали (рис. Б):

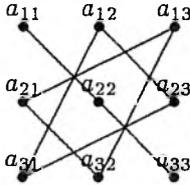


Рис. А

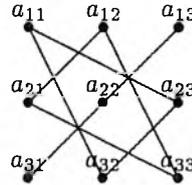


Рис. Б

### Решение задачи 10 из упр. 4

Представим систему с трёхдиагональной матрицей (1) коэффициентов в следующем виде:

$$\begin{aligned}
 b_1 x_1 + c_1 x_2 &= d_1, \\
 a_2 x_1 + b_2 x_2 + c_2 x_3 &= d_2, \\
 &\dots\dots\dots \\
 a_{n-1} x_{n-2} + b_{n-1} x_{n-1} + c_{n-1} x_n &= d_{n-1}, \\
 a_n x_{n-1} + b_n x_n &= d_n.
 \end{aligned} \tag{1}$$

Решим её: из первого уравнения системы (1) найдём  $x_1 = p_2 x_2 + q_2$ , где  $p_2 = -c_1/b_1$ ,  $q_2 = d_1/b_1$ ; из второго уравнения —  $x_2 = p_3 x_3 + q_3$ , где  $p_3 = -c_2/(b_2 + a_2 p_2)$ ,  $q_3 = (d_2 - a_2 q_2)/(b_2 + a_2 p_2)$  и т. д.

Если положить  $a_1 = 0$ ,  $c_n = 0$ ,  $p_1 = 0$  и  $q_1 = 0$ , то формулы для вспомогательных коэффициентов примут вид

$$p_{i+1} = \frac{-c_i}{b_i + a_i p_i}, \quad q_{i+1} = \frac{d_i - a_i q_i}{b_i + a_i p_i} \quad (i = 1, 2, \dots, n).$$

Из последнего уравнения системы (1) определим  $x_n = q_{n+1}$ , и на этом закончим прямой ход прогонки.

Значения неизвестных  $x_{n-1}, x_{n-2}, \dots, x_1$  найдём обратным ходом прогонки:

$$x_i = p_{i+1}x_{i+1} + q_{i+1} \quad (i = n, n-1, \dots, 1).$$

Решение системы запишем в вектор  $\mathbf{x} = (x_1, x_2, \dots, x_n, x_{n+1})$ , где элемент  $x_{n+1} = 0$  является вспомогательным.

Приведём текст алгоритма (см. также [17]):

{ \_\_\_\_\_ }

**алг ПРОГОНКА** (двточн массив  $a[lim], b[lim], c[lim],$   
 $d[lim]$ , цел  $lim, n$ , двточн массив  $x[lim]$ )  
**знач**  $a, b, c, d, lim, n$   
**рез**  $x$   
**двточн массив**  $p[lim], q[lim]$   
**цел**  $i$   
**двточн**  $t$   
**начало**  
 $p_1 \leftarrow 0 \quad q_1 \leftarrow 0$   
**для**  $i$  **от** 1 **до**  $n$  **шаг** 1 **выполнять**  
 $t \leftarrow 1/(b_i + a_i p_i)$   
 $p_{i+1} \leftarrow -c_i t$   
 $q_{i+1} \leftarrow (d_i - a_i q_i) t$   
**конец-цикл** { ео прямого хода прогонки }  
 $x_{n+1} \leftarrow 0$   
**для**  $i$  **от**  $n$  **до** 1 **шаг** -1 **выполнять**  
 $x_i \leftarrow p_{i+1} x_{i+1} + q_{i+1}$   
**конец-цикл** { ео обратного хода прогонки }  
**возвратиться**  
**конец** { ео алг Прогонка }  
 { \_\_\_\_\_ }

Требуют пояснения метAPERЕМЕННЫЕ  $lim$  — предельная длина (мощность) массива и  $n$  — текущая длина массива. Нижняя граница индекса равна (по умолчанию) единице.

## Предметный указатель

- алгоритм**  
 Бисекции 31  
 Вегстейна 65  
 Метод Гаусса 161  
 Метод Гаусса - Зейделя 178  
 Метод Гаусса - Лежандра 125  
 Метод итераций 49  
 Метод Ньютона 75  
 Метод средних прямоугольников 121  
 Метод хорд 62  
 Отделение корней 19  
 Парабола 123  
 Прогонка 196  
 Уточнение корней 30  
**арекосинус** 191  
**аресинус** 191  
**афоризм** 78  
**Вектор** 131  
**Геометрическая интерпретация**  
 единственности корня 16  
 метода бисекций 26  
 метода Гаусса - Зейделя 171  
 метода итераций 36, 42  
 метода касательных 66, 67  
 метода модифицированного касательных 73  
 метода парабол 113  
 метода прямоугольников 88, 89, 95  
 метода трапеций 109  
 метода хорд 54, 55  
 отделения корней 14, 16  
 сумм Дарбу 95, 96  
 существования корня 14  
**Данные внешние** 64  
**детерминант** 137  
**диагональ матрицы** 132  
**длина**  
 пути 84  
 эллипса 126, 127  
**доказательство** 25  
**дополнение алгебраическое** 139  
**Заряд электрический** 87  
**Интеграл, определение** 90  
**интегрируемость функций** 94–102  
 классы функций 98  
**итерация** 34, 35  
**Квадратура** 105  
**компьютерный  $\epsilon$**  52  
**константа Липшица** 39, 41  
**корень**  
 границы 10, 12

приближённое значение  
 22, 23  
 определение 9  
 отделение 13–20  
 этапы поиска 9, 10  
 косинус гиперболический  
 22, 190  
 коэффициент при неиз-  
 вестном 144  
**Лемма**  
 о пределах в неравен-  
 ствах 26  
 о свойствах сумм Дарбу  
 96, 97  
**Масса стержня** 87  
**матрица** 130–142  
 верхняя треугольная 133  
 действия над матрицей  
 133–136  
 диагональная 132  
 единичная 132, 136  
 квадратная 131  
 кососимметрическая 133  
 ленточная 132  
 неособенная 138  
 нулевая 132  
 особенная 138  
 перестановочные 136  
 противоположная 135  
 прямоугольная 130  
 равные 133  
 симметрическая 133  
 - столбец 131  
 - строка 131  
 сцепленные 134  
 транспонированная 133

**метод**  
 бисекций, дихотомии  
 25–32  
 Вегстейна 65  
 вычисления интеграла  
 102–105  
 Гаусса 153–168  
 Гаусса - Зейделя 170–182  
 итеративный 7, 145  
 касательных, Ньютона  
 65–78  
 касательных модифици-  
 рованный 73, 74  
 неопределённых коэффи-  
 циентов 192, 193  
 парабол, Симпсона 112,  
 115  
 прогонки 195, 196  
 простых итераций 34–52,  
 172  
 прямой 7, 145  
 равномерного перебора 19  
 сходящийся 38, 39  
 точный вычисления инте-  
 грала 91–94  
 хорд 54–71  
**минор** 139  
**Невязка**  
 система 167, 168  
 уравнение 48  
**Область допустимых зна-  
 чений** 13, 145  
**обсуждение метода**  
 бисекций 32  
 Гаусса 168  
 Гаусса - Зейделя 181

- итераций 47
- касательных 74
- методов решения уравнений 77–79
- хорд 62
- численного интегрирования 126, 117
- определитель 137–142
- свойства 139–140
- оценка
  - априорная 6, 22, 118
  - апостериорная 6, 23, 119
  - погрешности приближённого решения 22–23
  - практическая погрешности квадратурных формул 118, 119
- оценка погрешности в методе
  - Гаусса - Зейделя 175
  - Гаусса - Лежандра 125
  - итераций 42, 43
  - касательных 71, 72
  - парабол 115, 119
  - приближённого вычисления интегралов 119, 103
  - прямоугольников 107, 119
  - трапеций 111, 119
  - хорд 60, 61
- оценка качества решения системы 167, 168
- уравнения 32, 51, 64, 77
- оценка остатка ряда 103
- “отпечатки пальцев” 49
- Правило**
  - Декарта 10
  - кольца 11
  - Крамера 8, 146–152
  - Рунге 118, 119, 121
  - Саррюса 138
  - треугольника 194, 195
- переменная-флажок в методе
  - Гаусса - Зейделя 178
  - итераций 48
  - касательных 75
  - отделения корня 19
  - хорд 63
- площадь криволинейной трапеции 87
- предел интегральных сумм 90
- преобразование эквивалентное 153
- приближение начальное в методе
  - Гаусса - Зейделя 174
  - итераций 41, 35, 36, 39
  - касательных 68, 69
  - хорд 57, 54, 55
- приведение системы 176–177
- уравнения 45–46
- погрешность
  - метода 42, 43, 103
  - неустраняемая 42
  - округления 42, 103, 168
- порядок точности квадратурных формул 118
- последовательность итеративная 35
- сходящаяся 35, 36, 38

- фундаментальная 35, 40
- произведение
  - векторов 135
  - матриц 133–136
  - матрицы на вектор 134
- Работа** 85
- разбиение 89
- разложение определителя 141, 142
- резольвента 192
- ряд знакопеременный 103
- Самоисправляемость** 47, 62, 74, 182
- сигнум 192, 15, 34
- синус гиперболический 22, 191
- система линейных алгебраических уравнений 144
  - неоднородная 144
  - неопределённая 145
  - несовместная 145
  - однородная 144
  - определённая 145
  - решение 145
  - совместная 145
  - плохо обусловленная 147, 160, 167, 181
  - эквивалентная 145, 153, 158
- сходимость метода
  - Гаусса - Зейделя 173
  - итераций 39
  - касательных 69
  - хорд 58
- сумма
  - Дарбу 94–98
  - интегральная 89, 105
  - матриц 133
  - риманова 90
- схема единственного деления 153–159
- схема функциональная в методе
  - бисекций 30
  - Гаусса 160
  - Гаусса - Зейделя 178
  - итераций 48
  - касательных 48
  - отделения корней 19
  - средних прямоугольников 120
  - хорд 48
- Тангенс гиперболический** 22
- теорема
  - Больцано - Коши 14
  - Вейерштрасса 59, 70
  - вторая Вейерштрасса 23
  - Лагранжа о конечных приращениях 23, 40, 42, 60
  - об интегрируемости функций 98, 102
  - о поведении последовательности  $(x_n)$  52
  - основная алгебры 10
  - об единственности корня 15
  - о сходимости метода Гаусса - Зейделя 173, 174
  - о сходимости метода итераций 39
  - о сходимости метода ка-

- касательных 69, 70
- о сходимости метода хорд 58, 59
- оценка погрешности приближённого решения 23
- правило кольца 11
- разложения определителя 141
- точка
  - неподвижная 55–57
  - подвижная 55, 56
- точность расчётов 176
- Уравнение**
  - алгебраическое 10
  - биквадратное 192
  - возвратное 192
  - конечное 5
  - кубическое 189
  - линейное 144
  - эквивалентное 13, 35, 45, 48, 74
  - матричное 145
  - четвёртой степени 191
- Формула**
  - Кардано 7, 189
  - Крамера 146–152
  - Ньютона - Лейбница 91, 92, 93, 109
  - рекуррентная 34, 36, 56, 67, 73, 172, 173
- формула метода
  - Гаусса - Зейделя 173
  - Гаусса - Лежандра 124
  - итераций 36
  - касательных 67
  - модифицированного касательных 73
  - парабол 113, 114
  - простых итераций 172
  - прямоугольников 106
  - трапеций 111
  - хорд 56
- функция первообразная 92
- Ход**
  - прямой 155–158, 195
  - обратный 155, 196
- Число делений начальное в методе**
  - парабол 115
  - прямоугольников 107
  - трапеций 112
- число шагов (операций) в методе
  - бисекций 29
  - Гаусса 9, 159, 168
  - Гаусса - Зейделя 174
  - Крамера 9, 152
  - подсчёта определителя 142
- член остаточный формулы
  - парабол 115
  - прямоугольников 107
  - трапеций 111
- член свободный 144
- Эффект побочный, волновой** 120, 166, 167
- элемент**
  - ведущий 154, 155, 158, 168
  - главный 159, 160, 168, 169
  - диагональный 132, 169

## СПИСОК ЛИТЕРАТУРЫ

1. *Бабенко К. И.* Основы численного анализа. М.: Лаборатория Базовых Знаний, 2001.
2. *Батвалов Н. С., Жидков Н. П., Кобельков Г. М.* Численные методы. М.: Наука, 1987.
3. *Бронштейн И. Н., Семендяев К. А.* Справочник по математике. М.: Наука, 1986.
4. *Вержбицкий В. М.* Основы численных методов. М.: Наука, 2002.
5. *Волков Е. А.* Численные методы. М.: Наука, 1987.
6. *Выгодский М. Я.* Справочник по высшей математике. М.: Наука, 1977.
7. *Демидович Б. П., Марон И. А.* Основы вычислительной математики. М.: Наука, 1970.
8. *Дорот В. Л., Троицкий В. А., Шелест В. Д.* Элементы вычислительной математики. Л.: ЛПИ, 1979.
9. *Калиткин Н. Н.* Численные методы. М.: Наука, 1978.
10. *Казанер Д., Моулер К., Нэш С.* Численные методы и программное обеспечение. М.: Мир, 1998.
11. *Крылов В. И.* Приближённое вычисление определённых интегралов. М.: Наука, 1967.
12. *Тихонов А. Н., Костомаров Д. П.* Рассказы о прикладной математике. М.: Наука, 1979.
13. *Троицкий В. А., Иванова И. М., Старостин И. А., Шелест В. Д.* Инженерные расчёты на ЭВМ. М.: Наука, 1979.
14. *Фаддеев Д. К., Фаддеева В. Н.* Вычислительные методы линейной алгебры. М.-Л.: Физматгиз, 1963.
15. *Фаддеев Д. К., Никулин М. С., Соколовский И. Ф.* Элементы высшей математики для школьников. М.: Наука, 1987.
16. *Феликс Люсьен.* Элементарная математика в современном изложении. М.: Просвещение, 1967.
17. *Шелест В. Д.* Программирование. Структурный подход. Алгоритмы. Turbo Pascal. Borland C++. Современный Fortran. СПб.: ВHV, 2002.

*Шелест Вячеслав Дмитриевич  
Житомирский Михаил Сергеевич*

**Начала  
вычислительной  
математики**

Введение в численный эксперимент

Компьютерная вёрстка: *В. Д. Шелест, М. А. Смирнов*

Лицензия ЛР № 020593 от 07.08.97

Налоговая льгота — Общероссийский классификатор продукции  
ОК 005-93, т. 2; 95 3005 — учебная литература

---

Подписано в печать *09.09.2005*. Формат 60×84/16. Печать офсетная.  
Усл. печ. л. 12.75. Уч.-изд. л. 12.75. Тираж 200. Заказ *433*.

---

Отпечатано с готового оригинал-макета, предоставленного авторами,  
в типографии Издательства Политехнического университета.  
195251, Санкт-Петербург, Политехническая ул., 29.

