

УДК 517.518.12  
ББК 22.19  
Г 54

519.6(075)

Г 610

Рецензенты: доктор техн. наук, профессор,  
академик МАИ С. Д. Пашкеев  
и кафедра цифровой вычислительной техники и информатики  
С-Пб ГУТ им. проф. М. А. Бонч-Бруевича

Гловацкая А. П.

Г 34 Методы и алгоритмы вычислительной математики. Учеб.  
пособие для вузов. – М.: Радио и связь, 1999. – 408 с.: ил.  
ISBN 5-256-01458-7.

Излагаются основные численные методы решения широкого круга задач,  
возникающих в инженерной практике. Пособие составлено в соответствии с  
программами курсов, изучаемых студентами инженерно - технических спе-  
циальностей вузов.

Для облегчения понимания логической структуры рассматриваемых ме-  
тодов и их использования в учебном пособии приводится большое количе-  
ство задач и схем алгоритмов. Методы и алгоритмы иллюстрируются при-  
мерами численных расчетов.

Для студентов вузов, обучающихся по специальностям «Сети связи и  
системы коммуникации», «Многоканальные телекоммуникационные систе-  
мы», «Вычислительные машины, комплексы, системы и сети», «Программ-  
ное обеспечение вычислительной техники и автоматизированных систем»,  
«Прикладная математика», «Радиосвязь, радиовещание и телевидение»,  
«Автоматизация технологических процессов».

ББК 22.19

Учебное издание

Гловацкая Алеатина Петровна

МЕТОДЫ И АЛГОРИТМЫ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ

Учебное пособие

Редактор И. Н. Сулова

ИБ № 2883

ЛР № 010164 от 29.01.97

Сдано в набор 10.04.99

Формат 60×90/16

Усл. печ. л. 25,5

Изд. № 24180

Бумага газетная

Усл.-кр. отт. 26,0

Зак. № 55

Гарнитура Arial

Уч.-изд. л. 25,28

С-015

Подписано в печать 31.05.99

Печать офсетная

Тираж 2000 экз.

Издательство «Радио и связь». 103473 Москва, 2-й Щемилловский пер., 4/5  
Типография издательства «Радио и связь». 103473 Москва, 2-й Щемилловский пер., 4/5

ISBN 5-256-01458-7

© Гловацкая А. П., 1999

Научно-техническая  
Библиотека ВолгГТУ

## Предисловие

Одним из основных способов решения задач, возникающих в самых различных областях, от научных исследований до получения технологических решений, является математическое моделирование процессов и явлений. В свою очередь вычислительные эксперименты при математическом моделировании невозможны без использования методов вычислительной математики, таких, как методы аппроксимации, интегрирования, оптимизации и решения различных уравнений: алгебраических, трансцендентных, дифференциальных и интегральных.

Численные решения прикладных задач всегда интересовали математиков. Недаром самые известные численные методы связаны с именами великих математиков прошлого: Ньютона, Гаусса, Эйлера, Лобачевского, Чебышева и других.

Эффективное использование пакетов программ, предназначенных для решения вычислительных задач, возможно лишь при хорошем знании математического аппарата, составляющего основу прикладных методов.

Данное пособие объединяет изложение теоретической основы вычислительной процедуры с четким алгоритмом ее компьютерной реализации. Как правило, приводятся примеры решения конкретной задачи, иллюстрирующие изложенные алгоритмы.

Предлагаемое учебное пособие предназначено прежде всего студентам и инженерно-техническим работникам, не являющимся специалистами в области вычислительной математики, но пожелавшим решать прикладные задачи с помощью компьютера.

Для более углубленного изучения отдельных разделов вычислительной математики автор предлагает обращаться к специальной литературе, ссылки на которую приведены в тексте пособия. Номер в квадратных скобках отсылает к списку литературы, приведенному в конце книги.

Пособие разделено на главы и пункты. Нумерация таблиц и рисунков самостоятельная в каждой главе.

Автор выражает благодарность профессору С.Д. Пашкееву и коллегам за ценные замечания.

Автор признателен за помощь при оформлении пособия студентам С. Аксенову, В.Каширину, В. Сандомирскому и М. Чернявскому.

## 1. СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ

Системы линейных уравнений появляются во многих областях прикладной математики. Можно утверждать, что решение линейных систем является одной из самых распространенных задач вычислительной математики. В некоторых случаях эти системы уравнений непосредственно составляют ту задачу, которую необходимо решать, в других случаях задача сводится к такой системе.

Методы решения систем линейных уравнений можно разделить на три класса:

- точные методы, представляющие собой алгоритмы, позволяющие находить точные решения за конечное число шагов;
- итерационные методы, результат применения которых - решение с заданной точностью (при этом сходящийся бесконечный процесс ограничивается конечным числом итераций);
- вероятностные.

Классы задач, для решения которых применяются эти методы, можно условно назвать, соответственно, классами задач с малым, средним и большим числом неизвестных.

Заметим, что результаты даже точных методов являются приближенными из-за неизбежных округлений исходных данных и промежуточных вычислений. При использовании итерационных методов помимо погрешности округления возникает и погрешность метода.

### 1.1. Метод Гаусса

Наиболее известным из точных методов решения систем линейных уравнений является метод исключения Гаусса.

Большинство распространенных точных методов можно рассматривать как варианты метода Гаусса, отличающиеся между собой лишь некоторыми деталями.

В общем случае система имеет вид:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = c_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = c_2, \\ \dots \dots \dots \dots \dots \dots \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = c_n. \end{cases} \quad (1)$$

Процесс решения системы линейных уравнений сводится к построению эквивалентной системы, имеющей треугольную матрицу (прямой ход) с последующим нахождением решения (обратный ход).

Отличие определителя системы от нуля - необходимое и достаточное условие существования единственного решения системы (1).

В случае равенства нулю определителя система либо не имеет решения, либо имеет их бесчисленное множество.

Если определитель системы близок к нулю, говорят, что система плохо обусловлена. В этом случае найти численное решение системы трудно, а точность его весьма сомнительна.

Ведущим элементом на каждой итерации ( $k$ ) назовем элемент  $a_{kk} \neq 0$ , а строку, содержащую  $a_{kk}$ , назовем ведущей строкой.

Необходимым и достаточным условием приведения матрицы к треугольному виду является отличие от нуля всех ведущих элементов ( $a_{ii}, i = \overline{1, n}$ ).

Наибольшая точность решения достигается тогда, когда ведущий элемент строки имеет наибольшее значение. Поэтому строку с нулевым или малым ведущим элементом надо заменить на ту из стоящих под ней строк, в которой в том же столбце стоит элемент, имеющий наибольшее значение.

Процесс исключения неизвестной  $x_k$  из всех уравнений  $i > k$  называется прямым ходом метода Гаусса.

Формулы пересчета коэффициентов системы на  $k$ -м шаге имеют вид:

$$\tilde{a}_{ij}^{(k-1)} = \frac{a_{ij}^{(k-1)}}{a_{kk}^{(k-1)}}, \quad (2)$$

$$a_{ij}^{(k)} = a_{ij}^{(k-1)} - a_{ik}^{(k-1)}\tilde{a}_{kj}^{(k-1)}, \quad (3)$$

$$i = k+1, \dots, n; \quad j = k, \dots, n+1.$$

( $i$  - номер строки,  $j$  - номер столбца элемента исходной матрицы).

Определение неизвестных по формулам

$$x_i = \tilde{a}_{i, n+1}^{(i-1)} - \sum_{j=i+1}^n \tilde{a}_{ij}^{(i-1)}x_j, \quad i = n, n-1, \dots, 1 \quad (4)$$

называется обратным ходом метода Гаусса.

При условии, что  $a_{11} \neq 0$ , разделим первое уравнение на  $a_{11}$ . Затем вычтем из каждого  $i$ -го уравнения ( $i \geq 2$ ) полученное после деления первое, умноженное на  $a_{i1}$ .

### Структурная схема алгоритма метода Гаусса

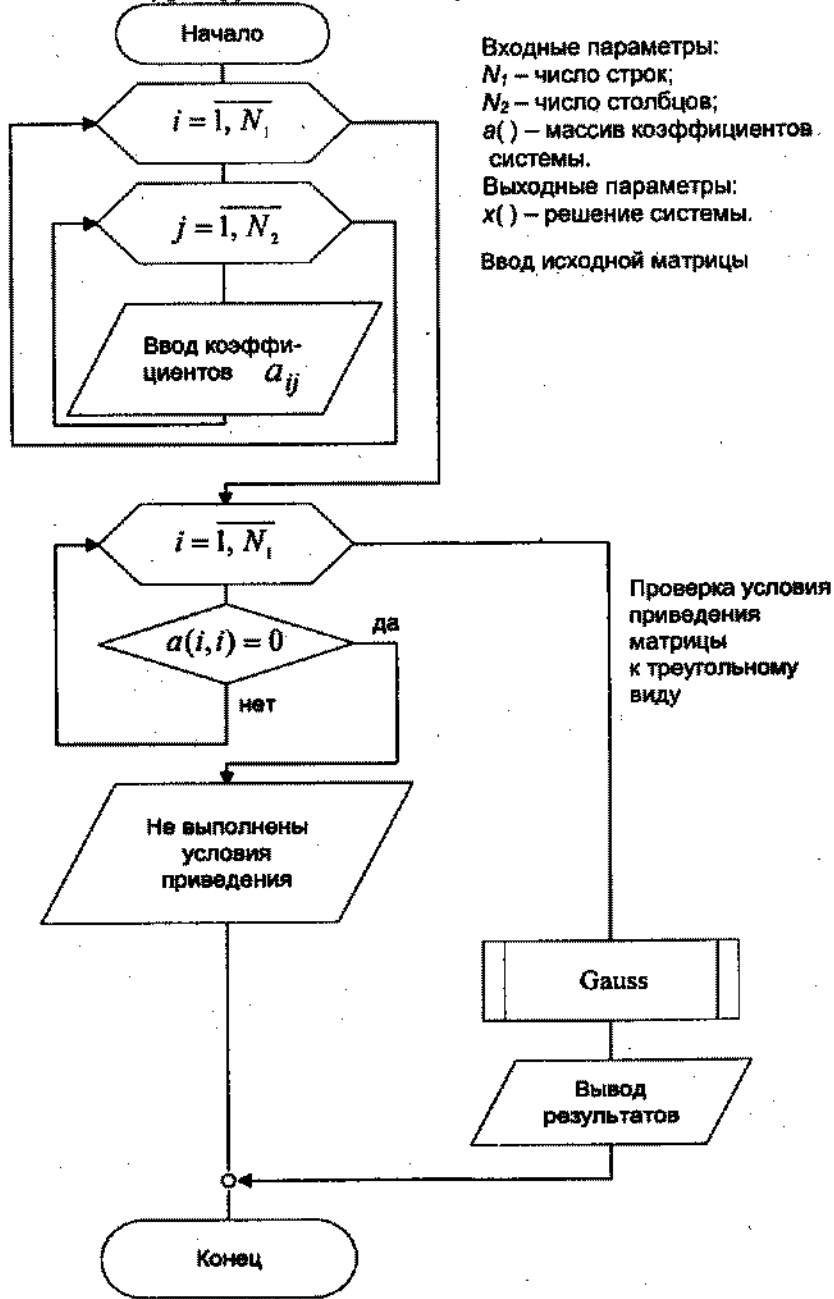


Рис. 1.1

### Процедура приведения матрицы к треугольному виду

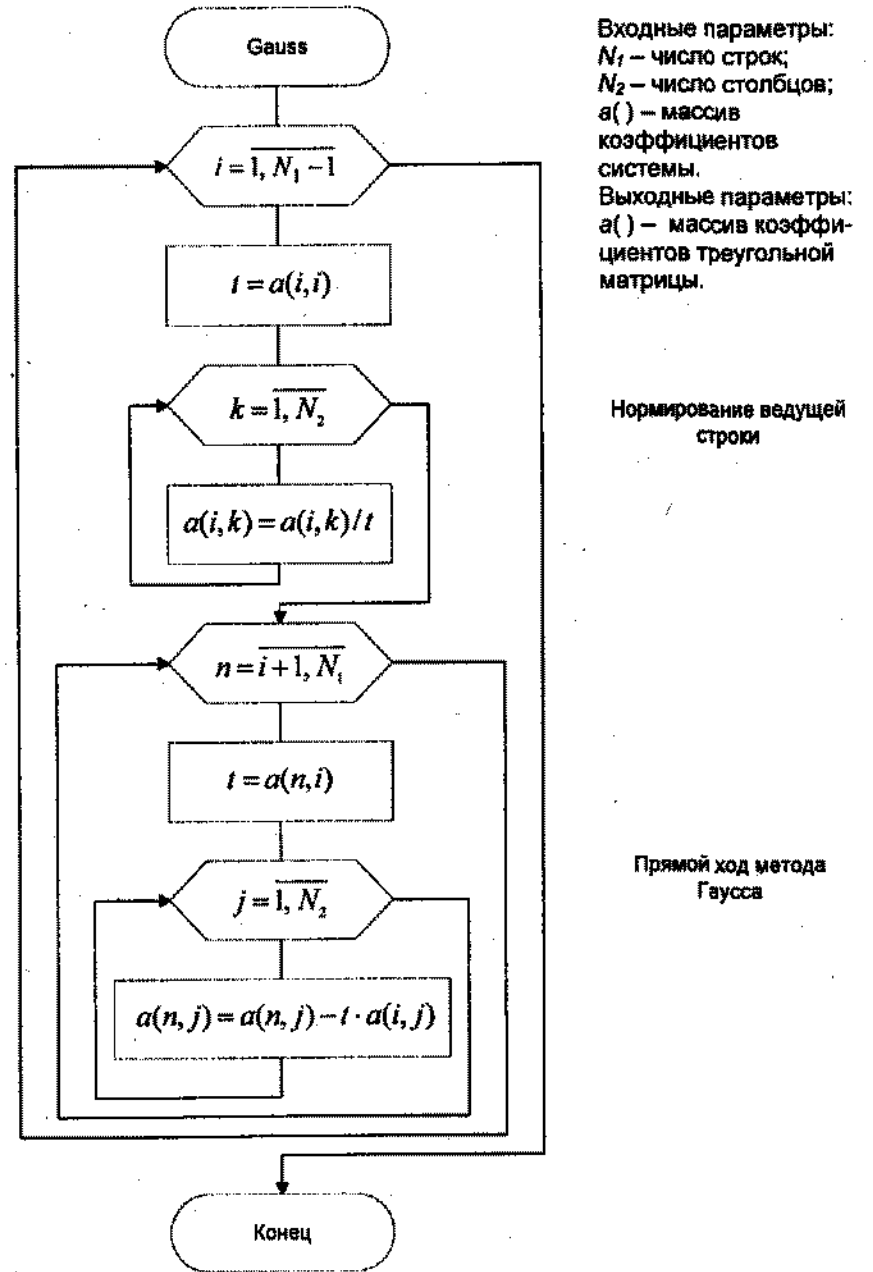


Рис.1.2

**Процедура решения системы уравнений, приведенной к треугольному виду**

Входные параметры:  
 $a()$  - треугольная матрица;  
 $N_1$  - число строк.

Выходные параметры:  
 $x()$  - решение системы.

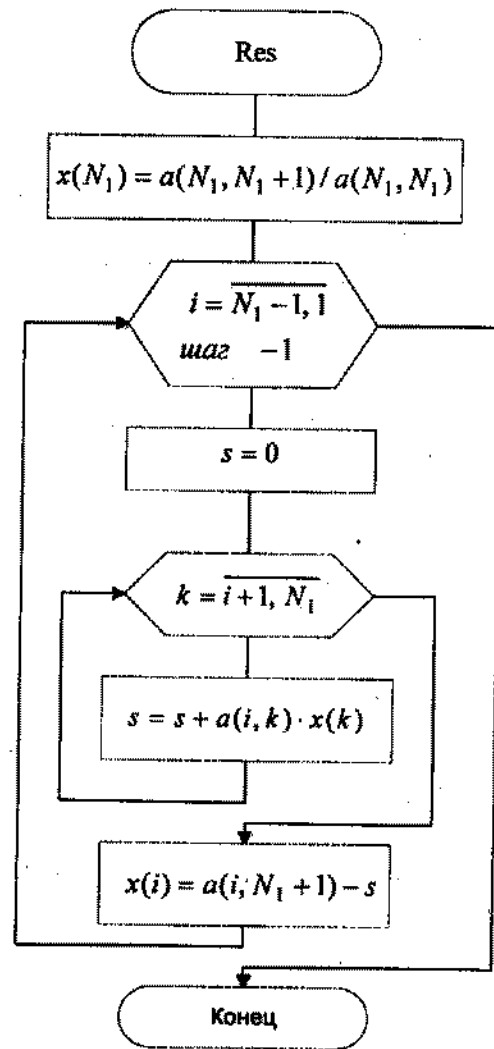


Рис.1.3

Обратный ход метода Гаусса

$$\begin{cases} x_1 + \tilde{a}_{12}x_2 + \dots + \tilde{a}_{1n}x_n = \tilde{a}_{1,n+1}, \\ 0 + a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n = a_{2,n+1}^{(1)}, \\ \dots \dots \dots \dots \dots \dots \\ 0 + a_{n2}^{(1)}x_2 + \dots + a_{nn}^{(1)}x_n = a_{n,n+1}^{(1)}. \end{cases}$$

где  $\tilde{a}_{ij} = \frac{a_{ij}}{a_{11}}$ ,  $a_{ij}^{(1)} = a_{ij} - a_{11}\tilde{a}_{ij}$ ,  $i = \overline{1, n}$ ,  $j = \overline{1, n+1}$ .

**Пример 1 [62].**

Решить систему уравнений методом Гаусса.

$$\begin{cases} x_1 + x_2 + x_3 - x_4 = 2, \\ x_1 - x_2 - x_3 + x_4 = 0, \\ 2x_1 + x_2 - x_3 + 2x_4 = 9, \\ 3x_1 + x_2 + 2x_3 - x_4 = 7. \end{cases}$$

Решение.

Запишем коэффициенты системы, обозначим уравнения буквами

$$\begin{array}{l} A_1 \quad 1 \quad 1 \quad 1 \quad -1 \quad | \quad 2 \quad (1^1) \\ A_2 \quad 1 \quad -1 \quad -1 \quad 1 \quad | \quad 0 \quad (1^2) \\ A_3 \quad 2 \quad 1 \quad -1 \quad 2 \quad | \quad 9 \quad (1^3) \\ A_4 \quad 3 \quad 1 \quad 2 \quad -1 \quad | \quad 7 \quad (1^4) \end{array}$$

Исключим члены, содержащие  $x_1$ , из всех уравнений, начиная со второго. Для этого коэффициенты уравнения  $(1^1)$  разделим на  $a_{11} = 1$ , затем элементы полученной строки, последовательно умноженные на  $a_{21}, a_{31}, a_{41}$  вычтем из  $(1^2), (1^3), (1^4)$ , получим

$$\begin{array}{l} B_1 = A_1 / 1 \quad 1 \quad 1 \quad 1 \quad -1 \quad | \quad 2 \\ B_2 = A_2 - B_1 \quad 0 \quad -2 \quad -2 \quad 2 \quad | \quad -2 \\ B_3 = A_3 - 2B_1 \quad 0 \quad -1 \quad -3 \quad 4 \quad | \quad 5 \\ B_4 = A_4 - 3B_1 \quad 0 \quad -2 \quad -1 \quad 2 \quad | \quad 1 \end{array}$$

Исключим члены, содержащие  $x_2$  из третьего и четвертого уравнений

$$\begin{array}{l} C_1 = B_1 \quad 1 \quad 1 \quad 1 \quad -1 \quad | \quad 2 \\ C_2 = B_2 / (-2) \quad 0 \quad 1 \quad 1 \quad -1 \quad | \quad 1 \\ C_3 = B_3 - (-1)C_2 \quad 0 \quad 0 \quad -2 \quad 3 \quad | \quad 6 \\ C_4 = B_4 - (-2)C_2 \quad 0 \quad 0 \quad 1 \quad 0 \quad | \quad 3 \end{array}$$

Исключение  $x_3$  дает

$$\begin{array}{l} B_1 \\ C_2 \\ D_3 = C_3 / (-2) \\ D_4 = C_4 - D_3 / 1 \end{array} \quad \begin{array}{ccc|c} 1 & 1 & 1 & -1 & 2 \\ 0 & 1 & 1 & -1 & 1 \\ 0 & 0 & 1 & -3/2 & -3 \\ 0 & 0 & 0 & 3/2 & 6 \end{array}$$

Затем получим

$$\begin{array}{l} B_1 \\ C_2 \\ D_3 \\ F_4 = D_4 / (3/2) \end{array} \quad \begin{array}{ccc|c} 1 & 1 & 1 & -1 & 2 \\ 0 & 1 & 1 & -1 & 1 \\ 0 & 0 & 1 & -3/2 & -3 \\ 0 & 0 & 0 & 1 & 4 \end{array}$$

Последняя система, записанная в форме уравнений:

$$\begin{cases} x_1 + x_2 + x_3 - x_4 = 2, \\ \quad x_2 + x_3 - x_4 = 1, \\ \quad \quad x_3 - \frac{3}{2}x_4 = -3, \\ \quad \quad \quad x_4 = 4. \end{cases} \quad (5)$$

Таким образом, система приведена к треугольному виду.

Из последнего уравнения (5) следует  $x_4 = 4$ , затем, последовательно получаем  $x_3 = 3$ ,  $x_2 = 2$ ,  $x_1 = 1$ .

Решение системы {1; 2; 3; 4}.

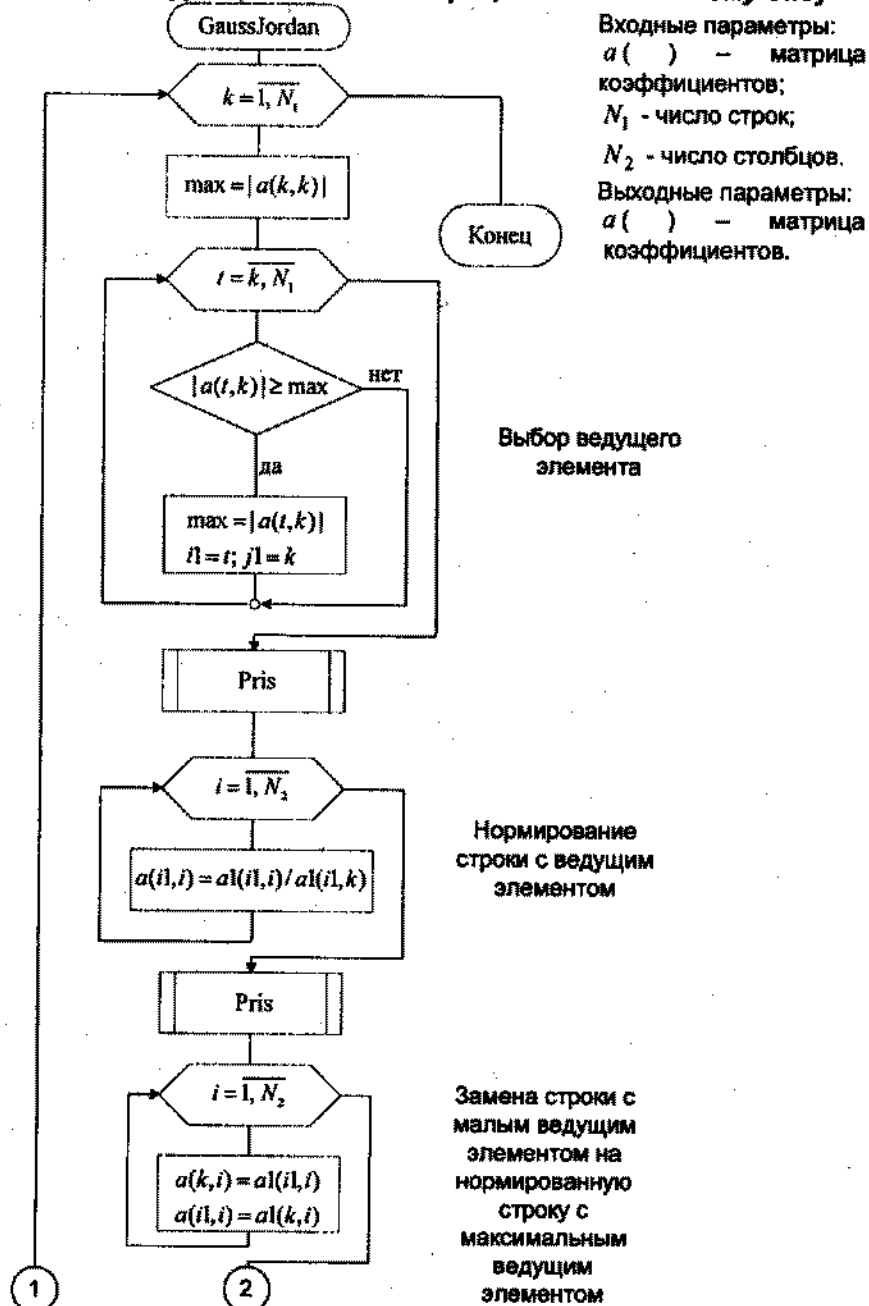
## 1.2. Метод Гаусса - Жордана

Суть метода – приведение матрицы исходной системы к диагональному виду путем преобразования коэффициентов уравнений, стоящих выше и ниже ведущего уравнения.

Применение описанного метода усложняется, если в каком-либо уравнении ведущий элемент равен нулю. В последнем случае уравнение нельзя нормировать. Однако изменив порядок, в котором расположены уравнения системы, эту трудность можно обойти.

Можно показать, что наибольшая точность достигается тогда, когда ведущий элемент имеет наибольшее значение. Поэтому строку с нулевым или малым ведущим элементом надо заменить на ту из стоящих под ней строк, в которой в том же столбце стоит элемент, имеющий наибольшее значение.

## Процедура приведения матрицы к диагональному виду



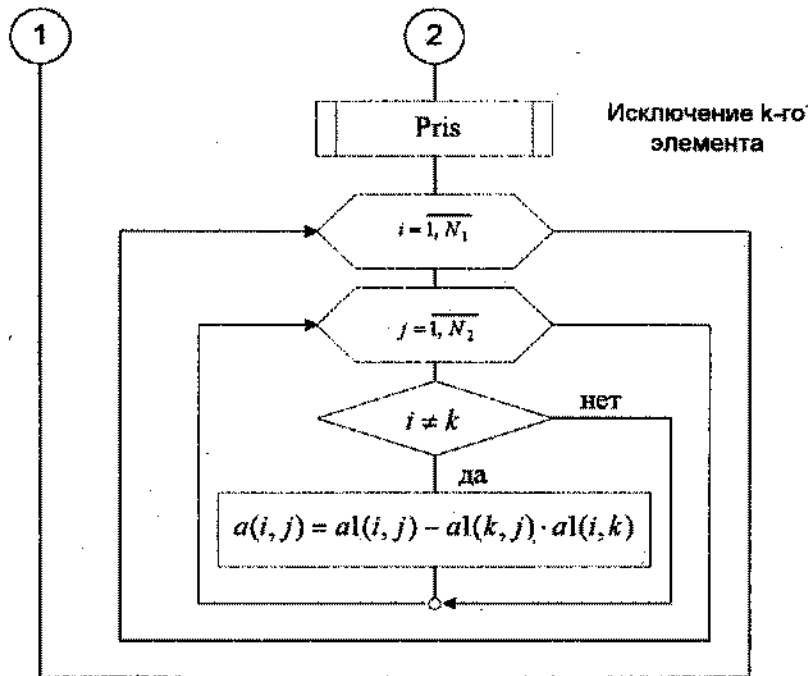


Рис.1.4

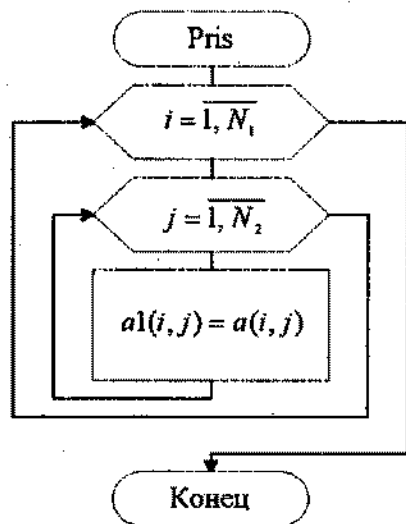


Рис.1.5

Процедура присваивает элементам массива  $a1(i, j)$  значения массива  $a(i, j)$ .

Входные параметры:

$a(i, j)$ .

Выходные параметры:

$a1(i, j)$ .

**Пример 2.**

Решить систему из примера 1 методом Гаусса-Жордана.

$$\begin{array}{l} A_1 \\ A_2 \\ A_3 \\ A_4 \end{array} \quad \left| \begin{array}{cccc} 1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 2 & 1 & -1 & 2 \\ 3 & 1 & 2 & -1 \end{array} \right| \begin{array}{l} 2 \\ 0 \\ 9 \\ 7 \end{array}$$

Решение.

Меняем местами первую и четвертую строку и исключаем  $x_1$  (первое уравнение – ведущее).

$$\begin{array}{l} B_1 = A_4 / 3 \\ B_2 = A_2 - B_1 \\ B_3 = A_3 - 2B_1 \\ B_4 = A_4 - B_1 \end{array} \quad \left| \begin{array}{cccc} 1 & 0.333 & 0.666 & -0.333 \\ 0 & -1.333 & -1.666 & 1.333 \\ 0 & 0.333 & -2.333 & 2.666 \\ 0 & 0.666 & 0.333 & -0.666 \end{array} \right| \begin{array}{l} 2.333 \\ -2.333 \\ 4.333 \\ -0.333 \end{array}$$

Исключаем  $x_2$  (второе уравнение – ведущее)

$$\begin{array}{l} C_1 = B_1 - C_2(0.333) \\ C_2 = B_2 / (-1.333) \\ C_3 = B_3 - C_2(0.333) \\ C_4 = B_4 - C_2(0.666) \end{array} \quad \left| \begin{array}{cccc} 1 & 0 & 0.25 & 0 \\ 0 & 1 & 1.25 & -1 \\ 0 & 0 & -2.75 & 3 \\ 0 & 0 & -0.5 & 0 \end{array} \right| \begin{array}{l} 1.75 \\ 1.75 \\ 3.75 \\ -1.5 \end{array}$$

Исключаем  $x_3$  (третье уравнение – ведущее)

$$\begin{array}{l} D_1 = C_1 - (0.25)D_3 \\ D_2 = C_2 - D_3(1.25) \\ D_3 = C_3 / (-2.75) \\ D_4 = C_4 - D_3(-0.5) \end{array} \quad \left| \begin{array}{cccc} 1 & 0 & 0 & 0.272 \\ 0 & 1 & 0 & 0.363 \\ 0 & 0 & 1 & -1.09 \\ 0 & 0 & 0 & -0.545 \end{array} \right| \begin{array}{l} 2.0909 \\ 3.454 \\ -1.363 \\ -2.181 \end{array}$$

Исключаем  $x_4$  из всех уравнений, кроме последнего

$$\begin{array}{l} F_1 = D_1 - (0.272)F_4 \\ F_2 = D_2 - (0.363)F_4 \\ F_3 = D_3 - (-1.09)F_4 \\ F_4 = D_4 / (3/2) \end{array} \quad \left| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right| \begin{array}{l} 1 \\ 2 \\ 3 \\ 4 \end{array} \quad (6)$$

Из записи системы (6) следует решение системы:  $\{1; 2; 3; 4\}$ .

Отличие метода Гаусса – Жордана от метода Гаусса в том, что преобразованная матрица имеет диагональный вид и нет необходимости в обратном ходе (как в методе Гаусса).

Запишем систему (1) в матричном виде:  $A \cdot \bar{X} = \bar{C}$ . Решение этой системы -  $\bar{X} = A^{-1} \cdot \bar{C}$ . Обозначим  $\bar{X}^*$  - решение, полученное по методу Гаусса, отличающееся от  $\bar{X}$  из-за погрешностей округления.

Степень отклонения полученного решения от точного можно охарактеризовать либо погрешностью  $\varepsilon = \bar{X} - \bar{X}^*$ , либо величиной  $r = C - A \cdot \bar{X}^*$ , называемой невязкой. Можно показать, что если одна из величин равна нулю, то и другая также равна нулю. Из малости погрешности  $\varepsilon$  обычно следует малость невязки  $r$ . Обратное утверждение справедливо не всегда.

### 1.3. Метод Холецкого решения систем линейных уравнений

Идея метода Холецкого заключается в преобразовании исходной расширенной матрицы коэффициентов -  $A$  системы к эквивалентной верхней треугольной матрице  $U$  с помощью преобразующей матрицы  $L$ .

$$A = \left( \begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & a_{1,n+1} \\ a_{21} & a_{22} & \dots & a_{2n} & a_{2,n+1} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} & a_{n,n+1} \end{array} \right),$$

$$U = \left( \begin{array}{cccc|c} 1 & u_{12} & u_{13} & \dots & u_{1n} & u_{1,n+1} \\ 0 & 1 & u_{23} & \dots & u_{2n} & u_{2,n+1} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & u_{n,n+1} \end{array} \right),$$

$$L = \left( \begin{array}{cccc} l_{11} & 0 & \dots & 0 \\ l_{21} & l_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ l_{n1} & l_{n2} & \dots & l_{nn} \end{array} \right),$$

$$(L) \cdot (U) = (A). \quad (7)$$

Из условия (7) вытекают соотношения:

$$1) \text{ Для первой строки} \\ a_{ij} = a_{ij} / a_{11}, \quad j = 2, 3, \dots, n+1. \quad (8)$$

2) Для  $m$ -го столбца и  $m$ -й строки ( $m = 2, 3, \dots, n$ )

а) для очередного  $l_{im}$  столбца

$$a_{im} = a_{im} - \sum_{k=1}^{m-1} a_{ik} a_{km}, \quad i = m, \dots, n; \quad (9)$$

б) для очередной  $u_{mj}$  строки

$$a_{mj} = \frac{a_{mj} - \sum_{k=1}^{m-1} a_{mk} a_{kj}}{a_{mm}}, \quad j = m+1, \dots, n+1. \quad (10)$$

$$3) \quad x_n = a_{n,n+1}, \quad (11)$$

$$x_i = a_{i,n+1} - \sum_{k=i+1}^n a_{ik} x_k, \quad i = n-1, \dots, 1. \quad (12)$$

Метод Холецкого предваряется частичной перестановкой строк исходной матрицы с тем, чтобы элементы главной диагонали матрицы коэффициентов ( $a_{mm}$ ) не были равны нулю.

#### Замечания

- Поскольку на каждом шаге используется только информация, полученная ранее, то результаты можно записывать на место исходных в матрицу  $A$ . Тогда первый столбец матрицы  $L$  сформируется автоматически ( $l_{i1} = a_{i1}$ ,  $i = \overline{1, n}$ ).
- Метод Холецкого экономичнее метода Гаусса - Жордана по обращению к оперативной памяти и по машинному времени.

### 1.4. Метод прогонки

При моделировании некоторых задач, а также при численном решении краевых задач для дифференциальных уравнений получаются системы с разреженной матрицей - матрица  $A$  содержит большое число нулевых элементов.

В частности, представляют интерес системы с трехдиагональной матрицей. Запишем систему уравнений в виде:

$$\begin{cases} b_1 x_1 + c_1 x_2 & & & & & = d_1, \\ a_2 x_1 + b_2 x_2 + c_2 x_3 & & & & & = d_2, \\ & a_3 x_2 + b_3 x_3 + c_3 x_4 & & & & = d_3, \\ \dots & \dots & \dots & \dots & \dots & \dots \\ & & a_{n-1} x_{n-2} + b_{n-1} x_{n-1} + c_{n-1} x_n & & & = d_{n-1}, \\ & & a_n x_{n-1} + b_n x_n & & & = d_n, \end{cases} \quad (13)$$

Структурная схема алгоритма метода Холецкого

Входные параметры:  
 $a()$  – матрица коэффициентов;  
 $N_1$  – число строк,  
 $N_2$  – число столбцов.  
 Выходные параметры:  
 $x()$  – решение системы.

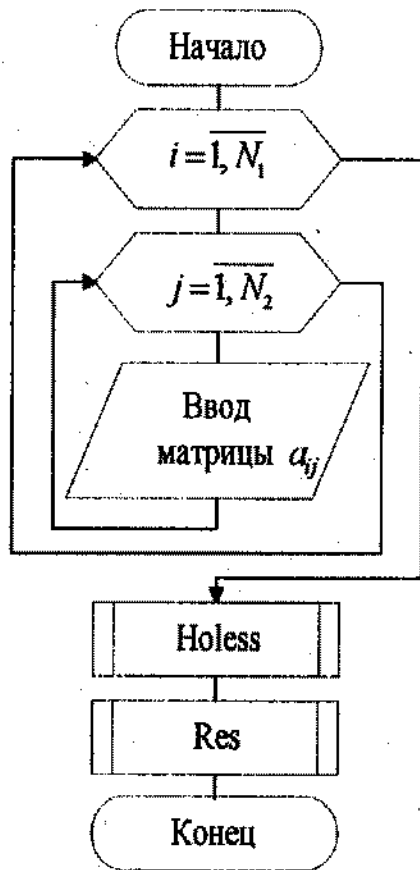
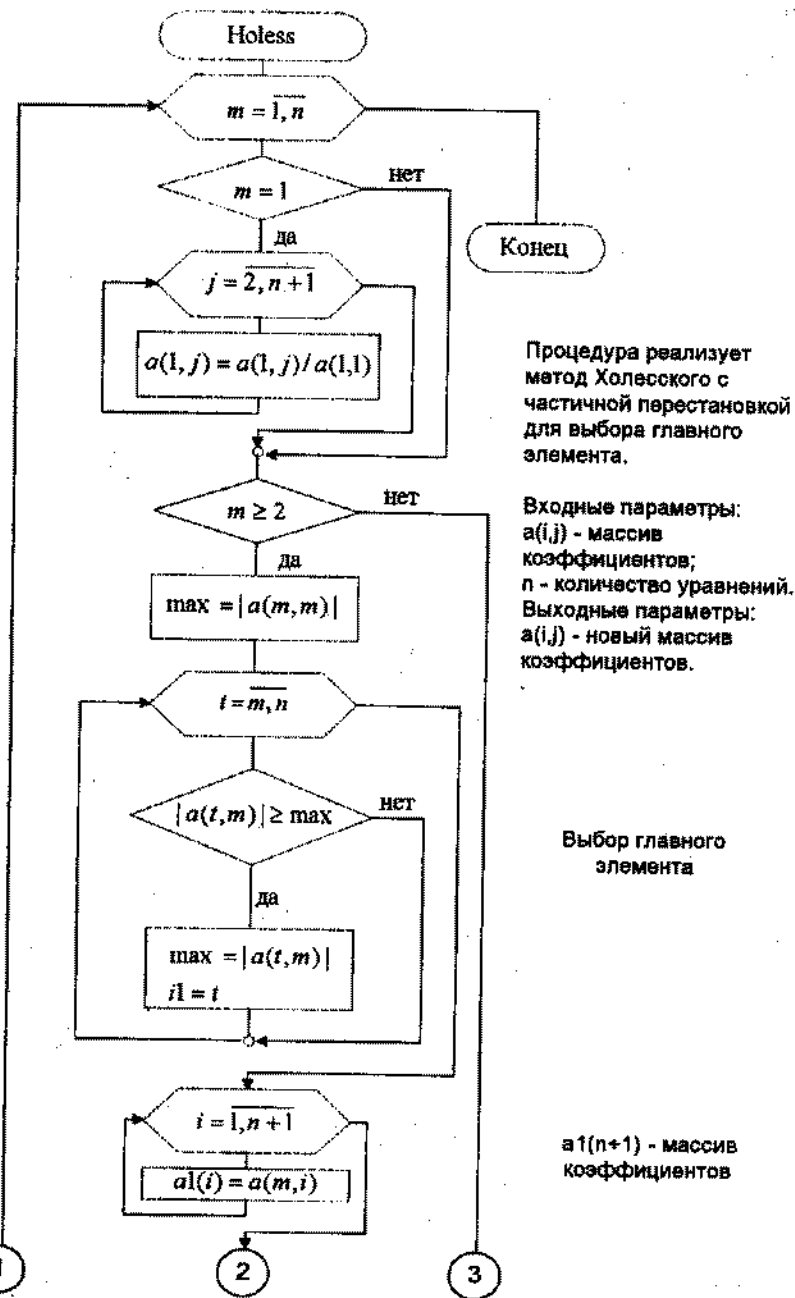


Рис. 1.6

9840458



Процедура реализует метод Холецкого с частичной перестановкой для выбора главного элемента.

Входные параметры:  
 $a(i, j)$  - массив коэффициентов;  
 $n$  - количество уравнений.  
 Выходные параметры:  
 $a(i, j)$  - новый массив коэффициентов.

Выбор главного элемента

$a(1(n+1))$  - массив коэффициентов

Рис. 1.7



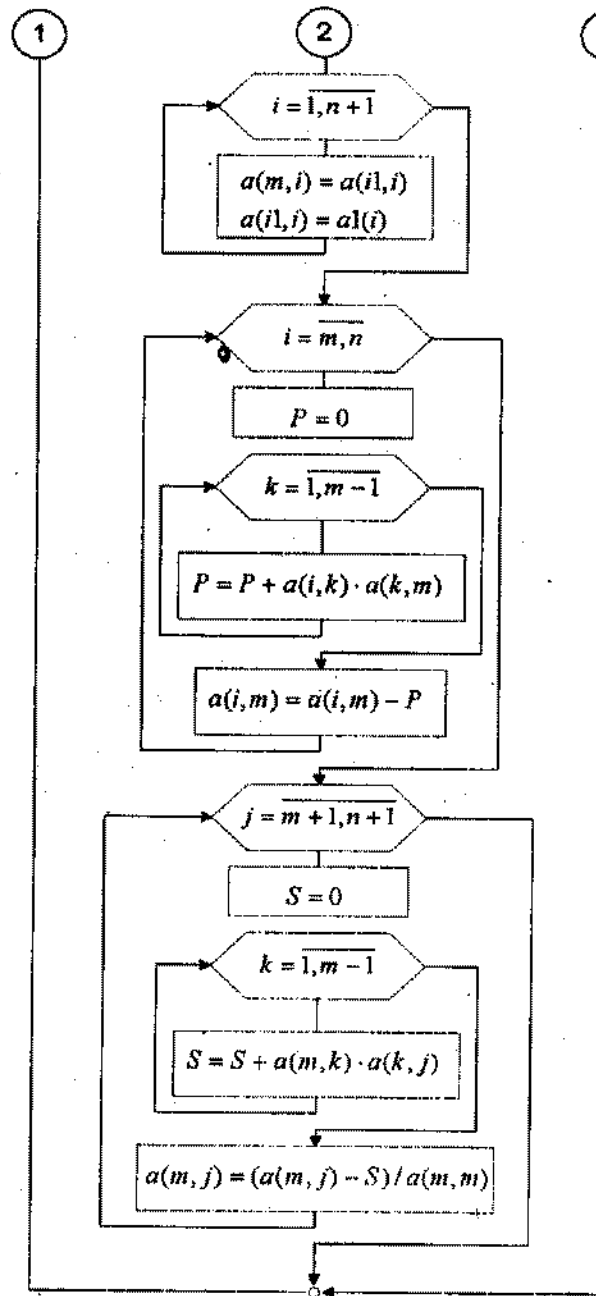


Рис.1.8

Замена ведущей строки на строку с выбранным главным элементом.

Формирование очередного  $l(i, m)$  столбца.

Формирование очередной  $a(m, j)$  строки.

### Процедура формирования решения системы уравнений

Входные параметры:

$a(i, j)$ -массив коэффициентов;  
 $n$ -число уравнений.

Выходные параметры:

$x(i)$ -массив решений системы.

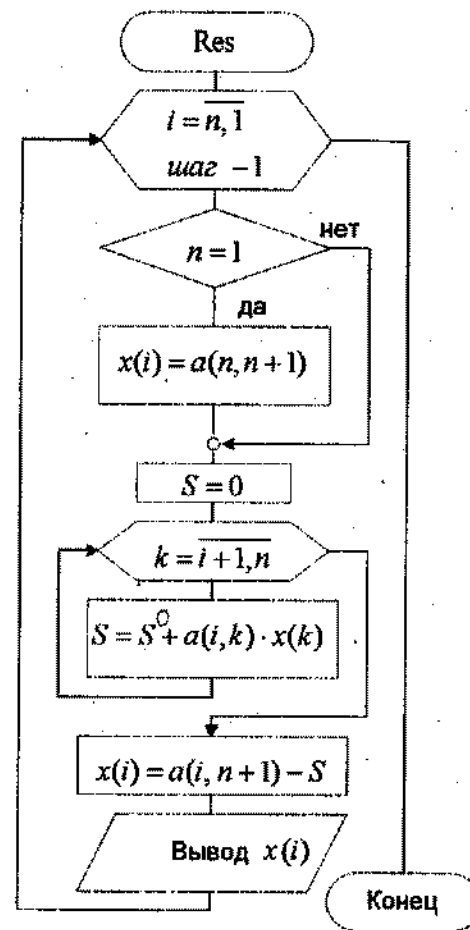


Рис. 1.9

$b_1, b_2, \dots, b_n$  - элементы главной диагонали ( $b_i \neq 0, i = \overline{1, n}$ ),

$c_1, c_2, \dots, c_{n-1}$  - элементы над главной диагональю,

$a_2, a_3, \dots, a_n$  - элементы под главной диагональю.

Метод прогонки состоит из двух этапов:

- прямая прогонка,
- обратная прогонка.

С помощью прогоночных коэффициентов  $A_i$  и  $B_i$  выражаем

$$x_i = A_i x_{i+1} + B_i, \quad i = \overline{1, n-1}. \quad (14)$$

Из первого уравнения (13) и выражения (14) при  $i=1$  следует

$$\begin{cases} x_1 = -\frac{c_1}{b_1} x_2 + \frac{d_1}{b_1}, \\ x_1 = A_1 x_2 + B_1, \end{cases}$$

$$\text{откуда } A_1 = -\frac{c_1}{b_1}, \quad B_1 = \frac{d_1}{b_1}. \quad (15)$$

$$\text{Далее } A_2 = -\frac{c_2}{l_2}, \quad B_2 = \frac{d_2 - a_2 B_1}{l_2},$$

$$l_2 = a_2 A_1 + b_2. \quad (\text{из (13) и (14) при } i=2).$$

В общем случае рекуррентные соотношения для любого номера  $i = 2, 3, \dots, n-1$

$$A_i = -\frac{c_i}{l_i}, \quad B_i = \frac{d_i - a_i B_{i-1}}{l_i}. \quad (16)$$

Обратная прогонка состоит в последовательном вычислении неизвестных  $x_i$ .

$$x_n = \frac{d_n - a_n B_{n-1}}{b_n + a_n A_{n-1}}. \quad (17)$$

Далее, используя (14), (15), (16), вычисляются  $x_{n-1}, x_{n-2}, \dots, x_1$ .

Схема алгоритма на рис. 10–11.

#### Замечания

- В формулах (16) возможно равенство нулю знаменателей. Доказано, что при выполнении условия

$$|b_i| \geq |a_i| + |c_i|, \quad (18)$$

причем хотя бы для одного значения  $i$  имеет место строгое неравенство, деления на ноль не возникает и система (13) имеет единственное решение.

- Условие (18) является достаточным условием устойчивости метода прогонки относительно погрешности округления.

- Метод прогонки позволяет решать большие системы уравнений.

#### Пример 3.

Решить систему уравнений

$$\begin{cases} 4x_1 - x_2 = 3, \\ 2x_1 + 6x_2 - x_3 = 7, \\ \quad 2x_2 - 3x_3 + x_4 = 0, \\ \quad \quad x_3 + 2x_4 = 3 \end{cases}$$

методом прогонки.

*Решение.*

Матрица системы – трехдиагональная. Условие сходимости (18) выполнено. Воспользуемся предложенным алгоритмом.

Запишем таблицу коэффициентов.

Таблица 1

$i$	$a_i$	$b_i$	$c_i$	$d_i$
1	–	4	–1	3
2	2	6	–1	7
3	2	–3	1	0
4	1	2	–	3

$$\text{Согласно (15) } A_1 = -\frac{c_1}{b_1} = \frac{1}{4}, \quad B_1 = \frac{d_1}{b_1} = \frac{3}{4}.$$

По формулам (16)

$$l_2 = a_2 A_1 + b_2 = \frac{1}{2} + 6 = \frac{13}{2},$$

$$A_2 = -\frac{c_2}{l_2} = \frac{2}{13}, \quad B_2 = \frac{d_2 - a_2 B_1}{l_2} = \frac{11}{13}.$$

$$l_3 = a_3 A_2 + b_3 = -\frac{35}{13},$$

$$A_3 = -\frac{c_3}{l_3} = \frac{13}{35}, \quad B_3 = \frac{d_3 - a_3 B_2}{l_3} = \frac{22}{35}.$$

$$x_4 = \frac{d_4 - a_4 B_3}{b_4 + a_4 A_3} = 1, \quad x_3 = A_3 x_4 + B_3 = 1,$$

$$x_2 = A_2 x_3 + B_2 = 1, \quad x_1 = A_1 x_2 + B_1 = 1.$$

Итак, решение системы:  $\{1; 1; 1; 1\}$ .

Структурная схема алгоритма метода прогонки

Входные параметры:  
 $a()$ ,  $b()$ ,  $c()$ ,  $d()$  – массивы  
 коэффициентов системы.  
 Выходные параметры:  
 $x()$  – массив решений.

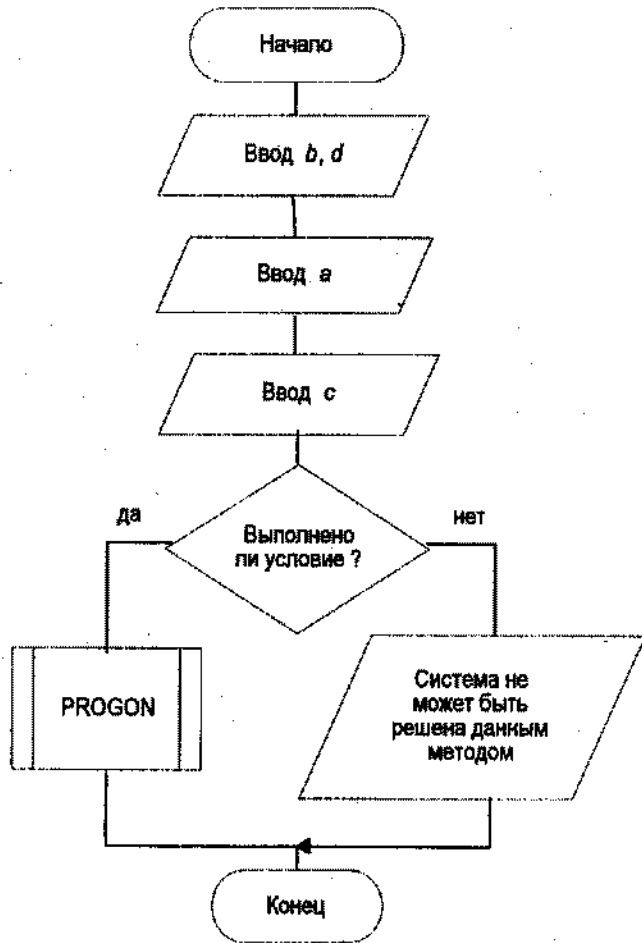


Рис 1.10

Процедура решения системы уравнений методом прогонки

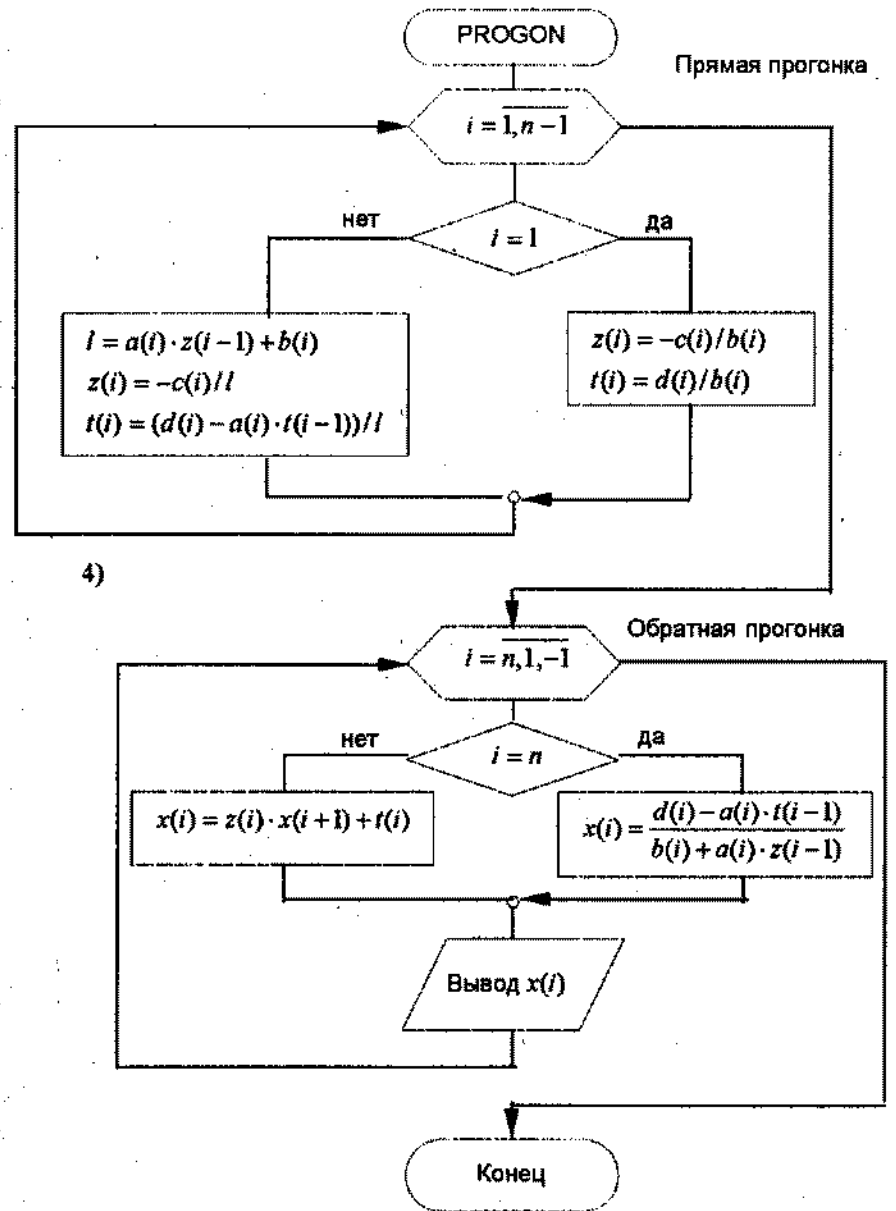


Рис. 1.11

## 1.5. Методы итерации

Итерационные схемы для систем линейных уравнений основаны на приведении их к виду:

$$x_i = \frac{1}{a_{ii}} (c_i - a_{i1}x_1 - a_{i2}x_2 - \dots - a_{i,i-1}x_{i-1} - a_{i,i+1}x_{i+1} - \dots - a_{in}x_n). \quad (19)$$

Система (19) получена из системы (1) в предположении, что диагональные элементы  $a_{11}, a_{22}, \dots, a_{nn}$  отличны от нуля.

Систему (1) также можно записать в виде

$$\bar{X} = A\bar{X} + \bar{C}. \quad (20)$$

Метод Якоби использует произвольно выбранные начальные приближения  $x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}$  для вычисления  $x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}$  и т.д. до тех пор, пока не будет достигнута точность, либо станет очевидно, что процесс расходится.

Приближения к решению вычисляются по формулам:

$$x_i^{(k)} = \sum_{j=1}^n a_{ij} x_j^{(k-1)} + c_i,$$

$k$  - номер итерации,  $k = 1, 2, \dots, i = \overline{1, n}$ . (21)

При этом замещение "старых" значений неизвестных на "новые" происходит одновременно.

Эффективнее использовать метод Гаусса - Зейделя - прием последовательного замещения уже вычисленных "новых" значений.

Согласно методу Гаусса - Зейделя  $k$ -е приближение к решению имеет вид

$$x_i^k = \sum_{j=1}^{i-1} a_{ij} x_j^k + \sum_{j=i+1}^n a_{ij} x_j^{(k-1)} + c_i. \quad (22)$$

Для сходимости итерационного процесса достаточно выполнения условий

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|, \quad i = \overline{1, n} \quad (23)$$

или  $\|A\| < 1$ .

Сходящийся процесс итерации обладает свойством самоисправляемости, т.е. отдельная ошибка в вычислениях не влияет на сходимость.

Итерации продолжаются, пока справедливо неравенство

$$\delta = \max_{1 \leq i \leq n} |x_i^{(k)} - x_i^{(k-1)}| > E, \quad \text{где } E - \text{ заданная точность.}$$

Оценка погрешности вычислений:

$$\|\bar{X}^* - \bar{X}^{(k)}\| \leq \frac{\|A^k\|}{1 - \|A\|} \|\bar{X}_1 - \bar{X}_0\|. \quad (24)$$

Если  $\bar{X}_0 = \bar{C}$  (система приведена к виду (20)), то

$$\|\bar{X}^* - \bar{X}^{(k)}\| \leq \frac{\|A^{k+1}\|}{1 - \|A\|} \|C\|,$$

где  $\bar{X}^*$  - точное решение,  $\|A\|$  - норма матрицы.

Число шагов ( $k$ ), обеспечивающих точность  $E$  решения системы, может быть определено из (24).

$$\|x^* - x^{(k)}\| \leq \frac{\|A^{k+1}\|}{1 - \|A\|} \|C\| \leq E. \quad (24')$$

### Пример 4.

Решить систему

$$\begin{cases} 4x_1 + 0.24x_2 - 0.08x_3 = 8, \\ 0.09x_1 + 3x_2 - 0.15x_3 = 9, \\ 0.04x_1 - 0.08x_2 + 4x_3 = 20 \end{cases}$$

методом Якоби с точностью  $\varepsilon = 0.001$ .

Решение.

Условие сходимости (23) выполнено.

Приведем систему к виду (19)

$$\begin{cases} x_1 = 2 - 0.06x_2 + 0.02x_3, \\ x_2 = 3 - 0.03x_1 + 0.05x_3, \\ x_3 = 5 - 0.01x_1 + 0.02x_2. \end{cases} \quad (25)$$

Выберем начальное приближение

$$x_1^{(0)} = 2; \quad x_2^{(0)} = 3; \quad x_3^{(0)} = 5.$$

Подставляя эти значения в правые части уравнений (25), получим первое приближение

$$\begin{cases} x_1^{(1)} = 2 - 0.06 \cdot 3 + 0.02 \cdot 5 = 1.92, \\ x_2^{(1)} = 3 - 0.03 \cdot 2 + 0.05 \cdot 5 = 3.19, \\ x_3^{(1)} = 5 - 0.01 \cdot 2 + 0.02 \cdot 3 = 5.04. \end{cases}$$

Дальнейшие вычисления приведены в таблице 2.

Таблица 2

$k$	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	2	3	5
1	1.92	3.19	5.04
2	1.9094	3.1944	5.0446
3	1.90923	3.19495	5.04485

**Пример 5.**

Решить систему

$$\begin{cases} 4x_1 + 0.24x_2 - 0.08x_3 = 8, \\ 0.09x_1 + 3x_2 - 0.15x_3 = 9, \\ 0.04x_1 - 0.08x_2 + 4x_3 = 20 \end{cases}$$

методом Гаусса – Зейделя, выполнив две итерации.

*Решение.*

Аналогично предыдущему примеру приведем систему к виду (25) и выберем те же начальные приближения.

Применяем процесс Гаусса – Зейделя

$$\begin{cases} x_1^{(1)} = 2 - 0.06 \cdot 3 + 0.02 \cdot 5 = 1.92, \\ x_2^{(1)} = 3 - 0.03 \cdot 1.92 + 0.05 \cdot 5 = 3.1924, \\ x_3^{(1)} = 5 - 0.01 \cdot 1.92 + 0.02 \cdot 3.1924 = 5.04465. \end{cases}$$

Находим второе приближение

$$\begin{cases} x_1^{(2)} = 2 - 0.06 \cdot 3.1924 + 0.02 \cdot 5.04465 = 1.90934, \\ x_2^{(2)} = 3 - 0.03 \cdot 1.90934 + 0.05 \cdot 5.04465 = 3.19495, \\ x_3^{(2)} = 5 - 0.01 \cdot 1.90934 + 0.02 \cdot 3.19495 = 5.04481. \end{cases}$$

Таблица 3

$k$	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	2	3	5
1	1.92	3.1924	5.04465
2	1.90934	3.19495	5.04481

Заметим, что метод Гаусса – Зейделя требует меньшего числа итераций при той же точности, чем метод Якоби.

Алгоритмы методов итерации рассмотрены в [26] и [58].

## 1.6. Исследование и решение систем линейных уравнений методом Жордана

В экономических исследованиях и планировании рассматриваются системы  $k$  уравнений первой степени с  $n$  неизвестными.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \quad (26)$$

Систему (26) перепишем в виде

$$\sum_{j=1}^n a_{ij}x_j = b_i, \quad i = \overline{1, m}. \quad (27)$$

Если система (26) имеет решения, она называется совместной, в противном случае – несовместной. Совместная система называется определенной или неопределенной в зависимости от того, имеет она одно или более решений.

Предположим, что в системе уравнений (26)  $a_{rs} \neq 0$ . Требуется исключить неизвестную  $x_s$  из всех уравнений системы, кроме  $r$ -ого уравнения. Назовем  $a_{rs}$  разрешающим (ведущим) коэффициентом, уравнение с номером  $r$  – разрешающим уравнением.

Исключение неизвестной  $x_s$  производится с помощью соотношений:

$$\begin{aligned} a'_{ij} &= a_{ij} - \frac{a_{is}a_{rs}}{a_{rs}}, & i \neq r, \\ a'_{rj} &= \frac{a_{rj}}{a_{rs}}, & i = r, \\ b'_i &= b_i - \frac{a_{is}}{a_{rs}}b_r, & i \neq r, \\ b'_r &= \frac{b_r}{a_{rs}}, & i = r, \end{aligned} \quad (28)$$

$$i = 1, 2, \dots, r-1, r+1, \dots, m, \quad j = \overline{1, n}.$$

Применяя формулы (28), приведем систему (26) к виду

$$x_i + \sum_{j=m+1}^n g_{ij}x_j = h_i, \quad i = \overline{1, m}, \quad m < n \quad (29)$$

или

$$\begin{cases} x_1 + \dots + g_{1,m+1}x_{m+1} + \dots + g_{1n}x_n = h_1, \\ \dots \\ x_2 + \dots + g_{2,m+1}x_{m+1} + \dots + g_{2n}x_n = h_2, \\ \dots \\ \dots \\ x_m + g_{m,m+1}x_{m+1} + \dots + g_{mn}x_n = h_m. \end{cases} \quad (30)$$

Система (30) эквивалентна системе (26).

Назовем систему (30) системой, приведенной к каноническому виду, переменные  $x_1, x_2, \dots, x_m$  - базисными,  $x_{m+1}, x_{m+2}, \dots, x_n$  - свободными.

Общее решение системы (30)

$$\begin{cases} x_1 = h_1 - (g_{1,m+1}x_{m+1} + \dots + g_{1n}x_n), \\ \dots \\ x_m = h_m - (g_{m,m+1}x_{m+1} + \dots + g_{mn}x_n). \end{cases} \quad (31)$$

Базисное решение, отвечающее нулевым значениям свободных неизвестных

$$x_1 = h_1, \quad x_2 = h_2, \quad \dots \quad x_m = h_m, \quad x_{m+1} = 0, \quad \dots \quad x_n = 0. \quad (32)$$

#### Замечание

Эквивалентное преобразование к виду (30) возможно, если система (26) совместна. В случае несовместности системы в процессе преобразования появляется уравнение вида

$$0 \cdot x_1 + 0 \cdot x_2 + \dots + 0 \cdot x_n = b'.$$

В этом случае решения системы (26) не существует.

Если в системе (26) и в процессе исключения неизвестных правые части уравнений неотрицательны, то базисные решения будут неотрицательными. Для того, чтобы свободные члены всех уравнений оставались неотрицательными, необходимо:

- в качестве разрешающей переменной выбирать переменную, при которой есть хоть один положительный коэффициент,
- в качестве разрешающего уравнения следует брать уравнение, соответствующее наименьшему среди отношений свободных членов уравнений к соответствующим положительным коэффициентам при разрешающей переменной.

Руководствуясь правилами исключения можно найти другое базисное решение. Всего базисных неизвестных  $m$ , свободных -  $(n-m)$ , количество базисных решений  $C_n^m$  - число сочетаний из числа неизвестных по числу оставшихся уравнений.

$$C_n^m = \frac{n!}{m!(n-m)!}. \quad (33)$$

#### Пример 6 [35].

Получить три базисных решения системы уравнений.

$$\begin{cases} 3x_1 + 9x_2 + 5x_3 - 2x_4 - 2x_5 = 7, \\ 2x_1 - 2x_2 - 2x_3 + x_4 - x_5 = 1, \\ 9x_1 - 8x_2 - 7x_3 + 3x_4 - 4x_5 = 3. \end{cases}$$

Решение.

В данной задаче существует  $C_5^3 = \frac{5!}{3!2!} = \frac{3! \cdot 4 \cdot 5}{3! \cdot 2} = 10$  базисных

решений.

Пусть  $a_{24} = 1$  - разрешающий элемент. Исключим с помощью формул (28) неизвестную  $x_4$  из всех уравнений, кроме второго.

$$\begin{cases} 7x_1 + 5x_2 + x_3 - 4x_5 = 9, \\ 2x_1 - 2x_2 - 2x_3 + x_4 - x_5 = 1, \\ 3x_1 - 2x_2 - x_3 - x_5 = 0. \end{cases}$$

В результате исключения  $x_4$  из всех уравнений, кроме третьего, получаем

$$\begin{cases} 10x_1 + 3x_2 - 5x_5 = 9, \\ -4x_1 + 2x_2 + x_4 + x_5 = 1, \\ -3x_1 + 2x_2 + x_3 - 4x_5 = 0. \end{cases}$$

Принимаем  $x_1$  за разрешающую переменную, первое уравнение - за разрешающее.

Тогда

$$\begin{cases} x_1 + 0.3x_2 - 0.5x_5 = 0.9, \\ +3.2x_2 + x_4 - x_5 = 4.6, \\ +2.9x_2 + x_3 - 0.5x_5 = 2.7, \end{cases}$$

где  $x_1, x_3, x_4$  - базисные переменные,  $x_2, x_5$  - свободные.

$$\begin{cases} x_1 = 0.9 - 0.3x_2 + 0.5x_5, \\ x_4 = 4.6 - 3.2x_2 + x_5, \\ x_3 = 2.7 - 2.9x_2 + 0.5x_5. \end{cases}$$

Базисное решение  $\bar{X}^{(0)} = \{0.9; 0; 2.7; 4.6; 0\}$ .

Аналогичными преобразованиями можно получить другие базисные решения

$$\bar{X}^{(1)} = \{0.47; 1.44; -1.47; 0; 0\}, \quad \bar{X}^{(2)} = \{-1; 0.31; 0; 0; -3.61\}.$$

**Пример 7.**

Исследовать систему уравнений

$$\begin{cases} x_1 - 2x_2 + 3x_3 = -2, \\ 5x_1 - 3x_2 + x_3 = 4, \\ 3x_1 + 8x_2 - x_3 = 4, \\ 6x_1 + 2x_2 + 3x_3 = 3, \\ 7x_1 + 4x_2 - 4x_3 = 11. \end{cases}$$

**Решение.**

В результате преобразования матрицы коэффициентов система имеет вид:

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

имеющей единственное решение  $\{1; 0; -1\}$ .**Пример 8.**

Исследовать систему

$$\begin{cases} x_1 + 2x_2 - x_3 + x_4 = 3, \\ 2x_1 + 7x_2 + 2x_3 - x_4 = 10, \\ x_1 + 5x_2 + 3x_3 - 2x_4 = 8. \end{cases}$$

**Решение.**

Вычитая из второго уравнения сумму первого и третьего уравнений, получим:

$$0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = 1.$$

Следовательно, система несовместна.

**2. НЕЛИНЕЙНЫЕ УРАВНЕНИЯ**

Пусть требуется решить уравнение

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

где  $f(x)$  - непрерывная функция в конечном или бесконечном интервале.Если  $f(x)$  представляет собой многочлен, то уравнение называют алгебраическим, в противном случае - трансцендентным.Всякое значение  $x = x^*$ , обращающее  $f(x)$  в ноль, называется корнем этого уравнения.

Займемся задачей отыскания изолированных корней уравнения (1) (корень изолированный, если существует окрестность, в которой этот корень единственный).

Решение состоит из двух этапов:

- отделение корней,
- уточнение корней.

При отыскании действительных корней этап отделения производится либо графически, либо аналитически, основываясь на известной теореме:

**Теорема 1.**Если непрерывная функция  $f(x)$  принимает значения разных знаков на концах отрезка  $[a; b]$ , т.е.

$$f(a)f(b) < 0, \quad (2)$$

то внутри этого отрезка существует по меньшей мере один корень уравнения  $f(x) = 0$ , т.е. такое значение  $x^*$ , что  $f(x^*) = 0$ .Корень  $x^*$  будет единственным на отрезке  $[a; b]$ , если производная  $f'(x)$  существует и сохраняет знак внутри интервала  $(a; b)$ .Практически отделение корней начинается с определения граничных точек  $x=a$  и  $x=b$  области существования функции  $f(x)$ . Затем вычисляют значения  $f(x)$  на  $[a; b]$  через промежутки произвольной длины  $h$ , учитывающие особенности функции  $f(x)$ , до смены знака при переходе от  $f(x)$  к  $f(x+h)$ .**Пример 1.**Отделить корни уравнения  $f(x) = \ln x + x^2 - 0.5 = 0$ .**Решение.**С учетом области определения ( $x > 0$ ) составим схему знаков функции.

$x$	0.1	0.5	1
$\text{sign } f(x)$	(-)	(-)	(+)

$f'(x) = \frac{1+2x^2}{x} > 0$  при  $x > 0$ , т.е. единственный корень принадлежит отрезку  $[0.5; 1]$ .

### Теорема 2.

Пусть  $x^*$  - точный, а  $x_n$  - приближенный корни уравнения (1), принадлежащие отрезку  $[a; b]$ , причем  $|f'(x)| \geq m_1 > 0$ .

Справедлива оценка  $|x_n - x^*| \leq \frac{|f'(x_n)|}{m_1}$ . (3)

Этап уточнения корней заключается в вычислении значения  $x_n$ , обеспечивающего выполнение условия  $|f(x_n)| < \varepsilon$  с некоторой погрешностью, не превышающей заданной точности  $\varepsilon$ .

Рассмотрим некоторые алгоритмы уточнения корня.

## 2.1. Метод половинного деления

Пусть требуется уточнить единственный корень уравнения  $f(x) = 0$ , принадлежащий отрезку  $[a; b]$  (отрезок неопределенности).

Точка  $c = \frac{a+b}{2}$  - середина отрезка  $[a; b]$ .

Если  $f(c) = 0$ , то корень найден. В противном случае для дальнейшего рассмотрения оставляем ту из половин  $[a; c]$  или  $[c; b]$ , на концах которой знаки функции  $f(x)$  различны. При этом получается последовательность вложенных отрезков, содержащих искомый корень.

На каждом шаге длина отрезка неопределенности уменьшается вдвое. Метод сходится всегда.

Условием окончания поиска корня может быть, например,

$|f(x)| < E$  или  $\frac{|b-a|}{2^n} < E$ , (4)

где  $E$  - точность,  $[a; b]$  - начальный отрезок,  $n$  - число итераций.

### Пример 2.

Методом половинного деления сделать шесть итераций уточнения корня уравнения  $f(x) = x^4 + 2x^3 - x - 1 = 0$ , принадлежащего отрезку  $[0; 1]$ .

Решение.

Таблица 1

$n$	$a$	$b$	$c$	$f(a)$	$f(b)$	$f(c)$
1	0	1	.5	-1	1	-1.19
2	0.5	1	0.75	-1.19	1	-0.59
3	0.75	1	0.875	-0.59	1	.05
4	0.75	0.875	0.812	-0.59	0.05	0.304
5	0.812	0.875	0.843	-0.304	0.05	-0.135
6	0.843	0.875	0.859	-0.135	0.05	0.045

После шести итераций корень локализован на отрезке  $[0.843; 0.875]$ .

### Геометрическая иллюстрация метода половинного деления

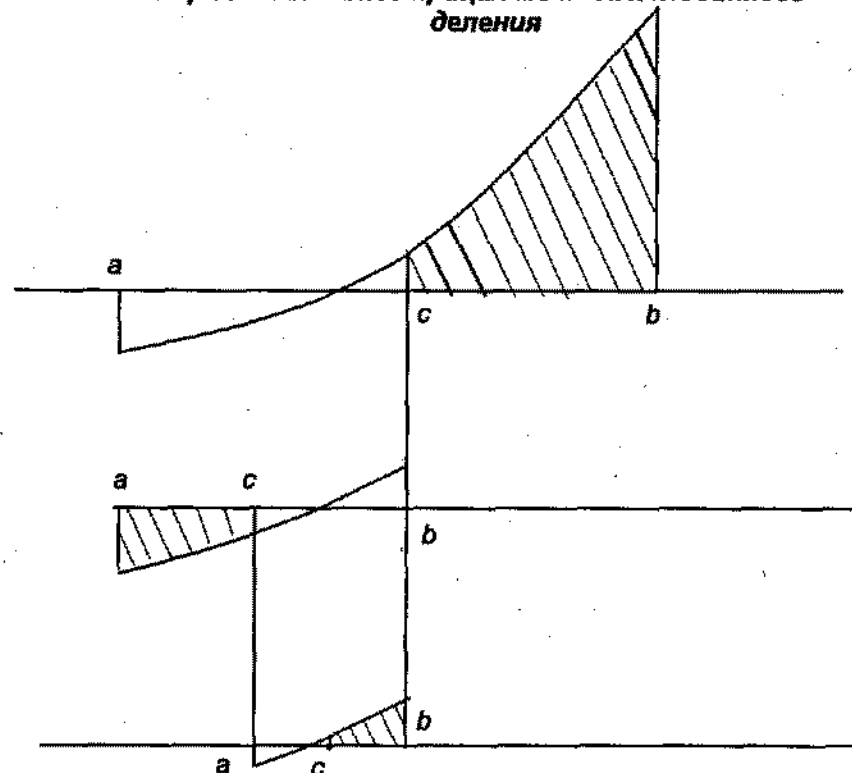


Рис. 2.1



**Структурная схема алгоритма метода половинного деления**

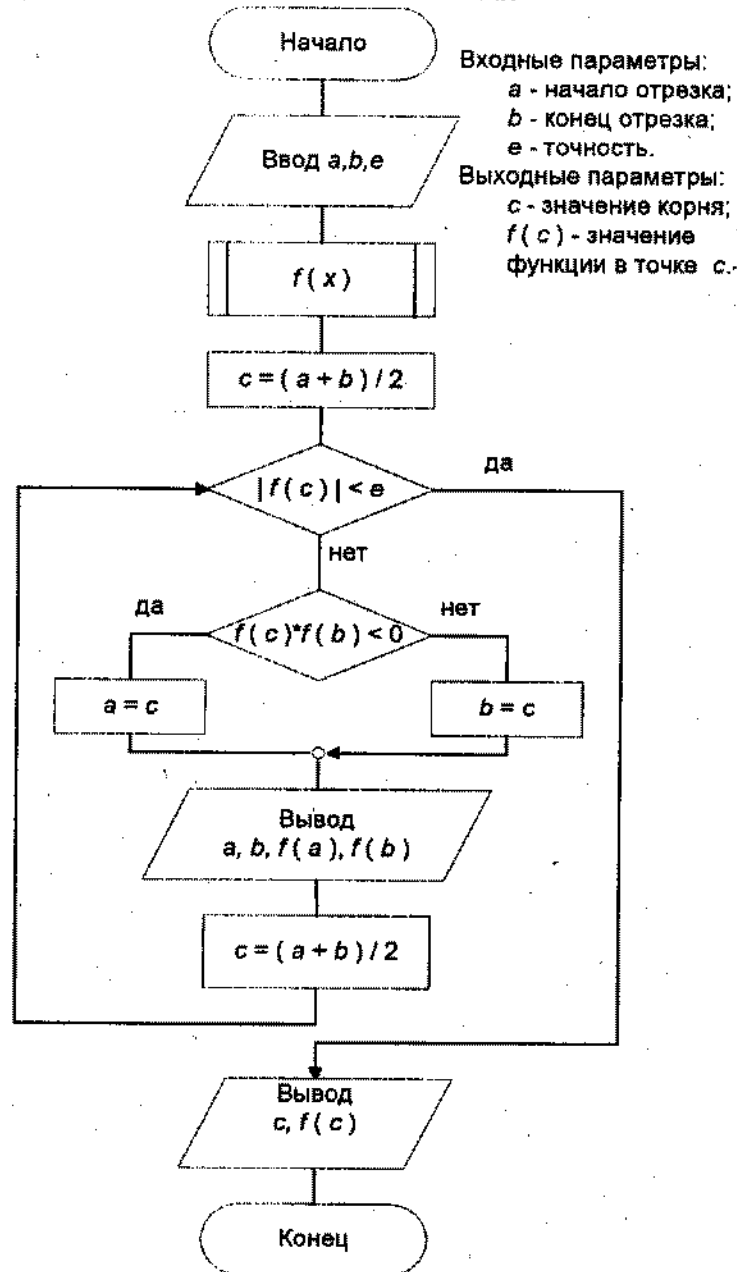


Рис. 2.2

**2.2. Метод итерации**

Пусть дано уравнение  $f(x) = 0$  и отрезок  $[a; b]$ , функция  $f(x)$  удовлетворяет условиям:

- $f(x)$  - дифференцируема на  $[a; b]$ , (5)
- $f(a)f(b) < 0$ , т.е.  $f(x)$  имеет разные знаки на концах отрезка,
- $f'(x) \neq 0$  на  $[a; b]$ .

Заменим уравнение  $f(x) = 0$  равносильным ему уравнением  $x = \varphi(x)$ . (6)

Выбирая произвольное значение  $x_0 \in [a; b]$ , вычислим

$$x_1 = \varphi(x_0),$$

$$x_2 = \varphi(x_1),$$

.....

$$x_{n+1} = \varphi(x_n), \quad n = 0, 1, \dots \quad (7)$$

Процесс (7) сходится к корню, если выполнены условия теоремы о сходимости (достаточные условия сходимости).

**Теорема 3.**

Пусть функция  $\varphi(x)$  определена и дифференцируема на  $[a; b]$ , причем все ее значения  $\varphi(x) \in [a; b]$ . Тогда, если существует правильная дробь  $q$ , такая, что

$$|\varphi'(x)| \leq q < 1 \quad (8)$$

при  $x \in (a; b)$ , то процесс итерации (7) сходится независимо от начального приближения  $x_0 \in [a; b]$  и предельное значение  $x^* = \lim_{n \rightarrow \infty} x_n$  является единственным корнем уравнения  $x = \varphi(x)$  на отрезке  $[a; b]$ .

Приведем один из способов равносильного преобразования уравнения  $f(x) = 0$  к виду  $x = \varphi(x)$ .

Построим функцию

$$\varphi(x) = \lambda \cdot f(x) + x, \quad (9)$$

$$\begin{cases} -\frac{1}{r} < \lambda < 0, & \text{если } f'(x) > 0, \\ 0 < \lambda < \frac{1}{r}, & \text{если } f'(x) < 0, \end{cases} \quad (10)$$

$$r = \max(|f'(a)|, |f'(b)|).$$

Тогда процесс итерации

$$x_{n+1} = \varphi(x_n), \quad n = 0, 1, \dots \quad (11)$$

сходится к корню  $x^*$  при любом  $x_0 \in [a; b]$ .

Скорость сходимости оценивается как

$$|x^* - x_n| \leq \frac{m}{1-q} q^n, \quad (12)$$

$$m = |x_0 - \varphi(x_0)|, \quad q = \lambda \cdot r + 1$$

или

$$q = \max |\varphi'(x)|, \quad x \in [a; b]. \quad (13)$$

Примером условий окончания могут быть

$$\begin{cases} \Delta x_n = |x_n - x_{n-1}| < E, \\ |f(x_n) - f(x_{n-1})| < E, \end{cases} \quad (14)$$

или

$$\frac{1-q}{q} \Delta x_n < E, \quad \text{где } E - \text{требуемая точность.}$$

Геометрически процесс итерации - нахождение абсциссы точки пересечения кривой  $y = \varphi(x)$  и прямой  $y = x$  (рис.3).

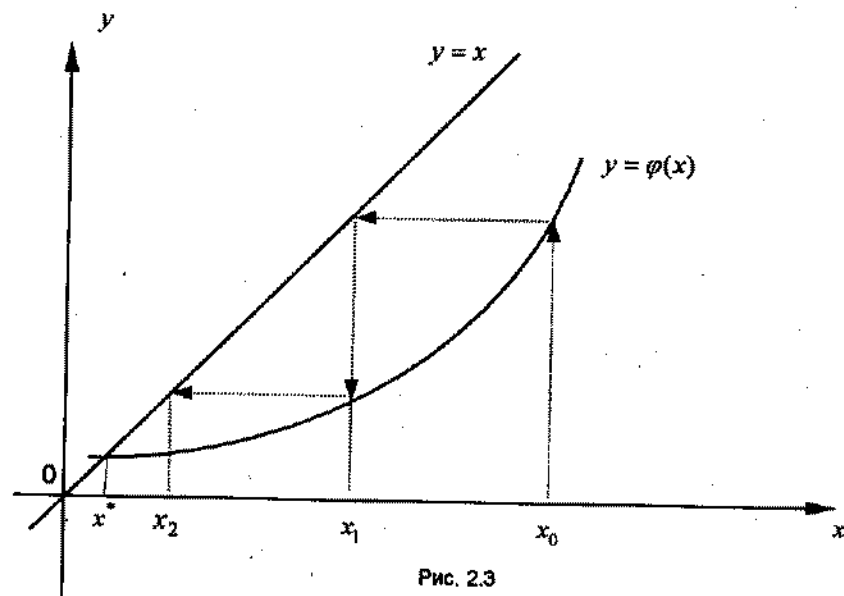


Рис. 2.3

### Пример 3.

Уточнить методом итерации корень уравнения

$$f(x) = 4x - 5 \ln(x) - 5 = 0,$$

принадлежащий отрезку  $[0.25; 0.75]$ .

Решение.

Выберем  $x_0 = 0.5$

Проверим выполнение условий (5).

$$f(a) = 1 - 5 \ln(0.25) - 5 = 1 + 5 \cdot 1.3863 - 5 > 0,$$

$$f(b) = 3 - 5 \ln(0.75) - 5 = 3 + 5(0.2877 - 1) < 0.$$

$$f'(x) = 4 - \frac{5}{x} < 0 \text{ на отрезке } [0.25; 0.75]$$

$$f'(0.25) = -16, \quad f'(0.75) = -\frac{8}{3}.$$

Согласно (10) и (11)  $r = 16$ ,  $0 < \lambda < \frac{1}{16}$ , возьмем  $\lambda = 0.05$ .

Из (9)  $\varphi(x) = 0.2x - 0.25 \ln(x) - 0.25 + x = 1.2x - 0.25(1 + \ln(x))$ .

Итак, рекуррентная формула метода итерации

$$x_{n+1} = 1.2x_n - 0.25(1 + \ln(x_n)), \quad n = 0, 1, \dots$$

Вычислим

$$x_1 = 1.2x_0 - 0.25(1 + \ln(x_0)) = 0.525.$$

Оценим первое приближение, для чего найдем  $m$  и  $q$  (13).

$$\varphi'(x) = 1.2 - \frac{0.25}{x}, \quad \varphi''(x) = \frac{0.25}{x^2} > 0,$$

$\varphi'(x)$  - возрастающая функция;  $\varphi'(0.75) = 0.9$ ,  $q = 0.9$ ,  $m = 0.025$ ,

$$|x^* - x_1| \leq \frac{0.025}{0.1} \cdot 0.9 = 0.225.$$

Процесс следует продолжать до тех пор, пока с заданной точностью не будет найден корень.

### 2.3. Усовершенствованный метод итерации

Перепишем (7) в виде  $x_{n+1} = x_n + \alpha \Delta x$ ,  $\Delta x = f(x_n) - x_n$ .

Полагаем  $x_{n+1} = x_n + \alpha \Delta x_n$ , где  $\alpha$  - искомый коэффициент, улучшающий сходимость итерации.

Геометрически процесс отыскания следующего приближения сводится к проведению хорды между точками  $(x_n; f(x_n))$ ,  $(x_{n+1}; f(x_{n+1}))$  и нахождению абсциссы точки пересечения этой хорды с прямой  $y = x$ .

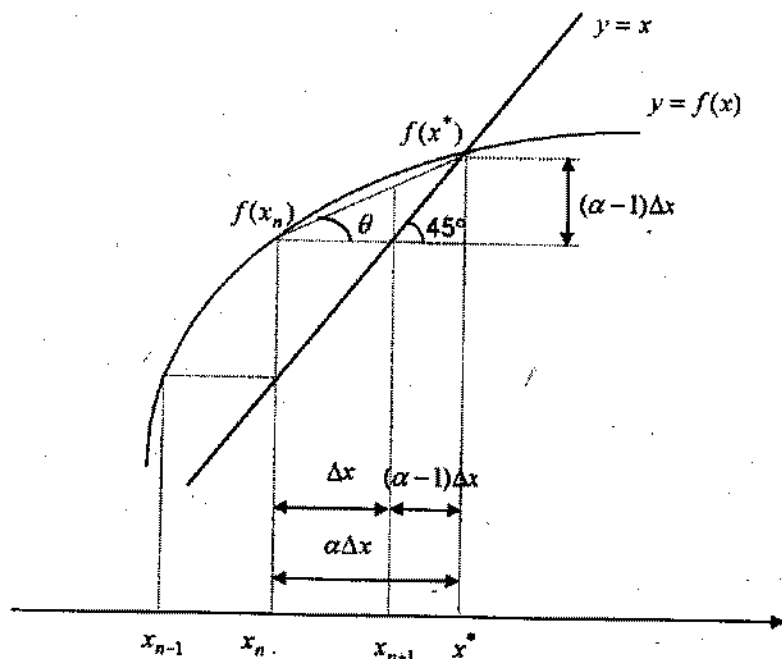


Рис. 2.4

Геометрическая иллюстрация для случая  $0 < f'(x) < 1$ . (рис. 4),

Наилучшим выбором  $\alpha$  будет тот, при котором  $x_{n+1} = x^*$ .

В этом случае итерации определяются по формулам

$$x_{n+1} = x_n + \alpha(f(x_n) - x_n), \quad (15)$$

$$\alpha = \frac{1}{1 - f'(\xi)}, \quad x_n \leq \xi < x^*, \quad (16)$$

Для  $f'(\xi)$  можно принять приближение

$$f'(\xi) = \frac{f(x_n) - x_n}{x_n - x_{n-1}}. \quad (17)$$

Правило выбора  $\alpha$ :

$$\alpha = \begin{cases} \alpha < 0, & \text{если } f'(x) > 1, \\ 0 < \alpha < \frac{1}{2}, & \text{если } f'(x) < -1, \\ \frac{1}{2} < \alpha < 1, & \text{если } -1 < f'(x) < 0, \\ \alpha > 1, & \text{если } 0 < f'(x) < 1. \end{cases} \quad (18)$$

Выбирая  $\alpha$  согласно (18), получим сходящийся процесс итерации. Описанная модификация принадлежит Вегстейну [68].

Итерационные методы обладают свойством самоисправляемости - ошибка на какой-то итерации может быть принята за очередное приближение, что влияет лишь на скорость вычисления корня. Ошибки округления не накапливаются.

## 2.4. Метод Ньютона - Рафсона

Рассмотрим модификация усовершенствованного метода.

Полагаем в (16)  $\xi = x_n$ , тогда из (15) следует

$$x_{n+1} = \frac{f(x_n) - x_n f'(x_n)}{1 - f'(x_n)}, \quad n = 0, 1, \dots \quad (19)$$

Обозначим  $F(x) = f(x) - x = 0$ , получим формулу метода Ньютона - Рафсона

$$x_{n+1} = x_n - \frac{F(x_n)}{F'(x_n)}. \quad (20)$$

Геометрическое толкование (19) (рис. 5.)

1)  $AB$  - касательная к графику  $y = f(x)$  в точке  $(x_n, f(x_n))$ .

2)  $B$  - точка пересечения касательной  $AB$  и прямой  $y = x$ .

3)  $CD$  - касательная, проведенная в точке  $C(x_{n+1}, f(x_{n+1}))$ .

Абсцисса точки пересечения касательной  $CD$  с прямой  $y = x$  - очередное приближение к корню  $x^*$  и т.д.

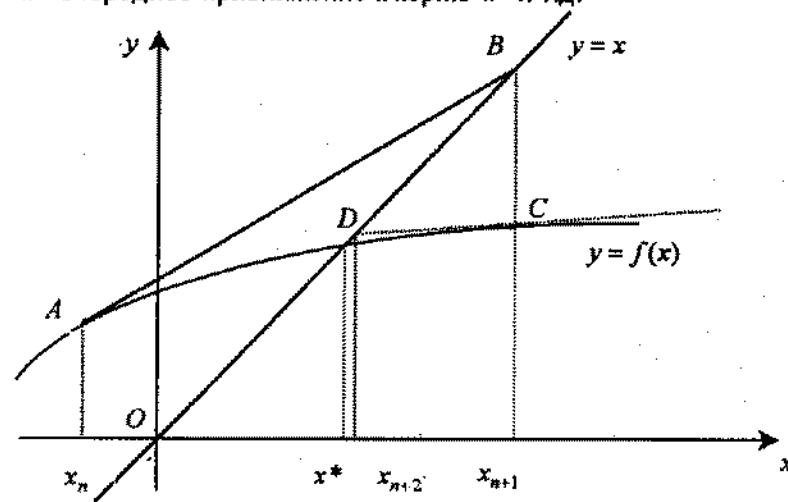


Рис. 2.5

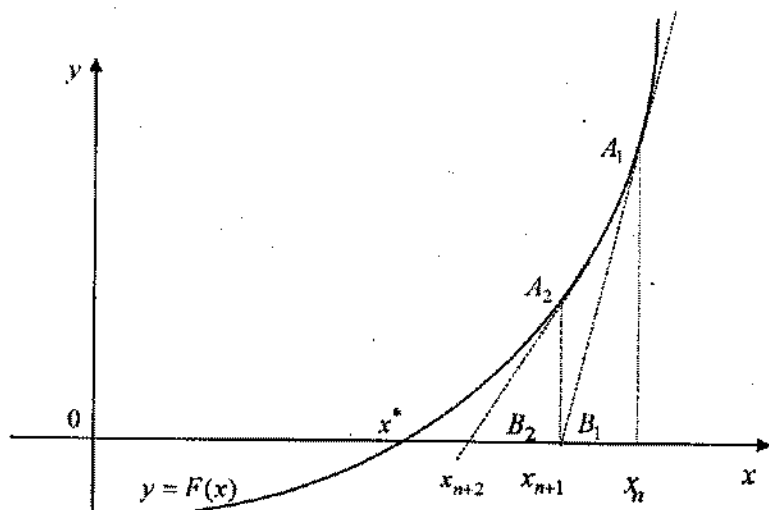


Рис. 2.6

Если уравнение задано в виде  $F(x) = 0$ , то, следуя (20), процесс выглядит так (рис.6).

- 1)  $A_1B_1$  - касательная к графику  $y = F(x)$  в точке  $A_1(x_n, F(x_n))$ .
- 2)  $B_1$  - абсцисса точки пересечения  $A_1B_1$  с осью  $ox$ .
- 3)  $A_2(x_{n+1}, F(x_{n+1}))$
- 4)  $A_2B_2$  - касательная к  $y = F(x)$  в точке  $A_2$  и т.д.

Метод Ньютона - Рафсона имеет большую скорость сходимости. Сходимость метода определена теоремой 4.

**Теорема 4.**

Если  $F(a) \cdot F(b) < 0$ , причем  $F'(x)$  и  $F''(x)$  отличны от нуля и сохраняют определенные знаки при  $a \leq x \leq b$ , то, исходя из начального приближения  $x_0$ , удовлетворяющего условию

$$F(x_0)F''(x) > 0, \quad (21)$$

можно вычислить методом Ньютона - Рафсона единственный корень уравнения  $x^*$  с любой степенью точности.

**Замечания**

- Теорема содержит достаточные условия сходимости.
- Если
  - 1)  $F(x)$  определена и непрерывна при  $-\infty < x < +\infty$ ,
  - 2)  $F(a) \cdot F(b) < 0$ ,

3)  $F'(x) \neq 0$  при  $a \leq x \leq b$ ,

4)  $F''(x)$  существует всюду и сохраняет постоянный знак, то

в качестве  $x_0$  может быть взята любая точка  $[a; b]$ .

• Если кривая  $y = F(x)$  вблизи корня имеет производную, близкую к нулю, метод Ньютона применять не рекомендуется.

Оценка погрешности метода такова:

$$|x^* - x_n| \leq \frac{M_2}{2m_1}(x_n - x_{n-1})^2, \quad (22)$$

где  $M_2$  - наибольшее значение  $|F''(x)|$ ,  $m_1$  - наименьшее значение  $|F'(x)|$  на отрезке  $[a; b]$ .

Если

$$|x_n - x_{n-1}| \leq \sqrt{\frac{2m_1 E}{M_2}}, \quad (23)$$

где  $E$  - требуемая точность, то полагают, что  $x^* = x_n$  с точностью  $E$ .

Остановимся на двух частных случаях для уравнения  $F(x) = 0$ , где  $F''(x)$  непрерывна и сохраняет постоянный знак при  $-\infty < x < +\infty$ .

1. Пусть  $F(x_0)F'(x) < 0$  и  $F(x_0)F''(x) < 0$ .

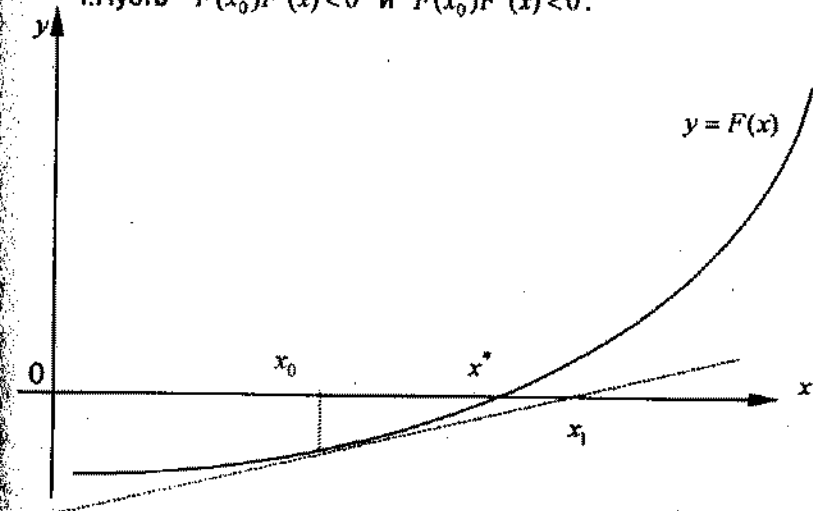


Рис. 2.7

Тогда уравнение имеет единственный корень  $x^* \in (x_0; x_1)$ , который может быть вычислен с заданной точностью методом Ньютона.

тона - Рафсона.  $x_1$  - абсцисса точки пересечения касательной, проведенной к кривой  $y = F(x)$  в точке  $x_0$ , с осью  $ox$ .

II. Пусть  $F'(x_0) = 0$ ,  $F(x_0)F''(x) < 0$ .

В этом случае уравнение  $F(x) = 0$  имеет два почти равных корня  $x_1^*$  и  $x_2^*$ .

Начальные приближения для этих корней

$$\bar{x}_0 = x_0 - \sqrt{\frac{2F(x_0)}{F''(x_0)}}, \quad (24)$$

$$\underline{x}_0 = x_0 + \sqrt{\frac{2F(x_0)}{F''(x_0)}}.$$

Геометрически  $\bar{x}_0$  и  $\underline{x}_0$  суть абсциссы точек пересечения параболы  $Y = F(x_0) + \frac{1}{2}F''(x_0)(x - x_0)^2$  с осью  $ox$ . Парабола - результат разложения левой части уравнения  $F(x) = 0$  по формуле Тейлора с сохранением величин второго порядка [69].

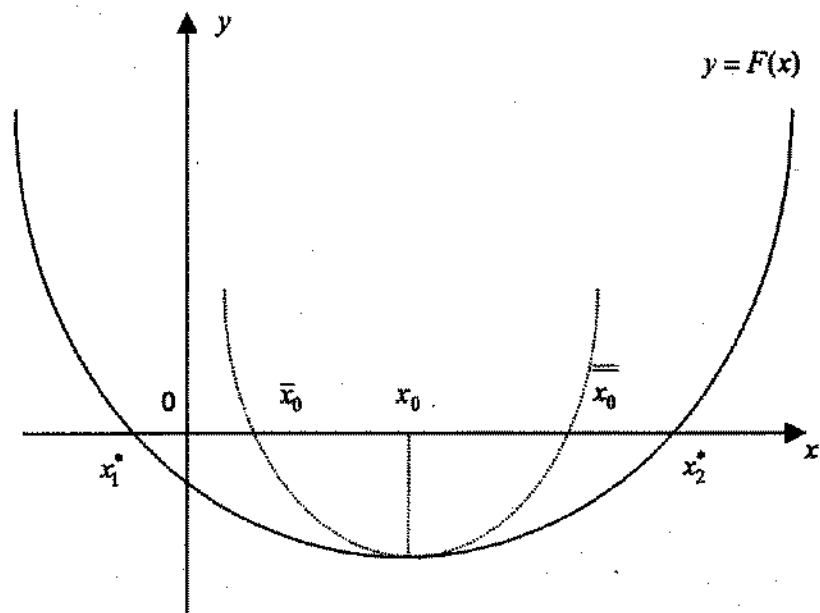


Рис. 2.8

#### Пример 4.

Вычислить положительный корень уравнения  $x^2 = \cos(x)$  и оценить погрешность второго приближения методом Ньютона.

Решение:

Отделение корней можно произвести, например, графически, определив точки пересечения графиков функций  $y_1 = x^2$  и  $y_2 = \cos(x)$ .

Перепишем исходное уравнение в виде

$$f(x) = x^2 - \cos(x) = 0.$$

На отрезке  $[0.5; 1]$   $f(0.5) \cdot f(1) < 0$ .

$$f'(x) = 2x + \sin(x), \text{ при } x \in [0.5; 1] \quad f'(x) > 0,$$

$$f''(x) = 2 + \cos(x) > 0 \text{ при } \forall x.$$

Согласно (21) полагаем  $x_0 = 1$ , т.к.  $f(1) > 0$  и  $f''(x) > 0$ .

По формуле (20)  $x_2 = 0.8242$ ,  $f(0.8242) = -0.1642 \cdot 10^{-3}$ .

Оценим погрешность, для чего вычислим

$$m_1 = \min_{x \in [0.5; 1]} |f'(x)| = 1.4794,$$

$$M_2 = \max_{x \in [0.5; 1]} |f''(x)| = 1.4597,$$

$$|x^* - x_2| \leq \frac{M_2}{2 \cdot m_1} \cdot (x_2 - x_1)^2,$$

$$|x^* - x_2| \leq 0.967 \cdot 10^{-4}.$$

Метод Ньютона критичен к выбору  $x_0$ , однако сходится быстрее других методов. В связи с этим часто используют следующий прием: начинают процесс поиска корня, например, методом половинного деления, а затем уточняют корень быстрым (в окрестности точного значения  $x^*$ ) методом Ньютона.

#### 2.5. Методы решения алгебраических уравнений

Отметим некоторые свойства алгебраических уравнений с действительными коэффициентами вида:

$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = 0. \quad (25)$$

• Уравнение степени  $n$  имеет ровно  $n$  корней (среди них могут быть и комплексные) при условии, что каждый корень считается столько раз, какова его кратность.

• Если коэффициенты уравнения действительные, то комплексные корни этого уравнения попарно комплексно – сопряженные, то есть, если число  $y = a + i \cdot b$  ( $a, b$  – действительные числа) – корень уравнения (25), то число  $\bar{y} = a - i \cdot b$  также является корнем уравнения.

• Алгебраическое уравнение (25) нечетной степени имеет по меньшей мере один действительный корень.

• Если уравнение (25) имеет только действительные корни, то справедливо соотношение  $a_k^2 > a_{k-1} \cdot a_{k+1}$ , ( $k = 1, 2, \dots, n-1$ ) (26)

• Грубая оценка модулей корней уравнения содержится в утверждении

$$\text{Пусть } A = \max\{|a_0|, |a_1|, \dots, |a_{n-1}|\}.$$

Тогда

$$|x_k| < 1 + \frac{A}{|a_n|} = R, \quad k = 1, 2, \dots, n. \quad (27)$$

$$\text{Пусть } a_0 \neq 0, \text{ и } B = \max\{|a_n|, |a_{n-1}|, \dots, |a_1|\}$$

$$|x_k| > \frac{1}{1 + \frac{B}{|a_0|}} = r. \quad (28)$$

Таким образом, все корни уравнения (25) расположены в круговом кольце

$$r < |x_k| < R. \quad (29)$$

Числа  $r$  и  $R$  являются соответственно нижней и верхней границами положительных корней, а  $(-R)$  и  $(-r)$  соответственно нижней и верхней границами отрицательных корней уравнения (25).

• Число положительных корней уравнения с учетом их кратностей равно числу перемен знаков в ряду коэффициентов уравнения (где коэффициенты, равные нулю, не учитываются) или меньше этого числа на четное число.

• Число отрицательных корней уравнения, если коэффициенты отличны от нуля, равно числу постоянств знаков в ряду коэффициентов или меньше этого числа на четное число.

#### Пример 5.

Определить состав и границы модулей корней уравнения

$$x^5 + 4x^4 - 9x^3 - 14x^2 + 50x - 600 = 0.$$

#### Решение

Число положительных корней равно числу перемен знаков в ряду коэффициентов 1, 4, -9, -14, 50, -600 или на четное число меньше, т.е. 1 или 3.

Число отрицательных корней 2 или 0.

Так как  $50^2 < (-14) \cdot (-600)$ , то уравнение имеет по крайней мере одну пару комплексных корней.

Согласно формулам (27) и (28) грубая оценка корней:

$$A = 600, \quad B = 50,$$

$$\frac{1}{1 + \frac{50}{600}} < |x_k| < 1 + \frac{600}{1},$$

$$0.9 < |x_k| < 601.$$

#### Замечание

Данное уравнение имеет корни:

$$3.3525; (-4.4784 - 1.327i); (-4.4784 + 1.327i); (0.8021 - 2.749i); (0.8021 + 2.749i).$$

Одним из способов решения алгебраических уравнений является метод понижения порядка. Он состоит в том, что после нахождения какого-либо корня  $c$  данное уравнение можно разделить на  $(x - c)$ , тем самым понизить его степень.

Для уменьшения погрешности лучше сначала находить меньшие по модулю корни уравнения и удалять их (понижать степень полинома). Все вышеуказанные методы решения трансцендентных уравнений могут быть использованы и для решения алгебраических уравнений.

## 2.6. Определение комплексных корней алгебраических уравнений

Рассмотрим методы их определения: метод Лина и метод Берстоу.

Согласно свойству 2 (п. 5) комплексные корни попарно сопряжены, поэтому при их исключении порядок уравнения уменьшается на 2.

Запишем уравнение

$$x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0 = 0 \quad (30)$$

в виде

$$(x^2 + px + q) \cdot (x^{n-2} + b_{n-1}x^{n-3} + \dots + b_3x + b_2) + b_1x + b_0 = 0. \quad (31)$$

Линейный остаточный член  $b_1x + b_0$  стремится к нулю при значении  $x$ , равном корню. Из сравнения коэффициентов при одинаковых степенях  $x$ , при условии  $b_0 = b_1 = 0$ , получаем:

$$\begin{cases} b_n = a_n = 1, \\ b_{n-1} = a_{n-1} - p, \\ b_{n-2} = a_{n-2} - p \cdot b_{n-1} - q, \\ \dots \\ b_{n-j} = a_{n-j} - p \cdot b_{n+1-j} - q \cdot b_{n+2-j}, \quad j = \overline{3, n-2}, \\ \dots \\ p = \frac{a_1 - q \cdot b_2}{b_2}, \\ q = \frac{a_0}{b_2}. \end{cases} \quad (32)$$

Задавая начальные значения  $p$  и  $q$ , последовательно вычисляем коэффициенты  $b_{n-1}, b_{n-2}, \dots$ . Из двух последних соотношений (32) найдем уточненные значения  $p$  и  $q$ , обозначим их  $p^*$  и  $q^*$ .

Процесс повторяется до тех пор, пока с заданной степенью точности не будут вычислены  $p$  и  $q$ , т.е. задача сводится к решению системы двух уравнений с двумя неизвестными методом простой итерации (методом Якоби).

#### Пример 6.

Найти методом Лина корни алгебраического уравнения

$$x^4 - 9x^3 - 14x^2 + 50x - 600 = 0.$$

**Решение.**

Применяя изложенный алгоритм, получим решение:

- комплексные корни  $(1.455351 + 3.346963 i)$ ;  $(1.455351 - 3.346963 i)$  попарно сопряжены,

-  $10.41447$ ;  $-4.32517$  - действительные корни.

Уравнение четвертой степени имеет четыре корня.

Метод Берстоу – обобщение метода Ньютона - Рафсона для нахождения комплексных корней уравнения.

Итерация Берстоу заключается в следующем.

Для коэффициентов квадратичного множителя  $x^2 + px + q$  выбирают начальные приближения. Затем эти значения уточняют с помощью корректировочных коэффициентов:

$$\begin{cases} p = p + \Delta p, \\ q = q + \Delta q. \end{cases} \quad (33)$$

Предположим, что  $b_1 = R(p, q)$ ,  $b_0 = S(p, q)$ , разложение этих функций в ряд Тэйлора:

$$\begin{cases} S(p + \Delta p, q + \Delta q) = S(p, q) + \frac{\partial S}{\partial p} \Delta p + \frac{\partial S}{\partial q} \Delta q + \dots, \\ R(p + \Delta p, q + \Delta q) = R(p, q) + \frac{\partial R}{\partial p} \Delta p + \frac{\partial R}{\partial q} \Delta q + \dots \end{cases} \quad (34)$$

Далее следует:

$$\begin{cases} S(p, q) + \frac{\partial S}{\partial p} \Delta p + \frac{\partial S}{\partial q} \Delta q = 0, \\ R(p, q) + \frac{\partial R}{\partial p} \Delta p + \frac{\partial R}{\partial q} \Delta q = 0. \end{cases}$$

Решая систему двух уравнений относительно  $\Delta p$  и  $\Delta q$ , получим:

$$\begin{cases} \Delta p = \frac{S \frac{\partial R}{\partial q} - R \frac{\partial S}{\partial q}}{\frac{\partial R}{\partial p} \frac{\partial S}{\partial q} - \frac{\partial S}{\partial p} \frac{\partial R}{\partial q}}, \\ \Delta q = \frac{R \frac{\partial S}{\partial p} - S \frac{\partial R}{\partial p}}{\frac{\partial R}{\partial p} \frac{\partial S}{\partial q} - \frac{\partial S}{\partial p} \frac{\partial R}{\partial q}}. \end{cases} \quad (35)$$

Используя соотношения (32), вычислим производные (36).

Задавая начальные приближения  $p_0$  и  $q_0$ , ищем решение системы (35).

Последовательные приближения вычисляются до тех пор, пока  $|\Delta p|$  и  $|\Delta q|$  не станут равными нулю с заданной степенью точности. Уточненные значения  $p$  и  $q$  - искомые коэффициенты трехчлена.

Метод Берстоу сходится быстрее, чем метод Лина, иногда дает сходимость даже в тех случаях, когда метод Лина расходится, однако критичен к выбору начальных приближений.

$$\frac{\partial b_n}{\partial p} = c_n = 0,$$

$$\frac{\partial b_{n-1}}{\partial p} = c_{n-1} = -1,$$

$$\frac{\partial b_{n-2}}{\partial p} = c_{n-2} = -b_{n-1} - p \cdot \frac{\partial b_{n-1}}{\partial p},$$

$$\dots$$

$$\frac{\partial b_{n-j}}{\partial p} = c_{n-j} = -b_{n-j+1} - p \cdot c_{n-j+1} - q \cdot c_{n-j+2},$$

$$\frac{\partial R}{\partial p} = \frac{\partial b_1}{\partial p} = c_1,$$

$$\frac{\partial S}{\partial p} = \frac{\partial b_0}{\partial p} = c_0,$$

$$\frac{\partial b_n}{\partial q} = d_n = 0,$$

$$\frac{\partial b_{n-1}}{\partial q} = d_{n-1} = 0,$$

$$\frac{\partial b_{n-2}}{\partial q} = d_{n-2} = -1,$$

$$\dots$$

$$\frac{\partial b_{n-j}}{\partial q} = d_{n-j} = -b_{n-j+2} - p \cdot d_{n-j+1} - q \cdot d_{n-j+2},$$

$$j = 3, n-2, \quad (36)$$

$$\frac{\partial R}{\partial q} = \frac{\partial b_1}{\partial q} = d_1,$$

$$\frac{\partial S}{\partial q} = \frac{\partial b_0}{\partial q} = d_0.$$

**Пример 7.**

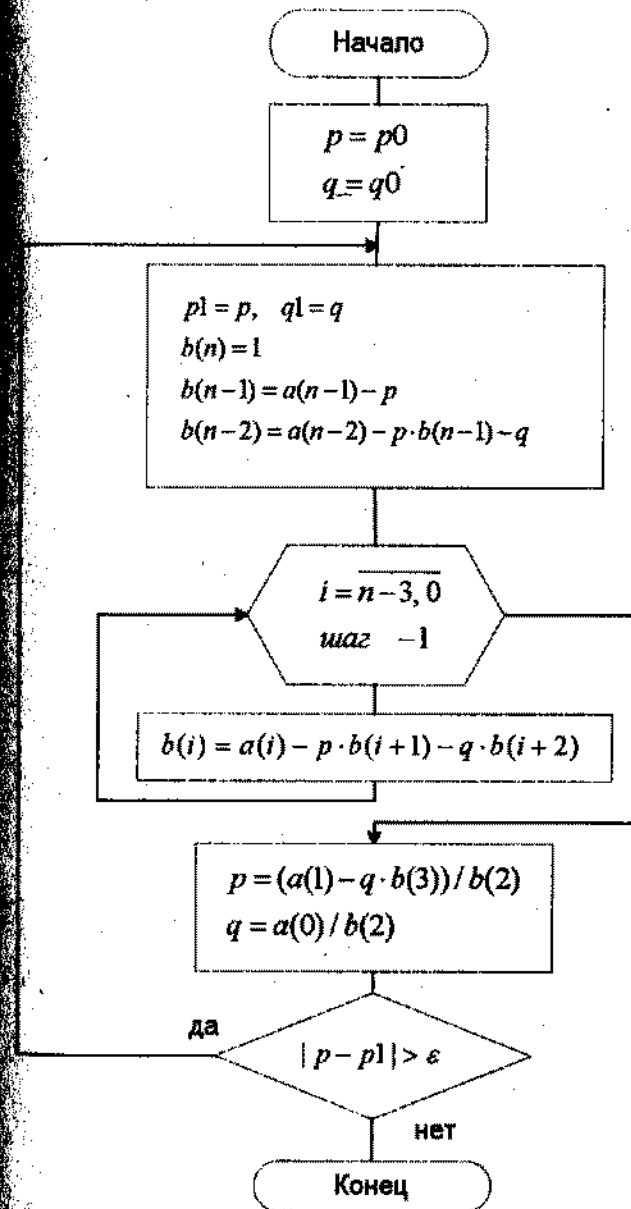
Найти корни уравнения  $x^5 + 4x^4 - 9x^3 - 14x^2 + 50x - 600 = 0$  методом Берстоу.

Решение.

$(-4.478485 + 1.327003 i); (-4.478485 - 1.327003 i); 3.352574;$

$(0.8021978 + 2.749415 i); (0.8021978 - 2.749415 i).$

**Структурная схема алгоритма метода Лина**



Входные параметры:  
 n – степень полинома;  
 a() – массив коэффициентов исходного полинома.  
 Выходные параметры:  
 b() – массив коэффициентов исходного полинома степени (n-1);  
 p, q – коэффициенты квадратного трехчлена.

Рис. 2.9



**Структурная схема алгоритма нахождения всех корней полинома  $n$ -ой степени**

Входные параметры:  
 $n$  – порядок полинома;  
 $a()$  – массив коэффициентов.  
 Выходные параметры:  
 $r1$  – массив действительных частей;

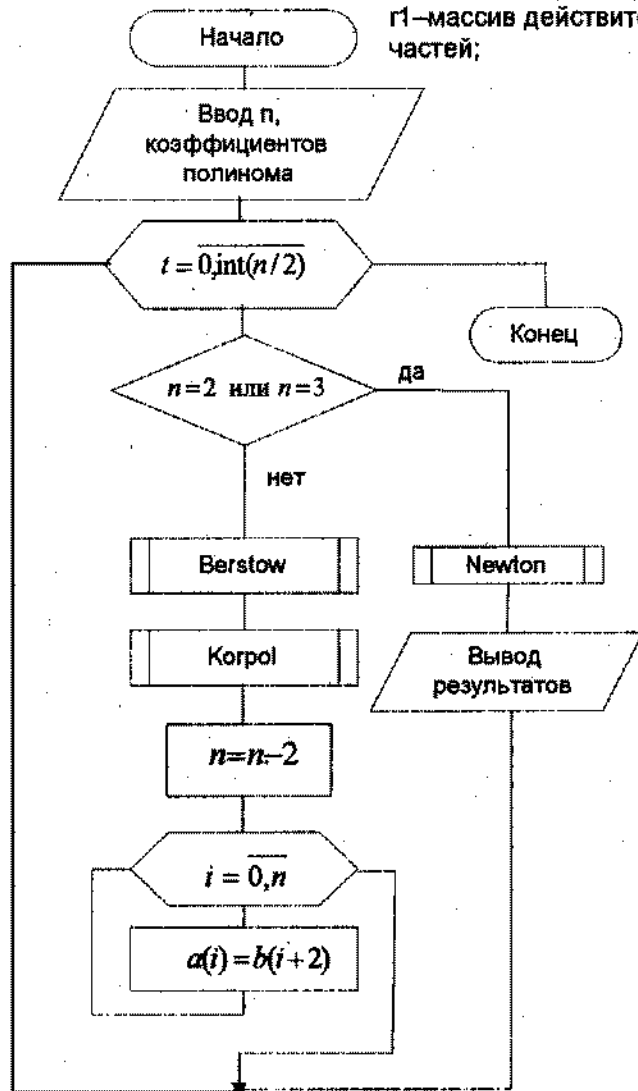
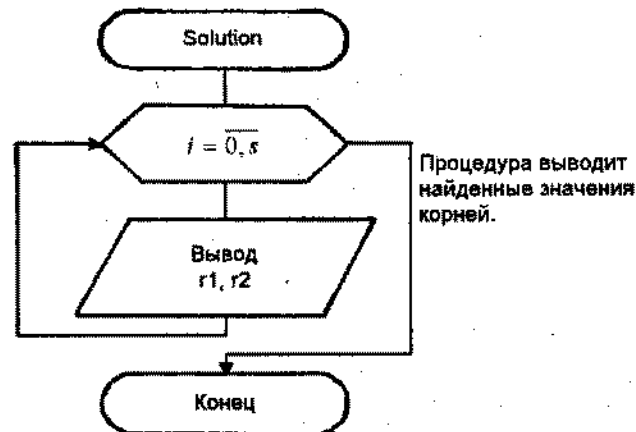


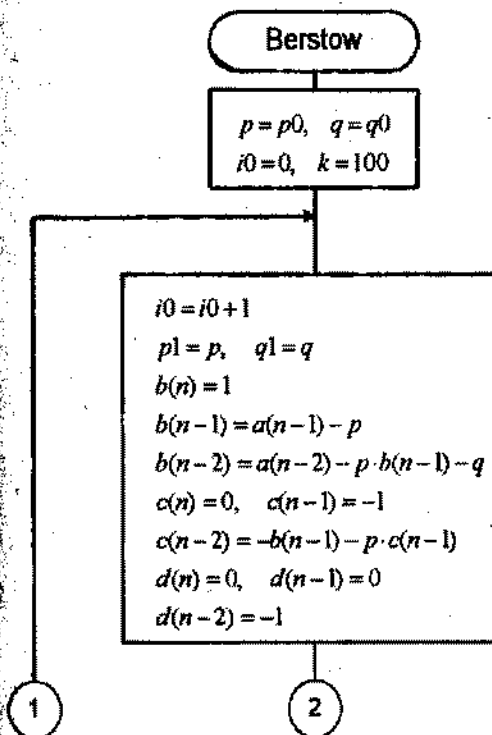
Рис. 2.9



Процедура выводит найденные значения корней.

Рис. 2.10

**Структурная схема метода Берстоу**



Процедура находит коэффициенты квадратного трехчлена ( $p$  и  $q$ ), используя итерационные формулы метода Берстоу.  
 Входные параметры:  
 $n$  – степень исходного полинома,  
 $a(n)$  – массив коэффициентов полинома,  
 $e$  – точность,  
 Выходные параметры:  
 $b(n)$  – массив новых коэффициентов,  
 $p, q$  – значения коэффициентов.

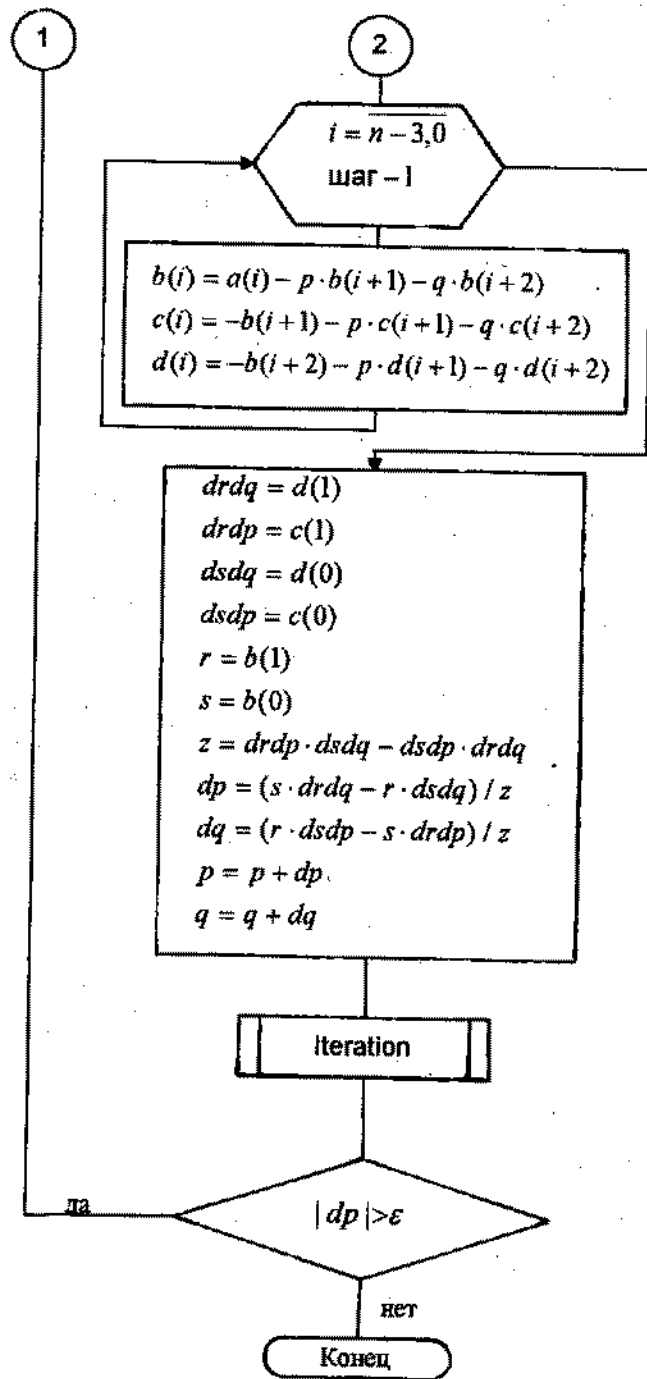


Рис. 2.11



Рис. 2.12

**Процедура вычисления корня полинома степени  $n \leq 3$  методом Ньютона**

Входные параметры:  
 $a()$  – массив коэффициентов полинома третьей степени.  
 Выходные параметры:  
 $s$  – число найденных корней;  
 $r1()$  – массив действительных частей корней;  
 $r2()$  – массив мнимых частей корней;  
 $p, q$  – коэффициенты полинома второй степени.

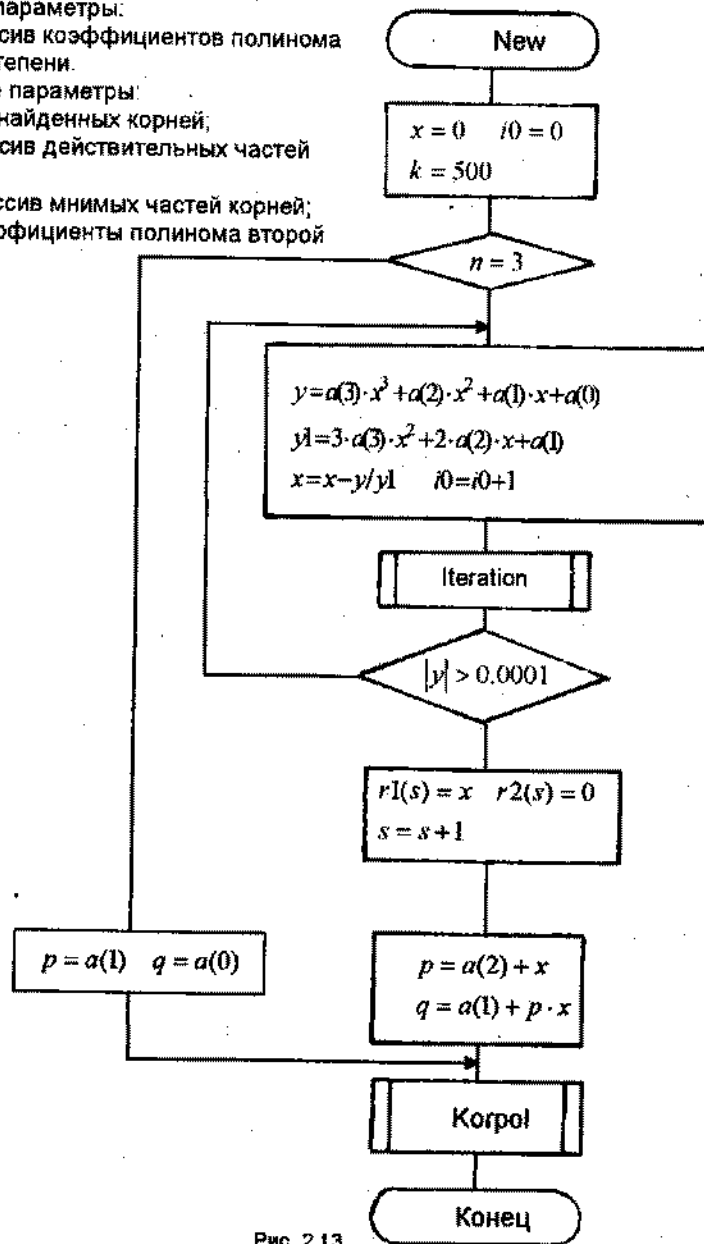


Рис. 2.13

**2.7. Метод парабол**

При решении уравнения  $f(x) = 0$  функцию  $f(x)$  аппроксимируют интерполяционным многочленом Ньютона второй степени, построенным для разделенных разностей (раздел 4.5) по точкам  $x_{k-2}, x_{k-1}, x_k$  [12]

$$P_2(t) = at^2 + bt + c, \quad (37)$$

где

$$a = \frac{f(x_k)(x_{k-1} - x_{k-2}) + f(x_{k-1})(x_{k-2} - x_k) + f(x_{k-2})(x_k - x_{k-1})}{(x_k - x_{k-1})(x_{k-1} - x_{k-2})(x_k - x_{k-2})},$$

$$b = a \cdot (x_k - x_{k-1}) + \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}, \quad (38)$$

$$c = f(x_k).$$

Меньший по модулю корень многочлена (37) ( $t_1$ ) определяет новое приближение к искомому корню уравнения  $f(x)$ .

$$x_{k+1} = x_k + t_1. \quad (39)$$

Из множества точек  $x_{k-2}, x_{k-1}, x_k$  выбирают три точки, отбрасывая точку, отвечающую наибольшему значению  $f(x)$ .

Очередное приближение ищется вновь построением многочлена  $P_2$ .

Оптимальное число итераций определяется в соответствии с приемом Гарвика. Выбирают некоторую величину  $\delta$  ( $\delta = m\varepsilon$ ,  $m \approx 2+5$ ,  $\varepsilon$  - точность вычисления корней  $f(x)$ ). Итерации ведут до выполнения условия  $|x_{k+1} - x_k| < \delta$ . Затем расчет продолжается до тех пор, пока  $|x_{k+1} - x_k|$  убывает. Первое же возрастание этой разности означает конец итерациям, причем последняя итерация не учитывается.

Сходимость метода парабол при произвольном начальном приближении не доказана, однако практика показывает, что всегда метод сходится к какому-либо корню уравнения, в том числе и к комплексному корню (если он существует) при действительном приближении. Это свойство – важное преимущество метода парабол перед другими методами.

Метод парабол является эффективным методом для вычисления корней многочлена степени  $n$ .

Вычислив один из корней ( $x_1$ ), его следует удалить и получить новое уравнение  $f_1(x) = 0$ . Удаление вычисленного корня про-

изводится делением исходного (результата предыдущего шага) на двучлен  $(x - x_1)$ , т.е.

$$f_1(x) = \frac{f(x)}{x - x_1}$$

Рекомендуется удаление корней начинать с меньшего, т.к. в этом случае погрешность округления при вычислении коэффициентов многочлена  $f_1(x)$  меньше.

Начальные приближения можно положить равными  $x_0 = -1$ ,  $x_1 = 0$ ,  $x_2 = 1$ . В этом случае итерации сходятся к меньшему корню.

Укажем способ деления целой рациональной функции

$$f(x) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n$$

на двучлен  $(x - x_0)$  (схема Горнера).

$$\frac{f(x)}{x - x_0} = b_0x^{n-1} + b_1x^{n-2} + \dots + b_{n-2}x + b_{n-1} + \frac{r}{x - x_0} \quad (40)$$

Если  $x_0$  - корень многочлена  $f(x)$ , то остаток  $r = 0$ .

Коэффициенты  $b_0, b_1, \dots, b_{n-1}$  (частного от деления) могут быть вычислены так:

$$\begin{aligned} b_0 &= a_0, \\ b_1 &= b_0x_0 + a_1, \\ b_2 &= b_1x_0 + a_2, \\ &\dots \end{aligned} \quad (41)$$

$$\begin{aligned} b_{n-1} &= b_{n-2}x_0 + a_{n-1}, \\ \text{Остаток } r &= b_{n-1}x_0 + a_n \end{aligned} \quad (42)$$

Структурная схема алгоритма метода парабол для вычисления корней многочлена приведена на рис. 14 - 18.

**Пример 8.**

Найти методом парабол положительный корень уравнения  $f(x) = x^2 - \cos(x) = 0$  (три итерации).

*Решение.*

Выбираем начальные точки  $x_0 = 0.5$ ,  $x_1 = 0.75$ ,  $x_2 = 1$ .

Пользуясь формулами (37) - (39) на первой итерации получим многочлен  $P_2(t) = 1.3648t^2 + 2.8568t + 0.4597$ .

Структурная схема алгоритма метода парабол

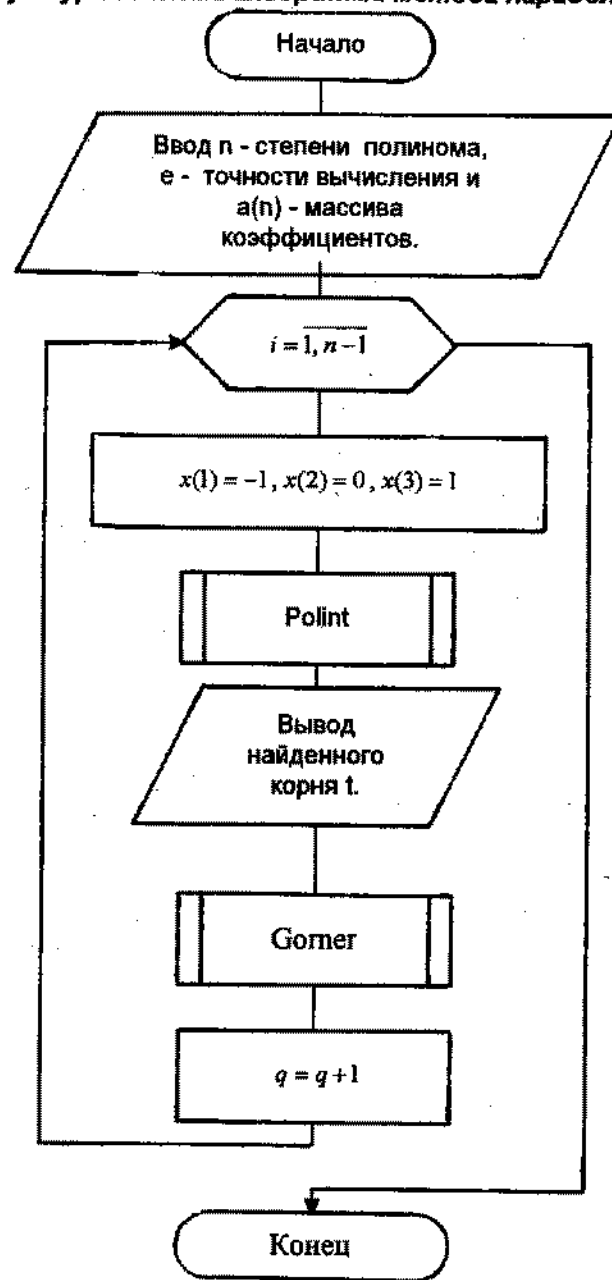
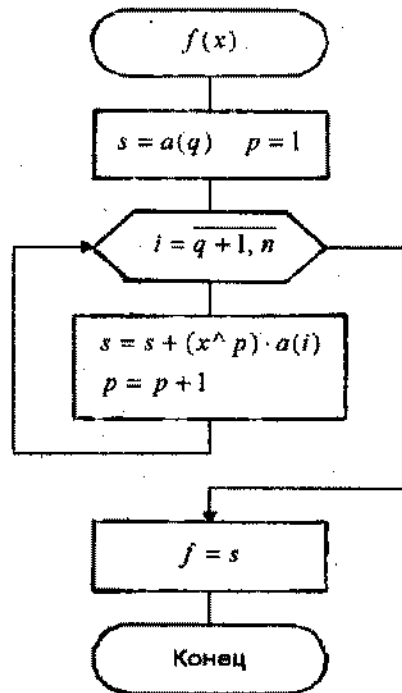
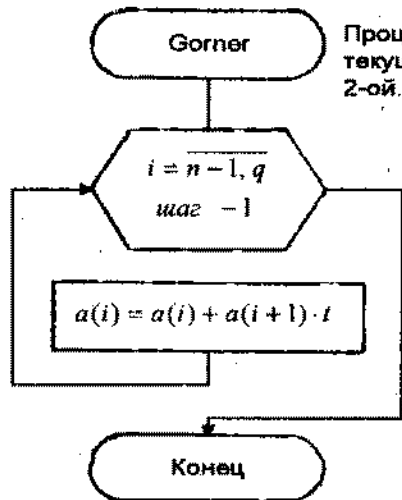


Рис. 2.14



Процедура вычисляет значение полинома в точке  $x$ .  
 Входные параметры:  
 $x$  - значение аргумента;  
 $a()$  - массив коэффициентов;  
 $q$  - число найденных корней.  
 Выходные параметры:  
 $f$  - значение функции.



Процедура вычисляет коэффициенты текущего полинома степени от  $(n-1)$  до 2-ой.  
 Входные параметры:  
 $a()$  - массив коэффициентов;  
 $t$  - найденный корень;  
 $n$  - степень исходного полинома;  
 $q$  - количество найденных корней.  
 Выходные параметры:  
 $a()$  - массив коэффициентов текущего полинома.

Рис. 2.15

### Процедура нахождения наименьшего по модулю корня квадратного уравнения

Входные параметры:  
 $a, b, c$  - коэффициенты квадратного уравнения;  
 $x()$  - массив узлов интерполяции.  
 Выходные параметры:  
 $t$  - найденный корень;  
 Flag - признак наличия комплексного корня.

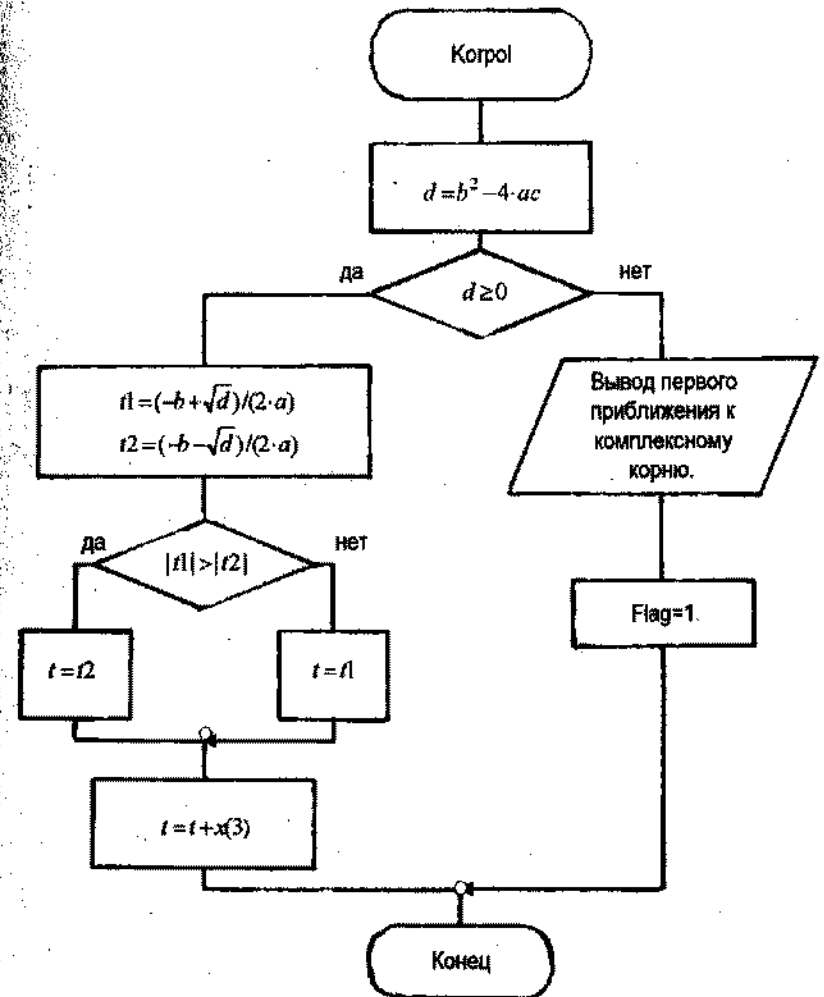


Рис. 2.16

**Процедура определения коэффициентов  
интерполирующего полинома**

Входные параметры:

$x()$  – массив узлов интерполяции;  
 $e$  – точность.

Выходные параметры:

$a, b, c$  – коэффициенты полинома;  
 $t$  – корень.

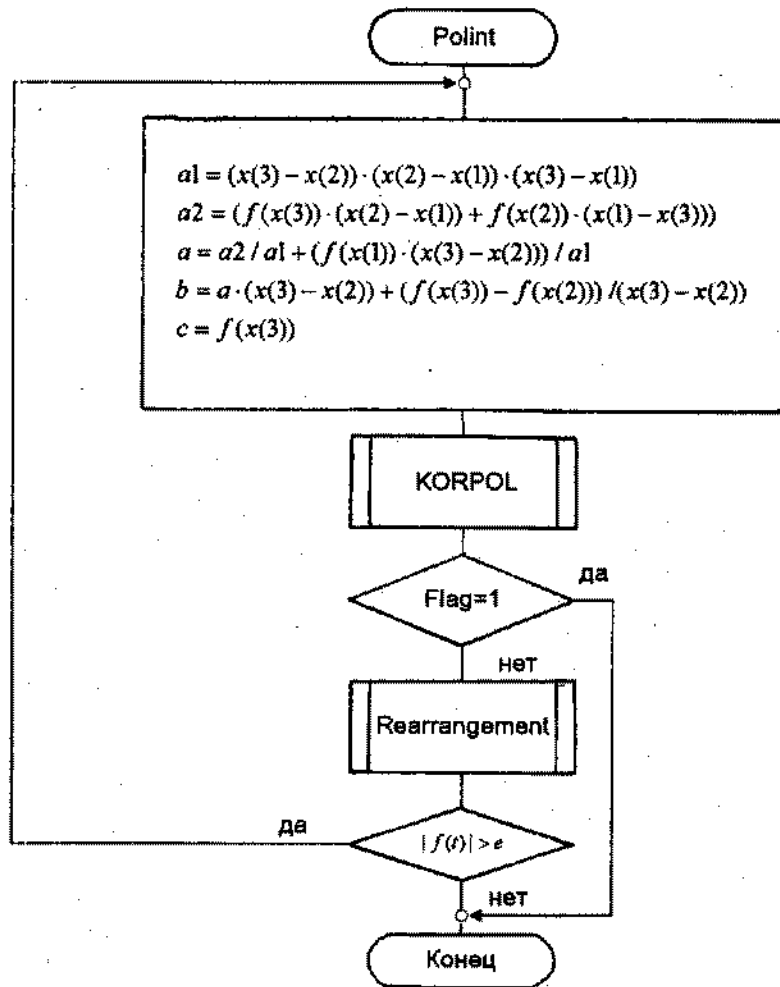
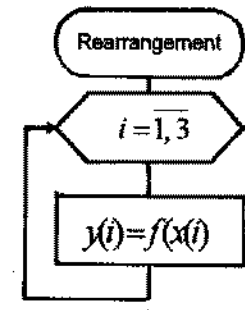


Рис. 2.17



Процедура производит  
выбор точек для  
последующей  
интерполяции

Входные параметры:  
 $x()$  - массив узлов  
интерполяции;  
 $t$  - корень.

Выходные параметры:  
 $x()$  - сформированный  
массив узлов.

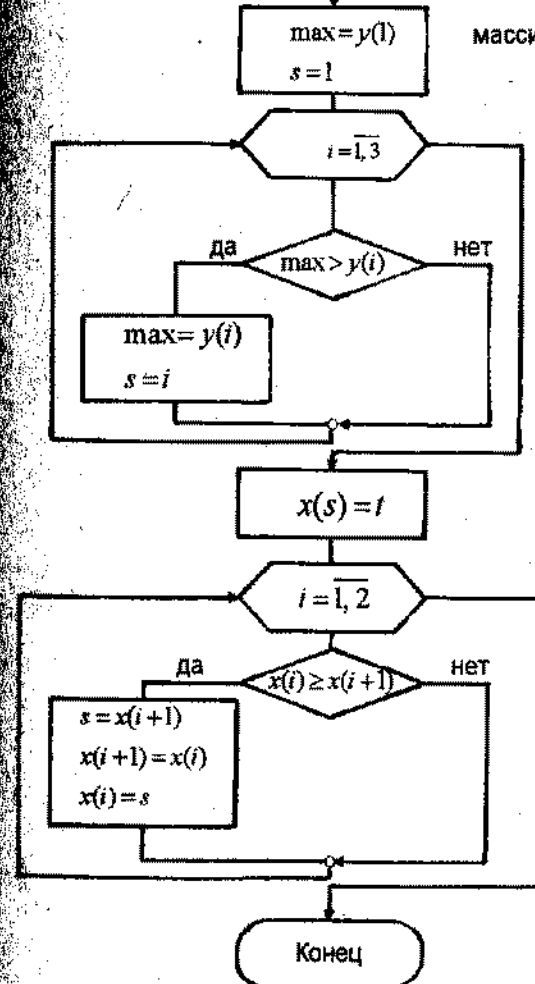


Рис. 2.18

Наименьший корень многочлена  $t_1 = -0.1728$ .

Тогда  $x_{k+1} = x_k + t_1 = 0.8272$ ,  $f(x_{k+1}) = 0.00736$ .

n	x	f(x)	t	x <sub>k+1</sub>	x <sub>k+1</sub> - x <sub>k</sub>
1.	0.5	-0.6276	-0.1728	0.8272	0.1728
	0.75	-0.1692			
	1	0.4597			
2.	0.5	-0.6276	-0.003	0.8242	0.003
	0.75	-0.1692			
	0.8272	0.00736			
3.	0.5	-0.6275	0.0001	0.8243	0.0001
	0.75	-0.1692			
	0.8242	0.0002			

### 3. СИСТЕМЫ НЕЛИНЕЙНЫХ УРАВНЕНИЙ

#### 3.1. Метод Ньютона

К решению систем нелинейных уравнений сводятся многие практические задачи. В отличие от систем линейных уравнений не существует прямых методов решения нелинейных систем. Общий метод решения системы уравнений должен состоять из двух этапов: предварительного определения расположения корней и дальнейшего уточнения решения. Для получения решения обычно используются итерационные методы.

Рассмотрим систему нелинейных уравнений:

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0, \\ f_2(x_1, x_2, \dots, x_n) = 0, \\ \dots \dots \dots \dots \dots \dots \\ f_n(x_1, x_2, \dots, x_n) = 0, \end{cases} \quad (1)$$

В матричном виде эта система имеет вид:

$$\bar{f}(\bar{x}) = 0 \quad (2)$$

где

$$\bar{f} = \begin{pmatrix} f_1 \\ f_2 \\ \dots \\ f_n \end{pmatrix}, \quad \bar{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}.$$

В предположении, что найдено  $k$ -е приближение к решению  $\bar{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$ , можно записать точное решение  $\bar{x}^*$  системы (1) в виде

$$\bar{x}^* = \bar{x}^{(k)} + \Delta \bar{x}^{(k)}, \quad (3)$$

$\Delta \bar{x}^{(k)} = (\Delta x_1^{(k)}, \Delta x_2^{(k)}, \dots, \Delta x_n^{(k)})$  - поправки к искомому решению.

Раскладывая левую часть

$$\bar{f}(\bar{x}^{(k)} + \Delta \bar{x}^{(k)}) = 0 \quad (4)$$

в ряд Тейлора, ограничиваясь линейными членами разложения, получаем рекуррентную формулу для вычисления поправок, а тем самым и формулу итерационного процесса определения решения системы (1) либо (2) методом Ньютона.

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} - W^{-1}(\bar{x}^{(k)}) \bar{f}(\bar{x}^{(k)}) = 0, \quad k = 0, 1, \dots \quad (5)$$

$$W(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \dots & \frac{\partial f_n}{\partial x_n} \end{pmatrix} \quad (6)$$

$W(x)$  - матрица Якоби,  $W^{-1}(x)$  - матрица, обратная матрице Якоби.

Процесс итерации (5) продолжается до тех пор, пока справедливо неравенство

$$|\Delta x_i^{(k)}| \geq E, \quad i = \overline{1, n}, \quad k = 0, 1, \dots, \quad (7)$$

либо число итераций  $S \leq M$ ;  $E$  - требуемая точность,  $M$  - максимальное число итераций.  $M \approx 10$  для расходящегося процесса.

Для системы двух уравнений

$$\begin{cases} f_1(x, y) = 0, \\ f_2(x, y) = 0 \end{cases}$$

формулу (5) можно записать так:

$$\begin{aligned} x^{(k+1)} &= x^{(k)} - \frac{1}{V(x^{(k)}, y^{(k)})} (f_1(x^{(k)}, y^{(k)}) \frac{\partial f_2}{\partial y}(x^{(k)}, y^{(k)}) - f_2(x^{(k)}, y^{(k)}) \frac{\partial f_1}{\partial y}(x^{(k)}, y^{(k)})), \\ y^{(k+1)} &= y^{(k)} - \frac{1}{V(x^{(k)}, y^{(k)})} (f_2(x^{(k)}, y^{(k)}) \frac{\partial f_1}{\partial x}(x^{(k)}, y^{(k)}) - f_1(x^{(k)}, y^{(k)}) \frac{\partial f_2}{\partial x}(x^{(k)}, y^{(k)})), \\ V(x^{(k)}, y^{(k)}) &= \frac{\partial f_1}{\partial x}(x^{(k)}, y^{(k)}) \frac{\partial f_2}{\partial y}(x^{(k)}, y^{(k)}) - \frac{\partial f_1}{\partial y}(x^{(k)}, y^{(k)}) \frac{\partial f_2}{\partial x}(x^{(k)}, y^{(k)}). \end{aligned} \quad (8)$$

### Пример 1.

Решить систему уравнений

$$\begin{cases} \sin(2x - y) - 1.2x = 0.4, \\ 0.8x^2 + 1.5y^2 = 1 \end{cases}$$

методом Ньютона.

Решение.

Отделение корней проведем графически.

$$\begin{cases} y_1 = 2x - \arcsin(1.2x + 0.4), \\ y_2 = \pm \sqrt{\frac{1 - 0.8x^2}{1.5}} \end{cases}$$

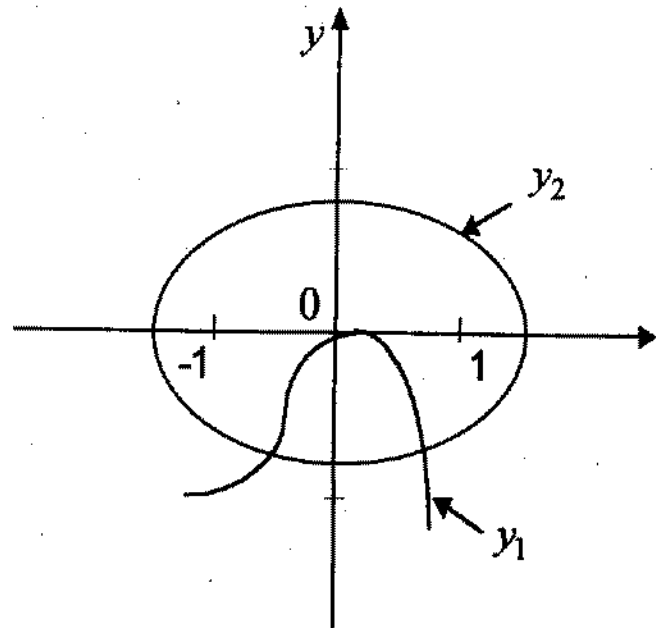


Рис. 3.1

Значения  $x$  из уравнений (с учетом ОДЗ)

$$\begin{cases} -1 \leq 1.2x + 0.4 \leq 1, \\ -\sqrt{1.25} \leq x \leq \sqrt{1.25}, \end{cases}$$

откуда  $-1.12 \leq x \leq 0.5$ .

Система имеет два решения. Уточним решение, принадлежащее области  $D$ .

$$D\{0.4 < x < 0.5; -0.76 < y < -0.73\}.$$

Выберем  $x_0 = 0.4, y_0 = -0.75$ , обозначим

$$\begin{cases} f_1 = \sin(2x - y) - 1.2x - 0.4, \\ f_2 = 0.8x^2 + 1.5y^2 - 1. \end{cases}$$

$$\frac{\partial f_1}{\partial x} = 2 \cos(2x - y) - 1.2; \quad \frac{\partial f_2}{\partial x} = 1.6x,$$

$$\frac{\partial f_1}{\partial y} = -\cos(2x - y); \quad \frac{\partial f_2}{\partial y} = 3y.$$

Матрица Якоби

$$W(x, y) = \begin{pmatrix} 2 \cos(2x - y) - 1.2 & -\cos(2x - y) \\ 1.6x & 3y \end{pmatrix}.$$



Вычисленные последовательные приближения по формулам (5), (6) сведены в таблицу:

Таблица 1

$k$	$x_k$	$\Delta x_k$	$y_k$	$\Delta y_k$
0	0.4	0.1	-0.75	0.017
1	0.5	-0.012	-0.7332	-2E-4
2	0.4914	-0.0001	-0.7334	-1E-5
3	0.4913		-0.7334	

Получаем решение системы:

$$\begin{cases} x^* = 0.491, \\ y^* = -0.733. \end{cases}$$

### Замечания

• Метод Ньютона эффективен только при достаточной близости начального приближения к решению системы. Требования к сходимости метода достаточно жесткие, теоремы о скорости сходимости, устойчивости приведены в [23]. Практически метод Ньютона применяется для уточнения решения, полученного каким-либо другим методом.

• Матрица Якоби содержит производные. Поскольку аналитическое дифференцирование в общем случае нежелательно, частные производные заменяют их приближенными конечно-разностными значениями.

$$\frac{\partial f_j}{\partial x_i} = \frac{f_j(x_1, x_2, \dots, x_{i-1}, x_i + H, x_{i+1}, \dots, x_n) - f_j(x_1, x_2, \dots, x_n)}{H}, \quad (9)$$

$i = \overline{1, n}$ ,  $H_i = E \cdot |x_i|$ . Пусть  $H = \max_i H_i = \text{const}$ .

Структурная схема алгоритма метода Ньютона приведена на рис. 2 - 4

### 3.2. Метод итерации

В предположении, что вектор-функция  $f(x)$  определена и непрерывна в окрестности точки  $\bar{x}$ , перепишем систему (2)  $\bar{f}(x) = 0$  так:

$$\bar{x} = \bar{x} + \Lambda \bar{f}(x). \quad (10)$$

### Структурная схема алгоритма метода Ньютона

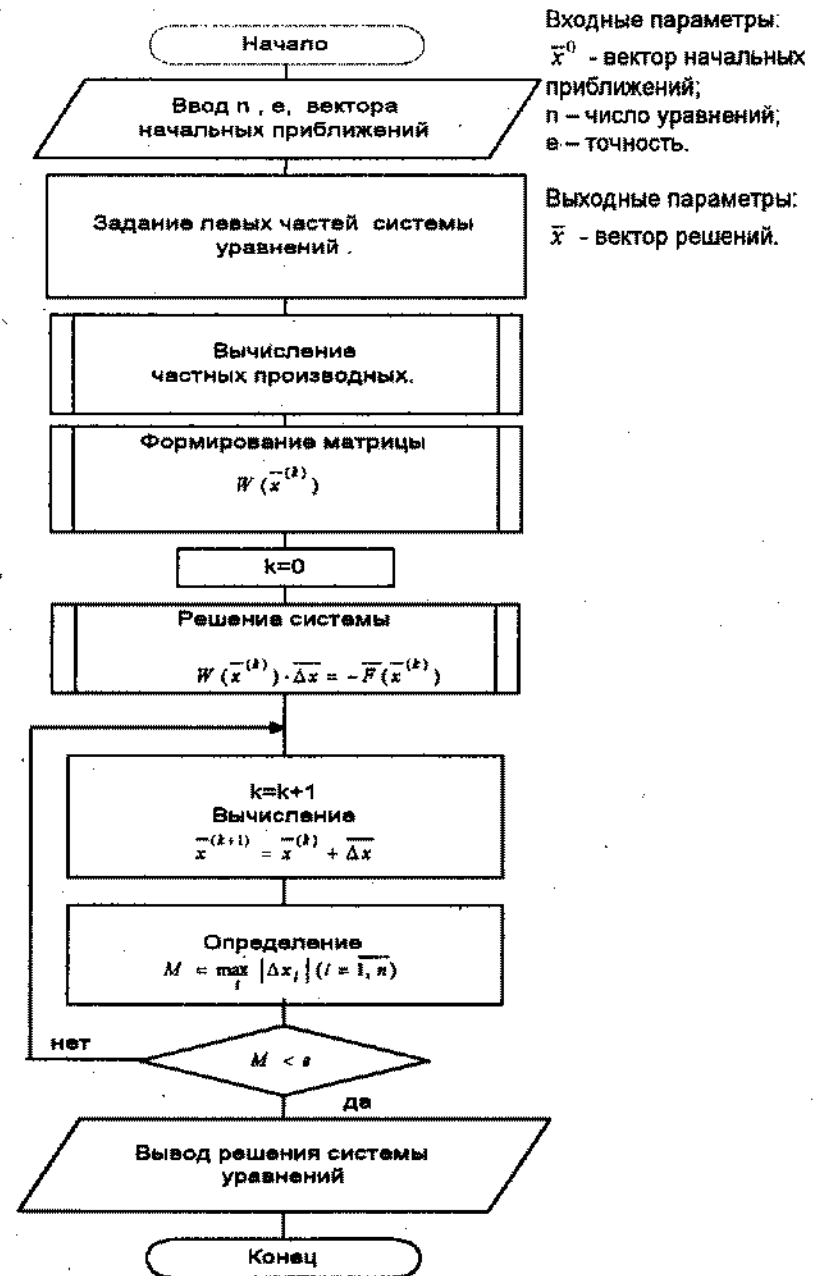


Рис. 3.2

**Процедура вычисления частных производных для формирования матрицы Якоби (CHP)**

Входные параметры:

- j - номер функции;
- i - номер аргумента;
- x(i) - аргумент;
- H - приращение аргумента.

Выходной параметр:

$$F_{ji} = \frac{\partial f_j}{\partial x_i}$$

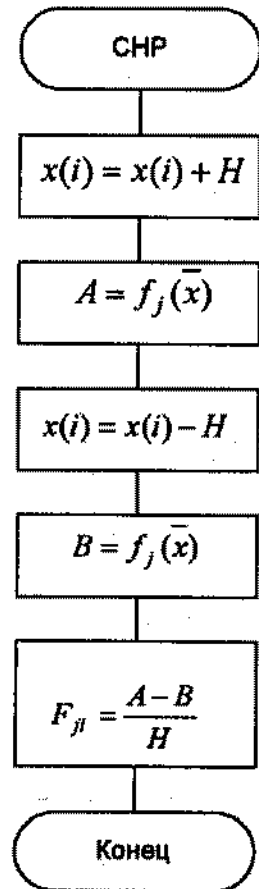


Рис. 3.3

**Процедура формирования матрицы Якоби (WM)**

Входные параметры:

- $\bar{x}$  - массив аргументов;
- n - размерность матрицы.

Выходной параметр: массив  $W_n$

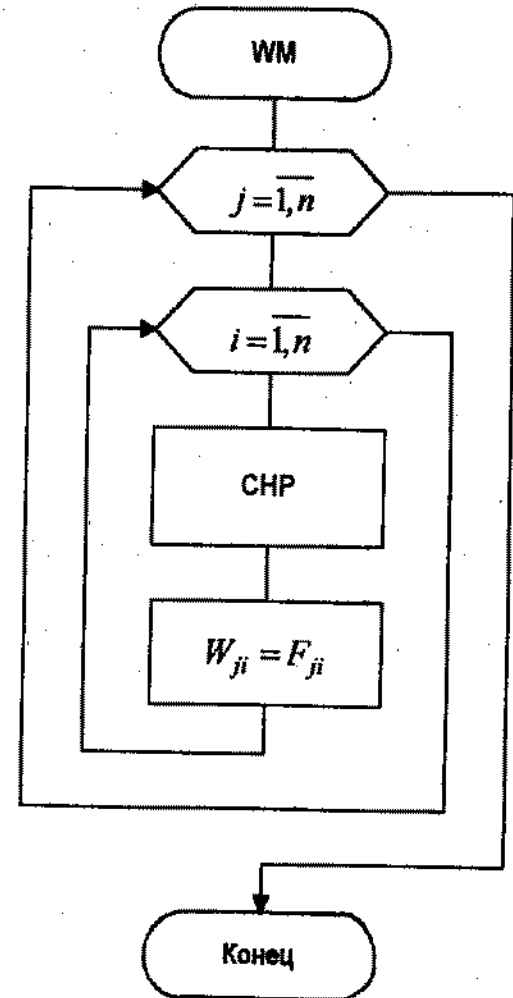


Рис. 3.4

Обозначим

$$\bar{x} + \Lambda f(\bar{x}) = \bar{\varphi}(\bar{x}), \quad (11)$$

тогда

$$\bar{x} = \bar{\varphi}(\bar{x}). \quad (12)$$

Параметр  $\Lambda$  выбираем из условия

$$\bar{\varphi}'(\bar{x}^{(0)}) = E + \Lambda f'(\bar{x}^{(0)}) = 0.$$

Если  $f'(\bar{x})$  - неособенная ( $\det f'(\bar{x}^{(0)}) \neq 0$ ), то

$$\Lambda = -[f'(\bar{x}^{(0)})]^{-1}. \quad (13)$$

Таким образом получаем формулу итерационного процесса

$$\bar{x}^{(k+1)} = \bar{\varphi}(\bar{x}^{(k)}), \quad k = 0, 1, \dots \quad (14)$$

**Замечания**

- Если  $\det f'(\bar{x}^{(0)}) = 0$ , следует выбрать другое начальное приближение.
- Выбор эквивалентной (2) зависимости (12) может быть сделан любым способом.

По определению норма матрицы

$$\|x\|_m = \max_i |x_i|, \quad (15)$$

$$\|x\|_l = \sum_{i=1}^n |x_i|. \quad (16)$$

Относительно области  $G$  введем нормы

$$\|\bar{\varphi}'(\bar{x})\|_m = \max_{x \in G} \|\bar{\varphi}'(\bar{x})\|_m, \quad (17)$$

$$\|\bar{\varphi}'(\bar{x})\|_l = \max_{x \in G} \|\bar{\varphi}'(\bar{x})\|_l, \quad (18)$$

где

$$\|\bar{\varphi}'(\bar{x})\|_m = \max_i \sum_{j=1}^n \left| \frac{\partial \varphi_i(\bar{x})}{\partial x_j} \right|, \quad \|\bar{\varphi}'(\bar{x})\|_l = \max_j \sum_{i=1}^n \left| \frac{\partial \varphi_i(\bar{x})}{\partial x_j} \right|.$$

Условие сходимости и оценка погрешности метода итерации формулирует

**Теорема** (достаточное условие сходимости).

Пусть функции  $\bar{\varphi}(\bar{x})$  и  $\bar{\varphi}'(\bar{x})$  непрерывны в области  $G$ , причем в  $G$  выполняется неравенство

$$\|\bar{\varphi}'(\bar{x})\|_l \leq q < 1, \quad (19)$$

$q$  - некоторая постоянная.

Если последовательные приближения

$$\bar{x}^{(k+1)} = \bar{\varphi}(\bar{x}^{(k)}), \quad k = 0, 1, \dots$$

не выходят из области  $G$ , то процесс итерации (14) сходится и предельный вектор

$$\bar{x}^* = \lim_{k \rightarrow \infty} \bar{x}^{(k)}$$

является в области  $G$  единственным решением системы (2).

**Следствие.**

Процесс итерации (14) сходится, если

$$\sum_{j=1}^n \left| \frac{\partial \varphi_i(\bar{x})}{\partial x_j} \right| \leq q_i < 1, \quad i = 1, 2, \dots, n \quad (20)$$

при  $x \in G$ .

Оценка погрешности определена неравенством

$$\left\| \bar{x}^* - \bar{x}^{(k)} \right\|_m \leq \frac{q^k}{1-q} \cdot \left\| \bar{x}^{(1)} - \bar{x}^{(0)} \right\|_m, \quad (21)$$

где  $\bar{x}^{(1)} = \bar{\varphi}(\bar{x}^{(0)})$ ,  $q = \max_i q_i$ .

**Пример 2.** [67]

Найти положительное решение системы уравнений методом итерации

$$\begin{cases} f_1(x, y) \equiv 2x^2 - xy - 5x + 1 = 0, \\ f_2(x, y) \equiv x + 3\lg(x) - y^2 = 0. \end{cases}$$

**Решение.**

Построим графики функций  $f_1(x, y)$  и  $f_2(x, y)$

Выберем начальное приближение  $x_0 = 3.5$ ;  $y_0 = 2.2$ .

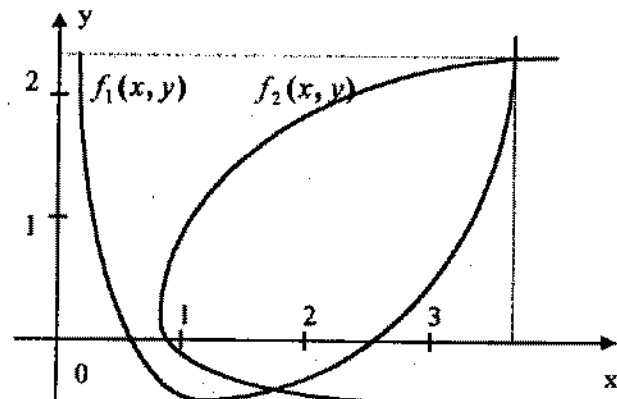


Рис. 3.5

Область G:  $\{|x-3.5| \leq 0.1, |y-2.2| \leq 0.1\}$

Запишем исходную систему в виде:

$$\begin{cases} x = \sqrt{\frac{x(y+5)-1}{2}} = \varphi_1(x, y), \\ y = \sqrt{x+3\lg x} = \varphi_2(x, y). \end{cases}$$

Вычислим производные функций  $\varphi_1(x, y)$  и  $\varphi_2(x, y)$ .

$$\frac{\partial \varphi_1}{\partial x} = \frac{y+5}{4\sqrt{\frac{x(y+5)-1}{2}}}, \quad \frac{\partial \varphi_2}{\partial x} = \frac{1 + \frac{3M}{x}}{2\sqrt{x+3\lg x}}$$

$$\frac{\partial \varphi_1}{\partial y} = \frac{x}{4\sqrt{\frac{x(y+5)-1}{2}}}, \quad \left| \frac{\partial \varphi_2}{\partial y} \right| = 0, \quad M = 0.43429.$$

В области G:

$$\left| \frac{\partial \varphi_1}{\partial x} \right| = \frac{2.3+5}{4\sqrt{\frac{3.4(2.1+5)-1}{2}}} < 0.54, \quad \left| \frac{\partial \varphi_1}{\partial y} \right| = \frac{3.6}{4\sqrt{\frac{3.4(2.1+5)-1}{2}}} < 0.27,$$

$$\left| \frac{\partial \varphi_2}{\partial x} \right| = \frac{1 + \frac{3 \cdot 0.43}{3.4}}{2\sqrt{3.4+3\lg 3.4}} < 0.31, \quad \left| \frac{\partial \varphi_2}{\partial y} \right| = 0.$$

Согласно (19), так как

$$\left| \frac{\partial \varphi_1}{\partial x} \right| + \left| \frac{\partial \varphi_2}{\partial x} \right| < 0.54 + 0.31 = 0.85 < 1, \quad \left| \frac{\partial \varphi_1}{\partial y} \right| + \left| \frac{\partial \varphi_2}{\partial y} \right| < 0.27 + 0 = 0.27 < 1,$$

то процесс итерации будет сходиться, если последовательные приближения не выйдут из области G, причем  $q = 0.85$ .

$$\begin{cases} x_{k+1} = \sqrt{\frac{x_k(y_k+5)-1}{2}}, \\ y_{k+1} = \sqrt{x_k+3\lg(x_k)} \end{cases} \quad (k = 0, 1, 2, \dots).$$

Таблица 2

k	0	1	2	3	4	5	6	7
$x_k$	3.5	3.484	3.483	3.484	3.485	3.486	3.487	3.487
$y_k$	2.2	2.252	2.258	2.259	2.260	2.261	2.261	2.261

Таким образом, принимаем

$$\begin{cases} x^* = 3.487, \\ y^* = 2.261. \end{cases}$$

По формуле (21) оценим погрешность

$$\left\| \begin{matrix} x^* - x^{(6)} \\ y^* - y^{(6)} \end{matrix} \right\|_{\infty} \leq 0.188.$$

При подстановке значений  $x^{(6)}$  и  $y^{(6)}$  в исходную систему невязки соответственно равны  $(-0.004256)$  и  $0.001244$ .

Структурная схема алгоритма метода итерации приведена на рис.

6-8.

### 3.2.1. Приведение системы $\bar{f}(x) = 0$ к виду, удобному для итерации

Если начальное приближение  $\bar{x}^{(0)}$  выбрано, то итерирующие функции  $\bar{\varphi}(\varphi_1, \dots, \varphi_m)$  следует выбирать так, чтобы условие (19) выполнялось с возможно меньшим значением  $q$ . В этом случае скорость сходимости будет наибольшей. Покажем на примере двух уравнений с двумя неизвестными практику выбора функций  $\varphi_1$  и  $\varphi_2$ .

$$\begin{cases} f_1(x, y) = 0, \\ f_2(x, y) = 0. \end{cases}$$

Будем искать  $\varphi_1$  и  $\varphi_2$  в виде

$$\begin{cases} \varphi_1(x, y) = x + \alpha \cdot f_1(x, y) + \beta \cdot f_2(x, y), \\ \varphi_2(x, y) = y + \gamma \cdot f_1(x, y) + \delta \cdot f_2(x, y). \end{cases} \quad (22)$$

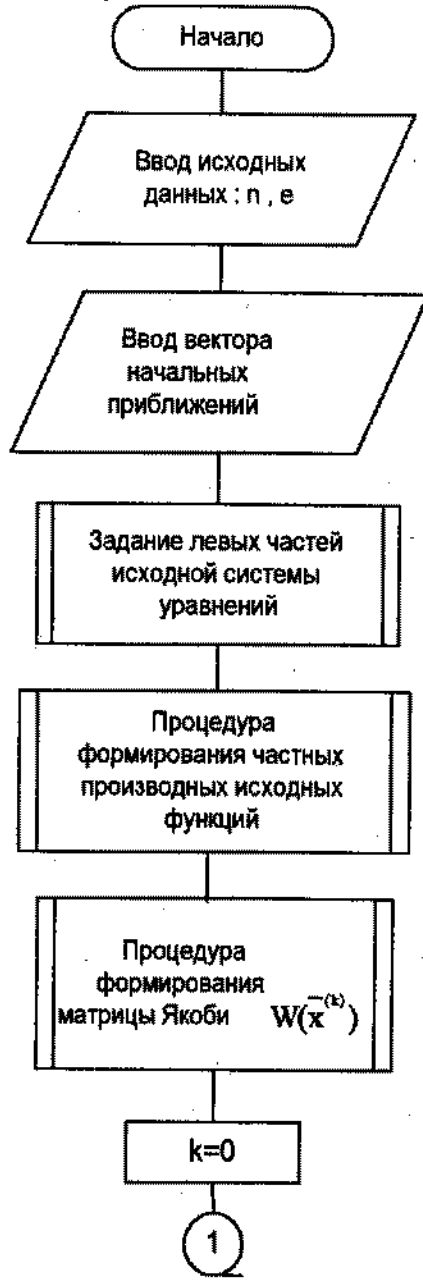
Потребуем, чтобы в точке  $\bar{x}^{(0)}$

$$\frac{\partial \varphi_1}{\partial x} = \frac{\partial \varphi_1}{\partial y} = \frac{\partial \varphi_2}{\partial x} = \frac{\partial \varphi_2}{\partial y} = 0. \quad (23)$$

Для определения  $\alpha, \beta, \gamma, \delta$  получим системы линейных уравнений:

$$\begin{cases} 1 + \alpha \frac{\partial f_1}{\partial x} + \beta \frac{\partial f_2}{\partial x} = 0, \\ \alpha \frac{\partial f_1}{\partial y} + \beta \frac{\partial f_2}{\partial y} = 0. \end{cases} \quad (24)$$

**Структурная схема алгоритма метода итерации  
для решения системы нелинейных уравнений**



Входные параметры:  
 $\bar{x}^0$  - вектор начальных приближений;  
 $n$  - число уравнений;  
 $e$  - точность.  
 Выходные параметры:  
 $\bar{x}$  - вектор решения.

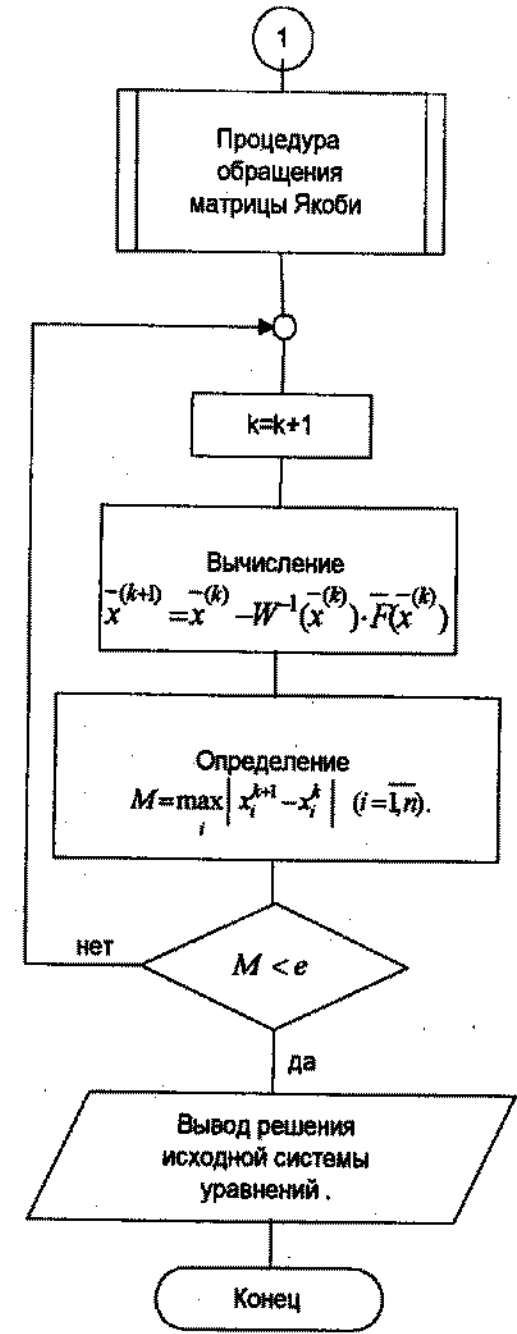


Рис. 3.6

### Процедура обращения матрицы методом Гаусса (OBRM)

Входные параметры:  $n$  - порядок матриц;  
 $A$  - исходная матрица.

Выходной параметр:  $B$  - матрица алгебраических дополнений матрицы  $A$ ;  
 $Q$  - матрица, обратная исходной.

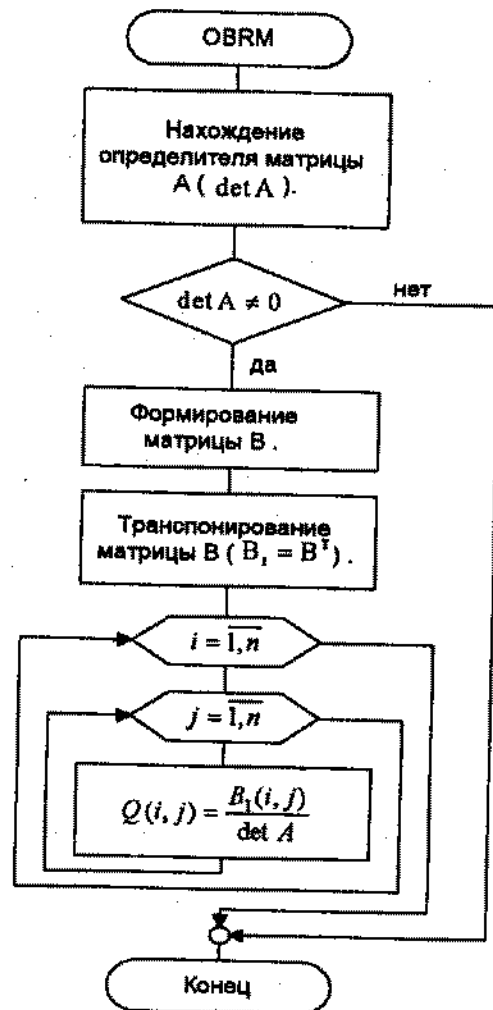


Рис. 3.7

$$\begin{cases} \gamma \cdot \frac{\partial f_1}{\partial x} + \delta \cdot \frac{\partial f_2}{\partial x} = 0, \\ 1 + \gamma \cdot \frac{\partial f_1}{\partial y} + \delta \cdot \frac{\partial f_2}{\partial y} = 0. \end{cases} \quad (25)$$

Решая системы (24) и (25), получим

$$\frac{\partial \varphi_1}{\partial x} + \frac{\partial \varphi_1}{\partial y} = 0 \quad \text{и} \quad \frac{\partial \varphi_2}{\partial x} + \frac{\partial \varphi_2}{\partial y} = 0.$$

В точке  $x^{-(0)}$  и в достаточно малой окрестности  $x^{-(0)}$  будет  $q \ll 1$ , то есть условие сходимости метода итерации выполняется.

#### Пример 3.

Дана система уравнений

$$\begin{cases} x^2 + y^2 - 1 = 0, \\ x^3 - y = 0 \end{cases}$$

и начальное приближение  $x^{-(0)} = (0.8; 0.55)$ . Привести к виду, удобному для итерации.

Решение.

Ищем  $\varphi_1$  и  $\varphi_2$  в виде

$$\varphi_1(x, y) = x + \alpha(x^2 + y^2 - 1) + \beta(x^3 - y),$$

$$\varphi_2(x, y) = y + \gamma(x^2 + y^2 - 1) + \delta(x^3 - y).$$

Имеем

$$\frac{\partial f_1}{\partial x} = 2x, \quad \frac{\partial f_1}{\partial y} = 2y, \quad \frac{\partial f_2}{\partial x} = 3x^2, \quad \frac{\partial f_2}{\partial y} = -1.$$

Следовательно, в точке  $x^{-(0)}$

$$\frac{\partial f_1}{\partial x} = 1.6, \quad \frac{\partial f_1}{\partial y} = 1.1, \quad \frac{\partial f_2}{\partial x} = 1.92, \quad \frac{\partial f_2}{\partial y} = -1.$$

Получаем системы

$$\begin{cases} 1 + 1.6 \cdot \alpha + 1.92 \cdot \beta = 0, \\ 1.1 \cdot \alpha - \beta = 0; \end{cases}$$

$$\begin{cases} 1.6\gamma + 1.92\delta = 0, \\ 1 + 1.1\gamma - \delta = 0. \end{cases}$$

Откуда

$$\alpha \approx -0.3, \quad \beta \approx -0.3, \quad \gamma \approx -0.5, \quad \delta \approx 0.4.$$

Итак,

$$\varphi_1(x, y) = x - 0.3 \cdot (x^2 + y^2 - 1) - 0.3 \cdot (x^3 - y),$$

$$\varphi_2(x, y) = y - 0.5 \cdot (x^2 + y^2 - 1) + 0.4 \cdot (x^3 - y).$$

Метод Ньютона и метод простой итерации являются итерационными методами. Метод Ньютона обладает большей скоростью сходимости. В то же время метод простой итерации имеет более простые достаточные условия сходимости и менее критичен к выбору начального приближения. Поэтому для высокоточных вычислений рекомендуется применить вначале метод простой итерации. После того, как найдено приближение, достаточно близкое к точному, применить метод Ньютона. Для обращения матрицы Якоби можно воспользоваться методом Гаусса.

Принципиально другой метод решения системы нелинейных уравнений - сведение к решению задачи оптимизации.

Составим функцию

$$Q(x_1, x_2, \dots, x_n) = \sum_{i=1}^n (f_i(x_1, x_2, \dots, x_n))^2. \quad (26)$$

Очевидно, что минимум функции  $Q(x)$  равен нулю и достигается в точке, соответствующей решению системы уравнений (1).

Таким образом, решив задачу оптимизации функции (26), получим решение системы нелинейных уравнений.

## 4. АППРОКСИМАЦИЯ ФУНКЦИЙ

В задачах теории колебаний, электродинамики, твердотельной электроники и других разделах прикладной математики широко применяется аппроксимация функций при описании физических параметров сред, для задания характеристик активных и пассивных элементов путем радиотехнических цепей и т. д. В задачах вычислительной математики аппроксимация функций является основой для разработки многих методов и алгоритмов.

Задача аппроксимации (приближения) функций заключается в замене некоторой функции  $y = f(x)$  другой функцией  $g(x, a_0, a_1, \dots, a_n)$  так, чтобы отклонение  $g$  от  $f(x)$  удовлетворяло в некоторой области (на множестве  $X$ ) определенному условию.

При этом функция  $g(x, a_0, a_1, \dots, a_n)$  обычно выбирается с учетом специфических особенностей рассматриваемой функции  $f(x)$ . Если множество  $X$  дискретно (состоит из отдельных точек), то приближение называется точечным, если же  $X$  есть отрезок  $a \leq x \leq b$ , то приближение называется интегральным.

Частным случаем аппроксимации является интерполяция.

### 4.1. Постановка задачи интерполяции

Значения функции  $f(x)$  известны в точках  $x_0, x_1, \dots, x_n$ , называемых узлами интерполяции.

Требуется построить некоторую функцию  $g(x, a_0, a_1, \dots, a_n)$  такую, что

$$g(x, a_0, a_1, \dots, a_n) = f(x), \quad i = \overline{0, n}. \quad (1)$$

Условие (1) — условие интерполяции.

В зависимости от класса, к которому принадлежит функция  $g$ , интерполяция называется параболической, тригонометрической и т. д.

Выделение класса параболического интерполирования вызвано наличием многочисленных приложений, а также тем фактом, что аппарат интерполирования многочленами является столь же важным в теории численного анализа, как формула Тейлора в классическом анализе.

В случае параболической интерполяции функция  $g(x, a_0, a_1, \dots, a_n)$  суть многочлен (полином) степени не выше  $n$

$$g(x, a_0, a_1, \dots, a_n) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \quad (2)$$

Условие интерполяции (1) будет иметь вид:

$$f(x_i) = a_0 + a_1x_i + \dots + a_nx_i^n, \quad i = \overline{0, n}. \quad (3)$$

Из решения системы (3) определяются коэффициенты  $a_0, a_1, \dots, a_n$ . Интерполяционный многочлен

$$f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$$

Задача имеет единственное решение, если все узлы интерполяции различны, т. к. в этом случае определитель системы отличен от нуля.

$$\Delta = \begin{vmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{vmatrix} \neq 0.$$

## 4.2. Интерполяционный полином Лагранжа

Интерполяционный многочлен Лагранжа имеет вид

$$L_n(x) = \sum_{i=0}^n f(x_i) \prod_{j \neq i} \frac{x - x_j}{x_i - x_j} = \sum_{i=0}^n \frac{(x - x_0)(x - x_1) \dots (x - x_{i-1})(x - x_{i+1}) \dots (x - x_n)}{(x_i - x_0)(x_i - x_1) \dots (x_i - x_{i-1})(x_i - x_{i+1}) \dots (x_i - x_n)} f(x_i). \quad (4)$$

Степень многочлена (4) не выше  $n$ . В силу условия (3) имеем:

$$L_n(x_j) = f(x_j).$$

Убедимся, что многочлен первой степени

$$L_1(x) = \frac{x - x_1}{x_0 - x_1} y_0 + \frac{x - x_0}{x_1 - x_0} y_1 \text{ удовлетворяет условию интерполяции.}$$

$$L_1(x_0) = y_0, \quad L_1(x_1) = y_1.$$

Многочлен  $L_2(x)$ , построенный по узлам

$x_0, x_1, x_2$ :

$$L_2(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} y_0 + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} y_1 + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} y_2,$$

$$L_2(x_0) = y_0, \quad L_2(x_1) = y_1, \quad L_2(x_2) = y_2.$$

Очевидно, что в точках, не являющихся узлами интерполяции, функции  $f(x)$  и  $L_n(x)$  отличаются и для определения отклонения этих функций необходима оценка погрешности.

В предположении непрерывности  $f^{(n+1)}(x)$

$$|f(x) - L_n(x)| \leq \frac{|f^{(n+1)}(x^*)| |\Pi_{n+1}(x)|}{(n+1)!}, \quad (5)$$

где

$$|\Pi_{n+1}(x)| = |(x - x_0) \cdot (x - x_1) \dots (x - x_n)|, \quad (6)$$

$$x \in [t_1; t_2],$$

$$t_1 = \min(x_0, x_1, \dots, x_n, x), \quad (7)$$

$$t_2 = \max(x_0, x_1, \dots, x_n, x),$$

$$f^{(n+1)}(x^*) = \max |f^{(n+1)}(x)|,$$

$f^{(n+1)}(x)$  —  $(n+1)$  производная функции  $f(x)$ .

Из формулы (6) видно, что, выбирая «удачно» узлы интерполяции, можно уменьшить  $|\Pi_{n+1}(x)|$ .

Задача о наиболее рациональном выборе узлов интерполяции  $x_i$  (при заданном их количестве), заключающаяся в том, чтобы полином  $\Pi_{n+1}(x)$  имел наименьшее по абсолютной величине значение на  $[a; b]$ , была решена математиком П. Л. Чебышевым. Он доказал, что наилучшим выбором узлов интерполяции является

$$x_i = \frac{b+a}{2} + \frac{b-a}{2} z_i, \quad i = 0, 1, \dots, n \quad (8)$$

$$z_i = -\cos \frac{2i+1}{2n+2} \pi$$

нули полинома Чебышева  $T_{n+1}(x)$ . В этом случае

$$|\Pi_{n+1}(x)| \leq 2 \left( \frac{b-a}{4} \right)^{n+1}.$$

Даже при таком выборе узлов в общем случае нельзя говорить, что абсолютная погрешность будет сколь угодно мала при достаточно большом  $n$ .

Если узлы функции  $f(x)$  уже заданы, иногда удается уменьшить погрешность интерполяции, последовательно выбирая в качестве  $x_0, x_1$  и т. д. узлы, наиболее близко расположенные к точке  $x$ , по возможности симметрично относительно этой точки. Такой прием позволяет уменьшить степень интерполяционного многочле-



на для достижения требуемой точности (не использовать все заданные узлы).

Практически оценку погрешности определяют исходя из приближенного равенства

$$|f(x) - L_m(x)| \approx |L_{m+1}(x) - L_m(x)|. \quad (9)$$

Поэтому  $\varepsilon_m = |L_{m+1}(x) - L_m(x)|$  можно принять за приближенную оценку погрешности интерполяционного многочлена Лагранжа.

Последовательно вычисляют  $L_0(x)$ ,  $L_1(x)$ ,  $\varepsilon_1$ ,  $L_2(x)$ ,  $\varepsilon_2, \dots$

Если при некотором  $m$  погрешность  $\varepsilon_m \leq E$  ( $E$  - заданная точность интерполяции), то вычисления заканчивают и полагают  $f(x) \approx L_m(x)$ .

Если это неравенство не выполняется ни при каком  $m$  ( $m \leq n$ ), то находят  $\varepsilon_{m\delta} = \min_m \varepsilon_m$  и полагают  $f(x) = L_{m\delta}(x)$ .

Если этот минимум достигается при нескольких  $m$ , то выбирают наименьшее  $m$ .

Если величины  $\varepsilon_m$ , начиная с некоторого  $m$ , имеют устойчивую тенденцию к увеличению, то вычисления  $L_m(x)$  и  $\varepsilon_m$  прекращают, полагая  $f(x) = L_m(x)$ .

### Пример 1.

Для таблично заданной функции  $y = f(x)$  на  $[1; 1.20]$  вычислить значение интерполирующего многочлена при  $x^* = 1.12$  и определить погрешность.

Таблица 1

$x$	1	1.05	1.10	1.15	1.20
$y = f(x)$	-0.0118	0	0.0112	0.0218	0.0318

### Решение.

Так как  $f(x)$  задана таблично, то для оценки погрешности будем пользоваться алгоритмом практической оценки погрешности.

Перенумеруем узлы ( $x_0$  - ближайших к  $x^*$  узел и т. д.)

Таблица 2

$i$	0	1	2	3	4
$x_i$	1.10	1.15	1.05	1.20	1
$y_i$	0.0112	0.0218	0	0.0318	-0.0118

Запишем  $L_1(x)$  по двум узлам  $x_0, x_1$  и вычислим  $L_1(x^*)$ .

$$L_1(x) = \frac{x-x_1}{x_0-x_1} y_0 + \frac{x-x_0}{x_1-x_0} y_1,$$

$$L_1(1.12) = 0.01544.$$

Многочлен  $L_2(x)$ , по узлам  $x_0, x_1, x_2$ :

$$L_2(x) = \frac{(x-x_1) \cdot (x-x_2)}{(x_0-x_1) \cdot (x_0-x_2)} y_0 + \frac{(x-x_0) \cdot (x-x_2)}{(x_1-x_0) \cdot (x_1-x_2)} y_1 + \frac{(x-x_0) \cdot (x-x_1)}{(x_2-x_0) \cdot (x_2-x_1)} y_2,$$

$$L_2(1.12) = 0.01551$$

$$\text{Вычислим } \varepsilon_2 = |L_2 - L_1| = 0.00007.$$

Запишем  $L_3(x)$  по узлам  $x_0, x_1, x_2, x_3$ :

$$L_3(x) = \frac{(x-x_1) \cdot (x-x_2) \cdot (x-x_3)}{(x_0-x_1) \cdot (x_0-x_2) \cdot (x_0-x_3)} y_0 + \frac{(x-x_0) \cdot (x-x_2) \cdot (x-x_3)}{(x_1-x_0) \cdot (x_1-x_2) \cdot (x_1-x_3)} y_1 + \\ + \frac{(x-x_0) \cdot (x-x_1) \cdot (x-x_3)}{(x_2-x_0) \cdot (x_2-x_1) \cdot (x_2-x_3)} y_2 + \frac{(x-x_0) \cdot (x-x_1) \cdot (x-x_2)}{(x_3-x_0) \cdot (x_3-x_1) \cdot (x_3-x_2)} y_3,$$

$$L_3(1.12) = 0.01551$$

$$\varepsilon_3 = |L_3 - L_2| = 0, \text{ полагаем } f(1.12) = L_2(1.12) = 0.01551.$$

### Пример 2.

С какой точностью можно вычислить  $\sqrt{115}$ , взяв в качестве узлов  $x_0 = 100$ ,  $x_1 = 121$ ,  $x_2 = 144$ ?

Решение.

Рассмотрим функцию  $y = \sqrt{x}$  и оценим  $R = |f(x) - L_2(x)|$ . По формуле (5)

$$R \leq \frac{M_3 \Pi_3(x)}{3!},$$

где  $M_3 = \max |y'''(x)|$ , при  $x \in [100; 144]$ ,

$$\Pi_3(x) = |(x-x_0) \cdot (x-x_1) \cdot (x-x_2)|.$$

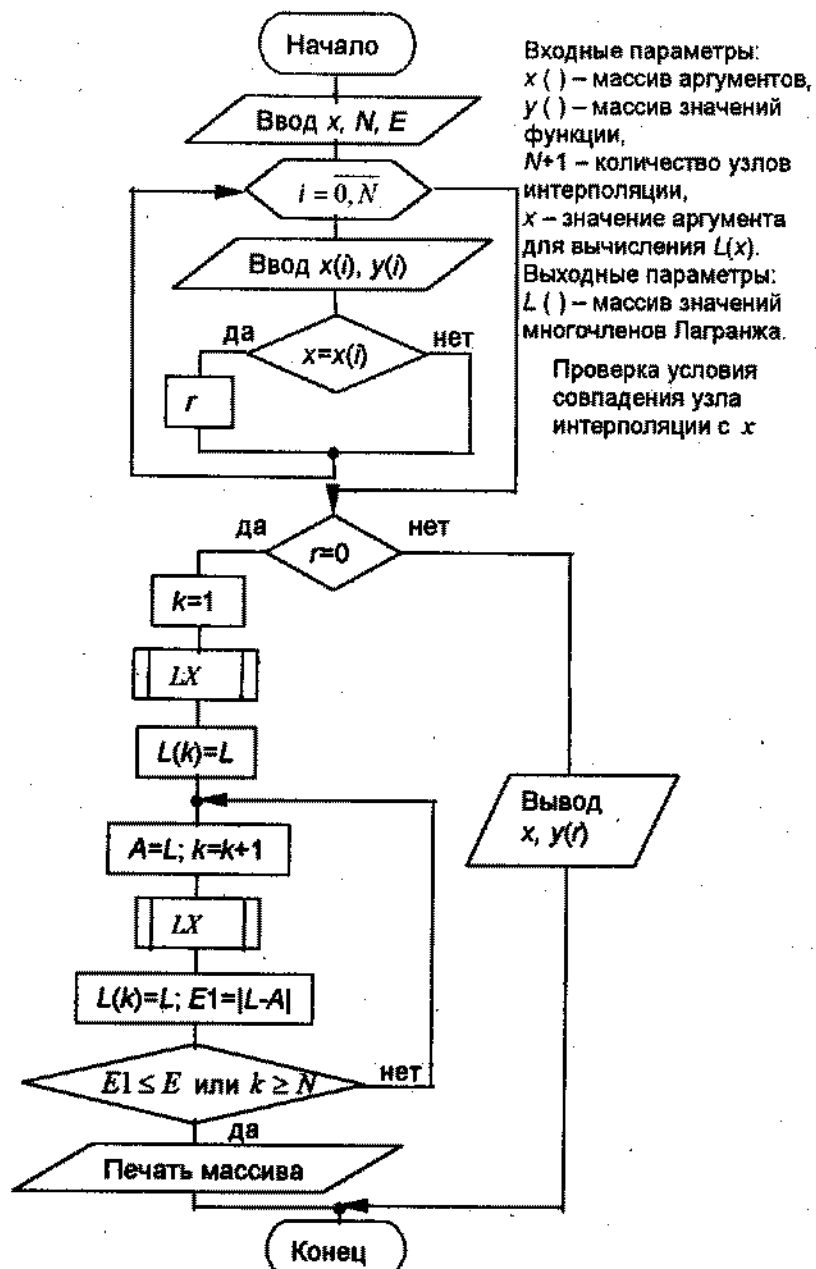
$$y'(x) = (\sqrt{x})' = \frac{1}{2} x^{-1/2}, \quad y''(x) = -\frac{1}{4} x^{-3/2}, \quad y'''(x) = \frac{3}{8} x^{-5/2},$$

$$M_3 = \frac{3}{8} \cdot 10^{-5},$$

$$R_3 \leq \frac{3}{8 \cdot 3!} \cdot 10^{-5} \cdot |(115-100)(115-121)(115-144)| \approx 1.6 \cdot 10^{-5}.$$

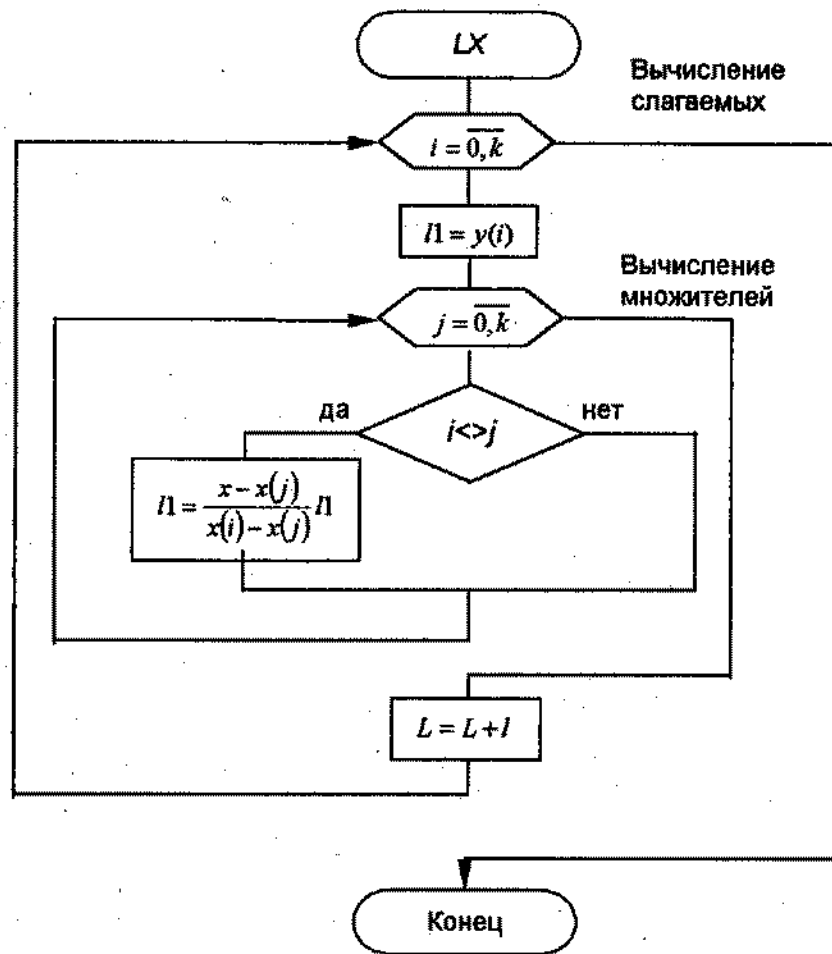
Итак, точность вычисления величины  $\sqrt{115}$  равна  $1.6 \cdot 10^{-5}$ .

### Схема алгоритма вычисления многочлена Лагранжа



### Процедура вычисления $L(x)$ – значения многочлена Лагранжа $k$ -ой степени в точке $x$

Входные параметры:  
 $k$  – порядок многочлена;  
 $x()$  – массив значений аргумента;  
 $y()$  – массив значений функции;  
 $x$  – значение аргумента для вычисления  $L(x)$ .  
 Выходные параметры:  
 $L$  – значение многочлена Лагранжа в точке  $x$ .



### 4.3. Интерполяционные формулы Ньютона

Если интерполируемая функция  $f(x)$  задана в  $(n+1)$  равноотстоящих узлах, т. е.  $x_{i+1} - x_i = \Delta x_i = h = const$ , то для ее интерполяции удобнее использовать формулы Ньютона.

Будем искать интерполирующий полином степени  $n$  в виде  $P_n(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)(x-x_1) + \dots + a_n(x-x_0)(x-x_1)\dots(x-x_{n-1})$  (10)

Для определения коэффициентов  $a_0, a_1, \dots, a_n$  воспользуемся условием интерполяции (3)

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

Определим конечные разности

$$\begin{aligned} \Delta y_0 &= y_1 - y_0, & \Delta y_1 &= y_2 - y_1, & \dots & \Delta y_i &= y_{i+1} - y_i, \\ \Delta^2 y_0 &= \Delta y_1 - \Delta y_0, & \Delta^2 y_1 &= \Delta y_2 - \Delta y_1, & \dots & \Delta^2 y_i &= \Delta y_{i+1} - \Delta y_i. \end{aligned}$$

В общем случае

$$\Delta^k y_i = \Delta^{k-1} y_{i+1} - \Delta^{k-1} y_i, \quad (11)$$

порядок конечной разности  $k = 1, 2, \dots$

Конечные разности удобно записывать в виде горизонтальной таблицы.

$x$	$y$	$\Delta y$	$\Delta^2 y$	$\Delta^3 y$	
$x_0$	$y_0$	$\Delta y_0$	$\Delta^2 y_0$	$\Delta^3 y_0$	
$x_1$	$y_1$	$\Delta y_1$	$\Delta^2 y_1$	$\Delta^3 y_1$	
$x_2$	$y_2$	$\Delta y_2$	$\Delta^2 y_2$	$\Delta^3 y_2$	

Для коэффициентов справедлива формула

$$a_i = \frac{\Delta^i y_0}{i! h^i}, \quad i = 0, 1, 2, \dots, n. \quad (12)$$

Первая формула Ньютона имеет вид [23]

$$P_n(x) = y_0 + q \Delta y_0 + \frac{q(q-1)}{2!} \Delta^2 y_0 + \dots + \frac{q(q-1)(q-2)\dots(q-n+1)}{n!} \Delta^n y_0, \quad (13)$$

$$q = \frac{x - x_0}{h}, \quad x_0 \text{ — ближайший к точке } x \text{ узел слева.} \quad (14)$$

**Замечание.**

Справедливы соотношения:

$$\Delta^n P_n(x) = n! a_0 h^n = const, \quad \Delta^s P_n(x) = 0 \text{ при } s > n.$$

Если таблица значений функции конечна, то наивысший порядок интерполяционного многочлена равен  $n$ , где  $(n+1)$  - количество узлов интерполяции.

Если дана неограниченная таблица, то число  $n$  в формуле (13) может быть любым.

Практически число  $n$  выбирают так, чтобы разность  $\Delta^n y_i$  была постоянна с заданной степенью точности. Считается, что конечные  $(k+1)$  порядки разности равны нулю, если выполняются условия

$$\begin{cases} |\Delta^{k+1} y_i| < E, \\ \left| \sum_{i=1}^n \Delta^{k+1} y_i \right| < E. \end{cases} \quad (15)$$

Формула (13) применяется для интерполяции в начале таблицы ("интерполяция вперед"), т. к. в этом случае имеется возможность, взяв большее число узлов, увеличить точность интерполирования.

Для интерполяции в конце таблицы ("интерполяция назад") используют вторую интерполяционную формулу Ньютона

$$P_n(x) = y_n + q \Delta y_{n-1} + \frac{q(q+1)}{2!} \Delta^2 y_{n-2} + \dots + \frac{q(q+1)(q+2)\dots(q+n-1)}{n!} \Delta^n y_0 \quad (16)$$

$$q = \frac{x - x_n}{h}, \quad (17)$$

$x_n$  - ближайший к точке  $x$  узел справа.

Оценки погрешностей формул (13) и (16) имеют соответственно вид (18) и (19).

$$R_n(x) = h^{n+1} \frac{q(q-1)\dots(q-n)}{(n+1)!} |f^{(n+1)}(x^*)|, \quad (18)$$

$$R_n(x) = h^{n+1} \frac{q(q+1)\dots(q+n)}{(n+1)!} |f^{(n+1)}(x^*)|, \quad (19)$$

$$x^* \in [x_0, x_n].$$

Практически интерполяционные формулы обрываются на членах, содержащих разности, постоянные в пределах заданной точности. Предполагая, что  $\Delta^{n+1} y \approx const$ ,  $h$  - достаточно мало и учитывая, что

$$f^{(n+1)}(x) = \lim_{h \rightarrow 0} \frac{\Delta^{n+1} y}{h^{n+1}}, \text{ полагают } f^{(n+1)}(x) \approx \frac{\Delta^{n+1} y_0}{h^{n+1}}.$$

Погрешность первой интерполяционной формулы (18)

$$R_n(x) \approx \frac{q(q-1)\dots(q-n)}{(n+1)!} |\Delta^{n+1} y_0|, \quad (20)$$

погрешность второй интерполяционной формулы —

$$R_n(x) \approx \frac{q(q+1)\dots(q+n)}{(n+1)!} |\Delta^{n+1} y_n|. \quad (21)$$

Степень интерполирующего полинома существенно зависит от шага таблицы (чем меньше шаг, тем меньше степень полинома). При возможности, шаг выбирают так, чтобы уже, например, разности второго порядка практически оказывались равными нулю. Тогда возможна линейная интерполяция. Формулы Ньютона позволяют проводить экстраполяцию, т. е. находить значения функции для аргументов, лежащих вне таблицы. Отличие в том, что для интерполяции по формуле (13) значение  $q > 0$ , при экстраполяции  $q < 0$ . Для формулы (16) наоборот: при интерполяции  $q < 0$ , при экстраполяции  $q > 0$ .

Интерполяционный многочлен Лагранжа в случае, если интерполируемая функция  $y = f(x)$  задана в  $(n+1)$  равноотстоящих узлах имеет вид:

$$L_n(x) = \sum_{i=0}^n f(x_i) \frac{(-1)^{n-i}}{i!(n-i)!} \frac{q(q-1)(q-2)\dots(q-n)}{q-i}. \quad (4)$$

$$q = \frac{x-x_0}{h}, \quad x_0 \text{ — ближайший к точке } x \text{ узел слева.}$$

### Пример 3.

Требуется вычислить с точностью  $\varepsilon = 10^{-5}$   $f(1043)$  и  $f(1098)$ , если функция  $y = f(x)$  задана таблично.

Таблица 3

x	y
1010	3.0043214
1020	3.0086002
1030	3.0128372
1040	3.0170333
1050	3.0218930
1060	3.0253059
1070	3.0293838
1080	3.0334238
1090	3.0374265
1100	3.0413927

Решение.

Составим таблицу конечных разностей.

Таблица 4

x	y	$\Delta y$	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$
1010	3.0043214	0.0042788	-0.0000418	0.0000009	-0.0000001
1020	3.0086002	0.0042370	-0.0000409	0.0000008	-0.0000001
1030	3.0128372	0.0041961	-0.0000401	0.0000007	0
1040	3.0170333	0.0041560	-0.0000394	0.0000007	0.0000001
1050	3.0218930	0.0041166	-0.0000387	0.0000008	-0.0000002
1060	3.0253059	0.0040779	-0.0000379	0.0000006	0.0000002
1070	3.0293838	0.0040400	-0.0000373	0.0000008	
1080	3.0334238	0.0040027	-0.0000365		
1090	3.0374265	0.0039662			
1100	3.0413927				

При  $k=3$  выполняются условия (15). Это означает, что функцию следует интерполировать полиномом третьей степени, полагая  $x_0 = 1040$  (ближайший к  $x = 1043$  узел слева)

$$x_0 = 1040, \quad y_0 = 3.0170333, \quad h = 10, \quad q = \frac{1043-1040}{10} = 0.3.$$

Тогда, следуя (13)

$$y(1043) = 3.0170333 + 0.3 \cdot 0.0041560 + \frac{0.3(0.3-1)}{2} (-0.0000394) + \frac{0.3(0.3-1)(0.3-2)}{6} 0.0000007 = 3.0182284.$$

Для вычисления  $y(1098)$  примем  $x_n = 1100$ ,  $y_n = 3.0413927$ ,

тогда

$$q = \frac{1098-1100}{10} = -0.2.$$

По формуле (16)

$$y(1098) = 3.0413927 + (-0.2) \cdot 0.0039662 + \frac{(-0.2)(-0.2+1)}{2} (-0.0000365) + \frac{(-0.2)(-0.2+1)(-0.2+2)}{6} 0.0000008 = 3.0406023$$

Сравнение методов Лагранжа и Ньютона позволяет сделать следующие выводы.

1. Формула Лагранжа может быть использована для произвольно заданных узлов интерполяции. Формулы Ньютона пригодны лишь в случае равноотстоящих узлов.

2. В формуле Лагранжа выбор узлов и свойства функции независимо друг от друга влияют на величину погрешности  $L_n(x)$ . Это облегчает оценку погрешности приближения функции многочленом Лагранжа.

3. Часто точность интерполяции повышают путем добавления к взятым узлам хотя бы одного нового узла. В формуле Лагранжа добавление узла повлечет за собой не только появление нового слагаемого в сумме (4), но и потребует «исправления» всех ранее найденных членов суммы.

4. В формулах Ньютона в случае добавления узла все найденные члены сохраняются и появляется новое слагаемое, представляющее собой не что иное, как поправку к уже вычисленному значению.

5. При интерполяции на малых участках слагаемые в формулах (13) и (16) будут расположены в порядке их малости, что облегчает использование формул Ньютона в вычислениях и позволяет судить о точности интерполяции.

#### 4.4. Обратная интерполяция

Наряду с задачей интерполяции в технических приложениях ставится задача обратного интерполирования. Пусть известна зависимость  $y = f(x)$ , в точках  $x_i$ ,  $i = \overline{0, n}$ , т. е. известны  $y_i = f(x_i)$ . Эта информация эквивалентна тому, что известны значения  $x_i = g(y_i)$  - обратной функции. При условии допустимости интерполяции по переменной  $y$  можно заменить обратную функцию  $g(y)$  интерполирующим многочленом  $L_n(y_i) = x_i$ ,  $i = \overline{0, n}$ .

##### Пример 4.

Требуется восстановить форму входного сигнала  $x(t)$  (предполагается, что это какой-то низкочастотный сигнал). Связь мгновенных значений входного сигнала  $x(t)$  и выходного сигнала  $y(t)$  определяется нелинейной динамической характеристикой  $y = \varphi(x)$ . Задачу решить методом обратного интерполирования.

$x$	-0.9	-0.3	0.3	0.9
$y = \varphi(x)$	0.31623	0.83666	1.14017	1.37840

**Решение.**

Строим многочлен Лагранжа третьего порядка

$$L_3(y) = \frac{(y-y_1)(y-y_2)(y-y_3)}{(y_0-y_1)(y_0-y_2)(y_0-y_3)}x_0 + \frac{(y-y_0)(y-y_2)(y-y_3)}{(y_1-y_0)(y_1-y_2)(y_1-y_3)}x_1 + \\ + \frac{(y-y_0)(y-y_1)(y-y_3)}{(y_2-y_0)(y_2-y_1)(y_2-y_3)}x_2 + \frac{(y-y_0)(y-y_1)(y-y_2)}{(y_3-y_0)(y_3-y_1)(y_3-y_2)}x_3.$$

Подставляя табличные значения, получаем

$$L_3(y) = -0.00638y^3 + 1.01572y^2 - 0.01232y - 0.9976.$$

Таким образом форма входного сигнала

$$x(y) = -0.9976 - 0.01232y + 1.01572y^2 - 0.00638y^3$$

для  $y \in [-0.5; 0.25]$ .

#### 4.5. Интерполяционная формула Ньютона для произвольных значений аргумента

На практике встречаются таблицы с переменным шагом, например, такой характер имеют некоторые эмпирические данные. Обобщением понятия конечной разности являются разделенные разности.

Пусть функция  $y = f(x)$  задана таблично

$x$	$x_0$	$x_1$	...	$x_n$	...
$y$	$y_0$	$y_1$	...	$y_n$	...

Разделенные разности первого и второго порядка описываются соотношениями

$$f(x_i, x_{i+1}) = [x_i, x_{i+1}] = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, \quad (22)$$

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

Разделенные разности  $n$ -го порядка определяются рекуррентными соотношениями

$$f(x_i, x_{i+1}, \dots, x_{i+n}) = [x_i, x_{i+1}, \dots, x_{i+n}] = \frac{[x_{i+1}, \dots, x_{i+n}] - [x_i, \dots, x_{i+n-1}]}{x_{i+n} - x_i}, \quad (23)$$

$$n = 1, 2, \dots, \quad i = 0, 1, n-1.$$

Свойства разделенных разностей.

Разделенные разности являются симметричными функциями своих аргументов, т.е.  $[x_i, x_{i+1}] = [x_{i+1}, x_i]$ .

Разделенная разность  $(n+1)$  порядка для полинома  $P_n(x)$  тождественно равна нулю, т.е.  $[x, x_0, x_1, \dots, x_n] = 0$  для любого набора различных между собой чисел  $x, x_0, \dots, x_n$ .

Интерполяционная формула Ньютона для произвольных значений аргумента [23]

$$P_n(x) = y_0 + [x_0, x_1] \cdot (x - x_0) + [x_0, x_1, x_2] \cdot (x - x_0) \cdot (x - x_1) + \dots + [x_0, x_1, \dots, x_n] \cdot (x - x_0) \cdot (x - x_1) \cdot \dots \cdot (x - x_{n-1}) \quad (24)$$

Погрешность интерполяции может быть вычислена по формулам ((5)-(7)).

В других обозначениях формула (24) имеет вид

$$P_n(x) = y_0 + f(x_0, x_1) \cdot (x - x_0) + f(x_0, x_1, x_2) \cdot (x - x_0) \cdot (x - x_1) + \dots + f(x_0, x_1, \dots, x_n) \cdot (x - x_0) \cdot (x - x_1) \cdot \dots \cdot (x - x_{n-1}) \quad (24')$$

**Пример 5 [15].**

Функция  $y = f(x)$  задана таблично

Таблица 5

$x$	0.103	0.108	0.115	0.120
$y$	2.01284	2.03342	2.06070	2.07918

Определить значение функции при  $x=0.112$ .

*Решение.*

При вычислении ограничимся разделенными разностями первого и второго порядков.

Вычислим разделенные разности

Таблица 6

$x_i$	$y_i$	$f(x_i, x_{i+1})$	$f(x_i, x_{i+1}, x_{i+2})$
0.103	2.01284	4.116000	-18.238166
0.108	2.03342	3.896142	-16.761833
0.115	2.06070	3.696000	-14.788461
0.120	2.07918	3.503750	

Найдем  $f(0.112)$  двумя способами, принимая за  $x_0$  вначале  $x=0.103$ , а затем  $x=0.108$ ,

$$f(0.112) = 2.01284 + 4.116(0.112 - 0.103) + (-18.238166)(0.112 - 0.108)(0.112 - 0.108) = 2.04923,$$

$$f(0.112) = 2.03342 + 3.897142(0.112 - 0.108) + (-16.761833)(0.112 - 0.108)(0.112 - 0.115) = 2.04921.$$

Принимаем  $f(0.112) = (2.04922 + )/2 = 2.04922$ .

## 4.6. Сплайн - интерполяция

Сплайн-интерполяция есть специальный вид многоинтервальной интерполяции многочленами, при котором интерполирующий многочлен обеспечивает не только выполнение равенства (3), но и непрерывность заданного числа первых производных на границах частичных интервалов. Многоинтервальная интерполяция заключается в интерполяции  $f(x)$  в ряде частичных интервалов отдельными полиномами невысокой степени.

«Сплайн» - гибкая линейка. Кубические сплайн-функции представляют собой некоторую математическую модель гибкого тонкого стержня из упругого материала, закрепленного в двух точках (соседних узлах интерполяции). Между точками закрепления этот стержень примет некоторую форму, минимизирующую его потенциальную энергию. В общем случае сплайн задается глобальным способом, т.е. с использованием всех узлов при любом их расположении.

Для локального сплайна непрерывны  $f(x)$  и  $f'(x)$ .

Кубический сплайн, заданный локально, - это интерполирующая функция в виде полинома третьей степени, вычисляемая по формулам

$$y(x) = \frac{(x_{i+1} - x)^2(2(x - x_i) + h)}{h^3} y_i + \frac{(x - x_i)^2(2(x_{i+1} - x) + h)}{h^3} y_{i+1} + \frac{(x_{i+1} - x)^2(x - x_i)}{h^2} m_i + \frac{(x - x_i)^2(x - x_{i+1})}{h^2} m_{i+1}, \quad (25)$$

$$y_i = f(x_i), \quad i = \text{int} \frac{x - a}{h},$$

$x$  - точка интерполяции,

$a = x_0$  - начальное значение,

$h$  - расстояние между узлами,

$m_i, m_{i+1}$  - первые производные функции  $f(x)$ .

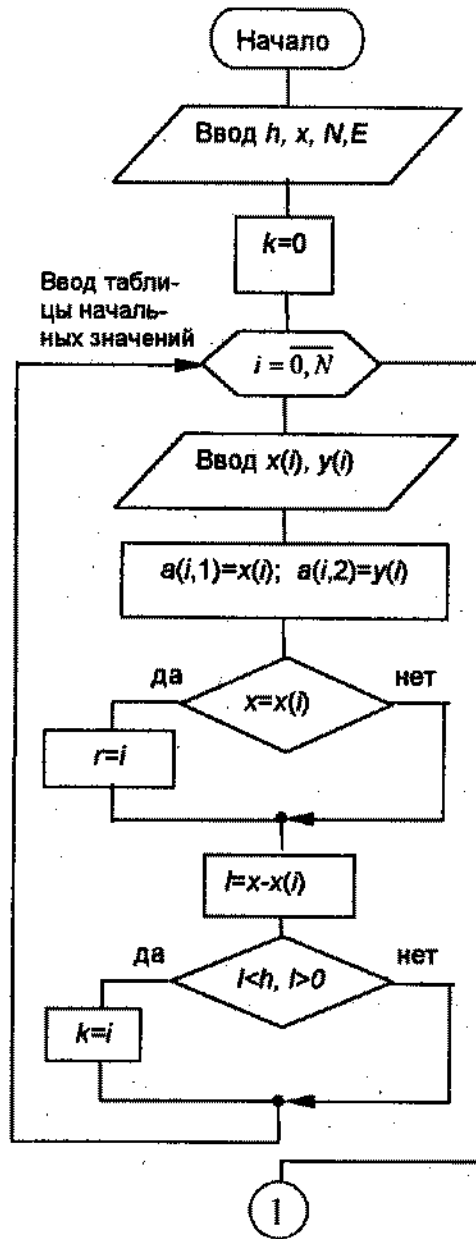
Производные могут вычисляться, например, с помощью формул численного дифференцирования по трем точкам.

$$m_i = \frac{y_{i+1} - y_{i-1}}{2h}, \quad i = 1, 2, \dots, n-1,$$

$$m_0 = \frac{4y_1 - y_2 - 3y_0}{2h}, \quad i = 0, \quad (26)$$

$$m_n = \frac{3y_n + y_{n-2} - 4y_{n-1}}{2h}, \quad i = n.$$

**Схема алгоритма интерполяции по формуле Ньютона для интерполяции в начале таблицы**



Входные параметры:  
 (N+1) – количество узлов;  
 E – требуемая точность интерполяции;  
 h – шаг таблицы аргументов;  
 x – значение аргумента для вычисления  $P_m(x)$   
 x() – массив аргументов;  
 y() – массив значений функции.  
 Выходные параметры:  
 $P_m(x)$  – значение интерполяционного многочлена при заданном

Проверка условия совпадения узла интерполяции с x

Нахождение ближайшего узла интерполяции

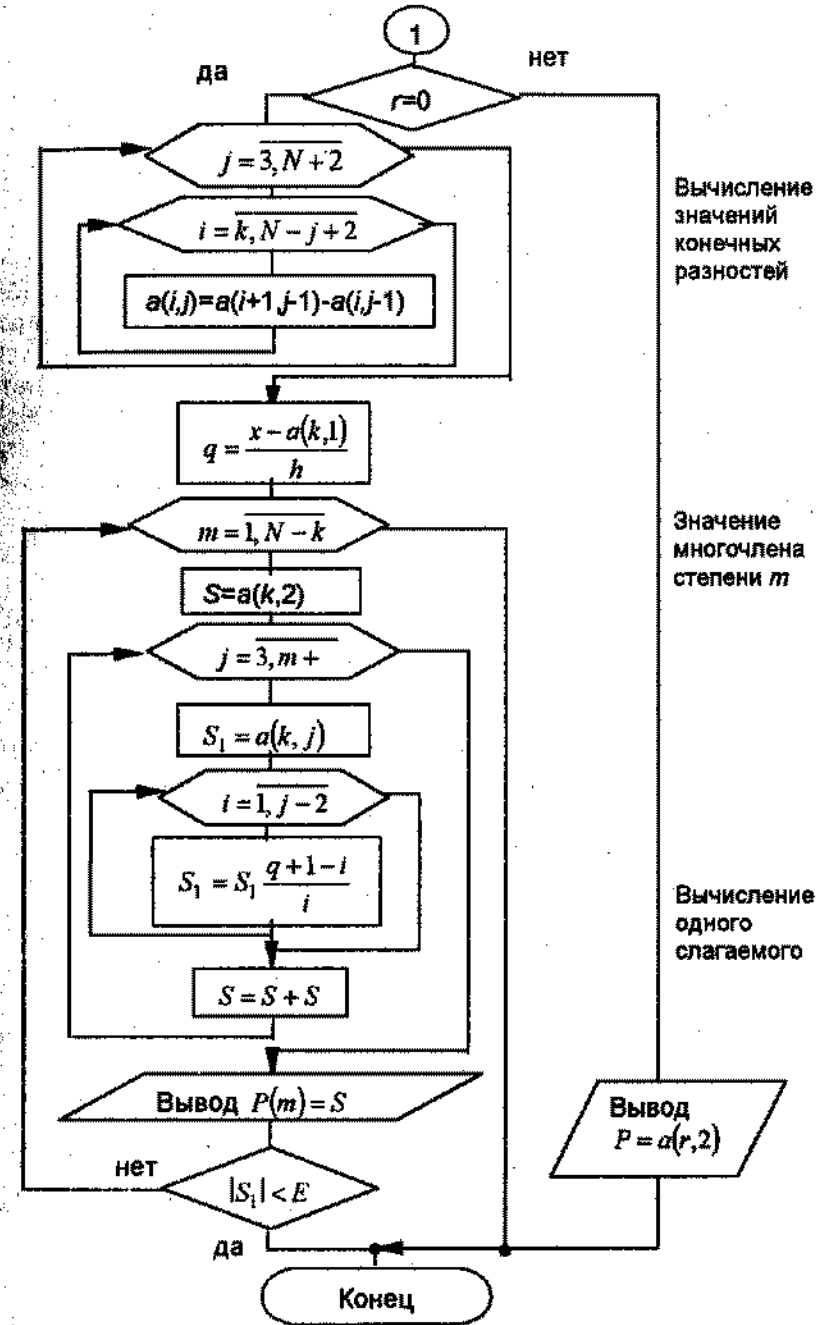


Рис. 4.3

Достоинством локально заданных сплайнов является описание свойств функции  $y(x)$  на каждом частичном интервале независимо от ее свойств на других интервалах. У обычной параболической аппроксимации изменение  $y(x)$  вблизи какого-либо узла оказывает влияние на ход всей кривой, что приводит к качественно неверному описанию некоторых физических явлений.

**Замечание**

Вывод формул (25) приведен в [13]. Отметим, что коэффициенты глобального кубического сплайна определяются из условий (3), условий непрерывности производных в узлах и условий на концах интервала. Вследствие этого коэффициенты сплайнов оказываются связанными системой линейных уравнений с трехдиагональной матрицей. Чаще всего такие системы решаются методом прогонки [58].

Согласно [47] оценки сходимости кубического сплайна, его первой и второй производных производятся в случае равномерной сетки ( $h = const$ ) по формулам

$$\begin{aligned} |f(x) - y(x)| &\leq \frac{5}{2} h^2 \omega(h, f''), \\ |f'(x) - y'(x)| &\leq 5h \omega(h, f'''), \\ |f''(x) - y''(x)| &\leq 5\omega(h, f'''). \end{aligned} \quad (27)$$

Величина

$\omega(h, \varphi) = \sup |\varphi(x') - \varphi(x'')|$  – максимальное изменение функции  $\varphi(x)$  на отрезке длины  $h$  в пределах  $[a, b]$ , где  $x', x'' \in [a, b], |x' - x''| \leq h$ .

**Пример 6.**

Провести аппроксимацию  $N$ -образной вольт-амперной характеристики туннельного диода, заданной в виде

Таблица 7

$x_i$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
$y_i$	0	10	4	2	1.5	2	5	9	13

в точках  $x = 0.1; x = 0.25$  кубическим сплайном.

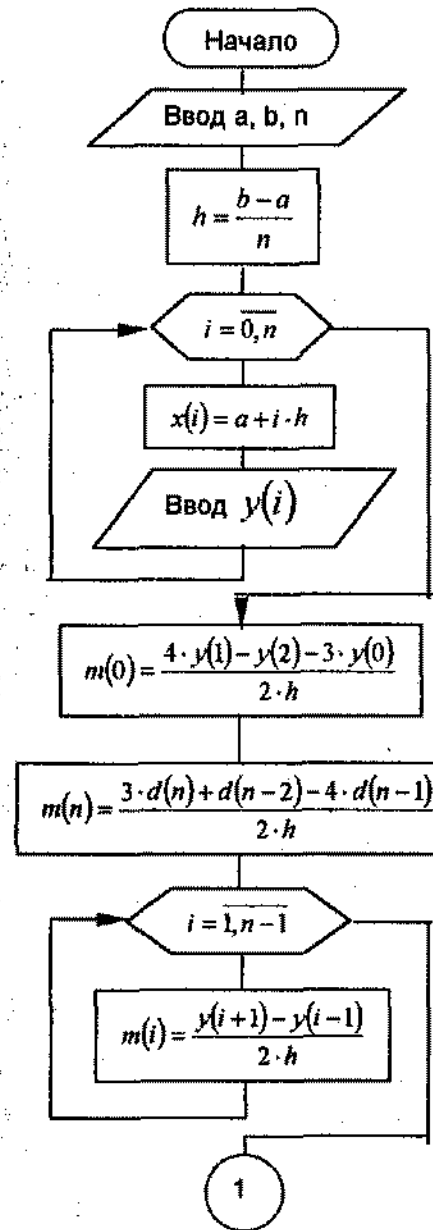
**Решение.**

Согласно формулам (25) и (26) вычислим  $y(0.1); y(0.25)$

Таблица 8

$x$	$i$	Используемые узлы				Значения в узлах				$m_i$	$m_{i+1}$	$y(x)$
		$x_{i-1}$	$x_i$	$x_{i+1}$	$x_{i+2}$	$y_{i-1}$	$y_i$	$y_{i+1}$	$y_{i+2}$			
0.1	1	0	0.1	0.2	0.3	0	10	4	2	20	-40	10
0.25	2	0.1	0.2	0.3	0.4	10	4	2	1.5	-40	-12.5	2.656

**Структурная схема сплайн-интерполяции**



Входные параметры:  
 $[a, b]$  – отрезок интерполяции;  
 $n$  – число узлов;  
 $x$  – заданное значение аргумента;  
 $y()$  – массив значений функции.  
 Выходные параметры:  
 $y$  – решение.

Формирование  
исходной  
таблицы

Вычисление  
первых  
производных

Рис. 4.5



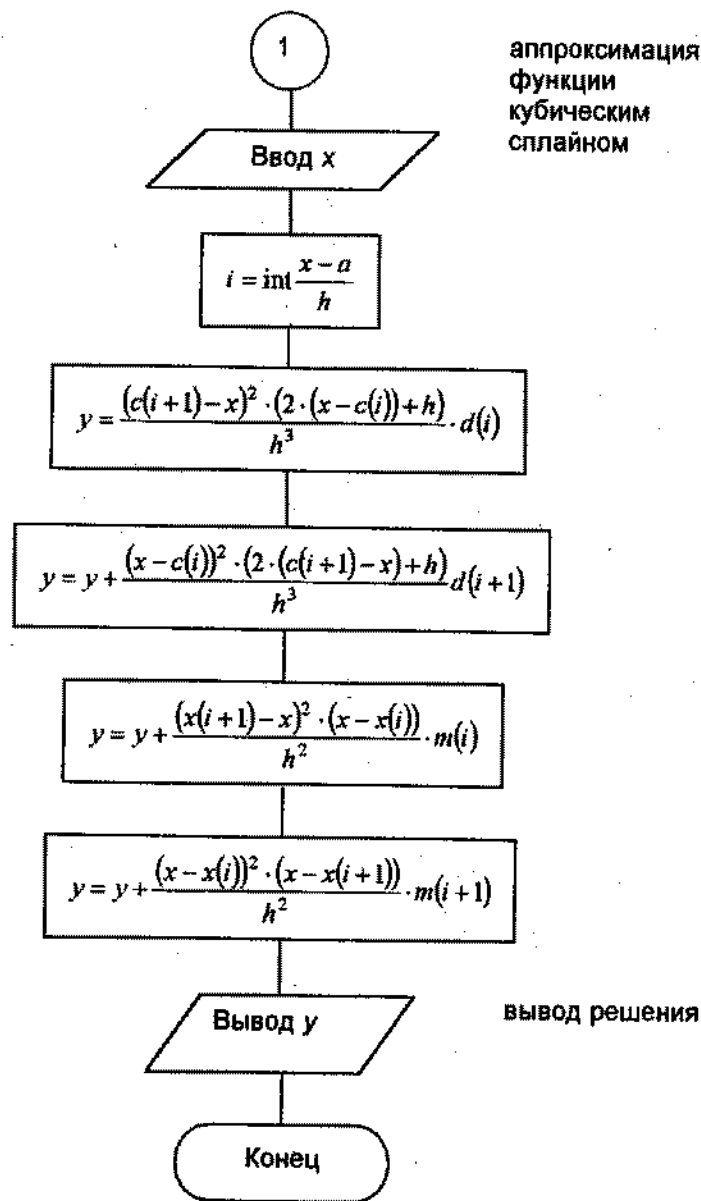


Рис. 4.6

#### 4.7. Метод наименьших квадратов

При интерполировании используются значения рассматриваемой функции в известных точках. В случае обработки опытных данных, полученных в результате наблюдений или вычисленных, нужно иметь в виду ошибки этих данных. Если набор экспериментальных данных получен со значительной погрешностью, то не имеет смысла использовать интерполяцию полиномами или сплайнами. В таком случае аппроксимирующую функцию  $\varphi(x)$  можно построить с помощью метода наименьших квадратов.

Наиболее распространен способ выбора функции  $\varphi(x)$  в виде

$$\varphi(x) = C_0 \varphi_0(x) + C_1 \varphi_1(x) + \dots + C_m \varphi_m(x), \quad (28)$$

$\varphi_0(x), \varphi_1(x), \dots, \varphi_m(x)$  - базисные функции,  $m \leq n$ ,

$\vec{C}^T = (C_0, C_1, \dots, C_m)$  - вектор коэффициентов, определяемых из условия минимизации

$$S = \sum_{i=0}^n \varepsilon_i^2 = \sum_{i=0}^n [\varphi(x_i) - f(x_i)]^2 \rightarrow \min, \quad (29)$$

$(n+1)$  - количество узловых точек.

Если погрешность исходных данных  $E$ , то количество базисных функций выбирается так, чтобы  $\sqrt{S} \approx E$ .

Условие минимума функции  $S$  приводит к системе линейных уравнений относительно параметров  $C_0, C_1, \dots, C_m$ . Система называется системой нормальных уравнений, ее матрица - матрицей Грама. Элементы матрицы Грама являются скалярными произведениями базисных функций

$$(\varphi_j, \varphi_k) = \sum_{i=0}^n \varphi_j(x_i) \varphi_k(x_i).$$

Выбор конкретных базисных функций зависит от свойств аппроксимируемой функции, таких, как свойство симметрии, периодичности и т. п.

Доказано [23], что если функции  $\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x)$  линейно-независимы на  $[a, b]$ , то единственное решение соответствует наименьшему квадратичному отклонению.

Один из возможных базисов - степенной:

$$\varphi_0(x) = 1, \quad \varphi_1(x) = x, \dots, \quad \varphi_m(x) = x^m.$$

Степень аппроксимирующего полинома  $m \ll n$ .

В случае степенного базиса матрица Грама

$$G = \begin{pmatrix} n+1 & \sum_{i=0}^n x_i & \sum_{i=0}^n x_i^2 & \dots & \sum_{i=0}^n x_i^m \\ \sum_{i=0}^n x_i & \sum_{i=0}^n x_i^2 & \sum_{i=0}^n x_i^3 & \dots & \sum_{i=0}^n x_i^{m+1} \\ \dots & \dots & \dots & \dots & \dots \\ \sum_{i=0}^n x_i^m & \sum_{i=0}^n x_i^{m+1} & \sum_{i=0}^n x_i^{m+2} & \dots & \sum_{i=0}^n x_i^{2m} \end{pmatrix}, \quad (30)$$

столбец правых частей системы нормальных уравнений

$$G \cdot \bar{C} = \bar{B}$$

$$\bar{B} = \begin{pmatrix} \sum_{i=0}^n y_i \\ \sum_{i=0}^n x_i y_i \\ \dots \\ \sum_{i=0}^n x_i^m y_i \end{pmatrix}. \quad (31)$$

Если  $m \leq 5$ , то такие системы можно решать, например, методом Гаусса. Специальный метод решения - метод сингулярного разложения [70].

В качестве меры уклонения заданных значений функции  $y_0, y_1, \dots, y_n$  от многочлена степени  $m$

$$\varphi(x) = C_0 \varphi_0(x) + C_1 \varphi_1(x) + \dots + C_m \varphi_m(x)$$

принимается величина

$$\rho = \sqrt{\frac{1}{n+1} \sum_{i=0}^n [\varphi_m(x_i) - f(x_i)]^2}. \quad (32)$$

$(n+1)$  - количество узлов,  $m$  - степень аппроксимирующего полинома,  $n+1 \geq m$ .

Система линейных уравнений для нахождения коэффициентов многочлена  $\varphi(x) = a_0 + a_1 x$  (линейная аппроксимация):

$$\begin{cases} (n+1) a_0 + (\sum_{i=0}^n x_i) a_1 = \sum_{i=0}^n y_i, \\ (\sum_{i=0}^n x_i) a_0 + (\sum_{i=0}^n x_i^2) a_1 = \sum_{i=0}^n x_i y_i. \end{cases}$$

Введем обозначения  $\bar{x} = \frac{\sum_{i=0}^n x_i}{n+1}$ ,  $\bar{y} = \frac{\sum_{i=0}^n y_i}{n+1}$  - средние значения

исходных данных.

Решением системы уравнений являются

$$a_0 = \bar{y} - a_1 \bar{x},$$

$$a_1 = \frac{\sum_{i=0}^n x_i y_i - (n+1) \bar{x} \bar{y}}{\sum_{i=0}^n x_i^2 - (n+1) (\bar{x})^2}.$$

### Пример 7.

Аппроксимировать следующие данные многочленом второй степени, используя метод наименьших квадратов.

Таблица 9

x	0.78	1.56	2.34	3.12	3.81
y	2.50	1.20	1.12	2.25	4.28

Решение.

Запишем в таблицу 10 элементы матрицы Грама и столбец свободных членов

Таблица 10

	$x^0$	$x^1$	$x^2$	$x^3$	$x^4$	y	xy	$x^2 y$
1	0.78	0.608	0.475	0.370	2.50	1.950	1.520	
1	1.56	2.434	3.796	5.922	1.20	1.872	2.920	
1	2.34	5.476	12.813	29.982	1.12	2.621	6.133	
1	3.12	9.734	30.371	94.759	2.25	7.020	21.902	
1	3.81	14.516	55.306	210.72	4.28	16.307	62.129	
$\Sigma$	5	11.61	32.768	102.76	341.75	11.35	29.770	94.604

Система нормальных уравнений

$$\begin{cases} 5C_0 + 11.61C_1 + 32.768C_2 = 11.350, \\ 11.61C_0 + 32.768C_1 + 102.761C_2 = 29.770, \\ 32.768C_0 + 102.761C_1 + 341.750C_2 = 94.604. \end{cases}$$

Решение системы:

$$C_0 = 5.022; \quad C_1 = -4.014; \quad C_2 = 1.002.$$

Искомая аппроксимирующая функция

$$\varphi(x) = 5.022 - 4.014x + 1.002x^2.$$

Сравним исходные значения для y со значениями  $\bar{y}$ , полученными путем вычисления аппроксимирующего многочлена.

Таблица 11

$x$	$y$	$\tilde{y} = \varphi(x)$
0.78	2.50	2.501
1.56	1.20	1.199
2.34	1.12	1.117
3.12	2.25	2.255
3.81	4.28	4.278

Вычислим невязку  $\rho = \sqrt{\frac{1}{n+1} \sum_{i=0}^n [\varphi_m(x_i) - f(x_i)]^2}$ ,  $\rho = 0.001218$ .

Рассмотрим одну из классических задач радиотехники: на входе безынерционного нелинейного устройства с характеристикой преобразования  $y = f(x)$  действует сигнал  $x(t)$ ; требуется определить спектр выходного сигнала  $y(x)$ .

Характеристику  $y = f(x)$  аппроксимируем полиномом:

$$y \approx \varphi_m(x) = C_0 + C_1 x + \dots + C_n x^m.$$

Входной сигнал задается в виде:

$$x(t) = \sum_{k=1}^q \alpha_k \cos \omega_k t + \beta_k \sin \omega_k t,$$

Далее по известным формулам вычисляются амплитуды спектральных составляющих выходного сигнала  $y(t)$  с частотами  $\omega_k$ , а также с кратными и комбинаторными частотами.

При аппроксимации, прежде всего, возникает вопрос о качестве приближения. В данной задаче необходимо аппроксимировать так, чтобы выполнялось условие близости вычисленных по  $\varphi_m(x)$  и  $\varphi_m^*(x)$  спектров, если два полинома  $\varphi_m(x)$  и  $\varphi_m^*(x)$  обеспечивают малые (и при этом близкие) расстояния до  $f(x)$ , т.е.

$$\rho = \sqrt{\frac{1}{n+1} \sum_{i=1}^n [\varphi_m(x_i) - f(x_i)]^2}$$

- среднеквадратичное отклонение - невязки. Однако, как показывает практика, задача вычисления спектра на основе среднеквадратичной степенной аппроксимации является неустойчивой.

Фактор «структурного дефекта аппроксимации» приводит к заметным различиям пересчитанных спектров в высокочастотной области.

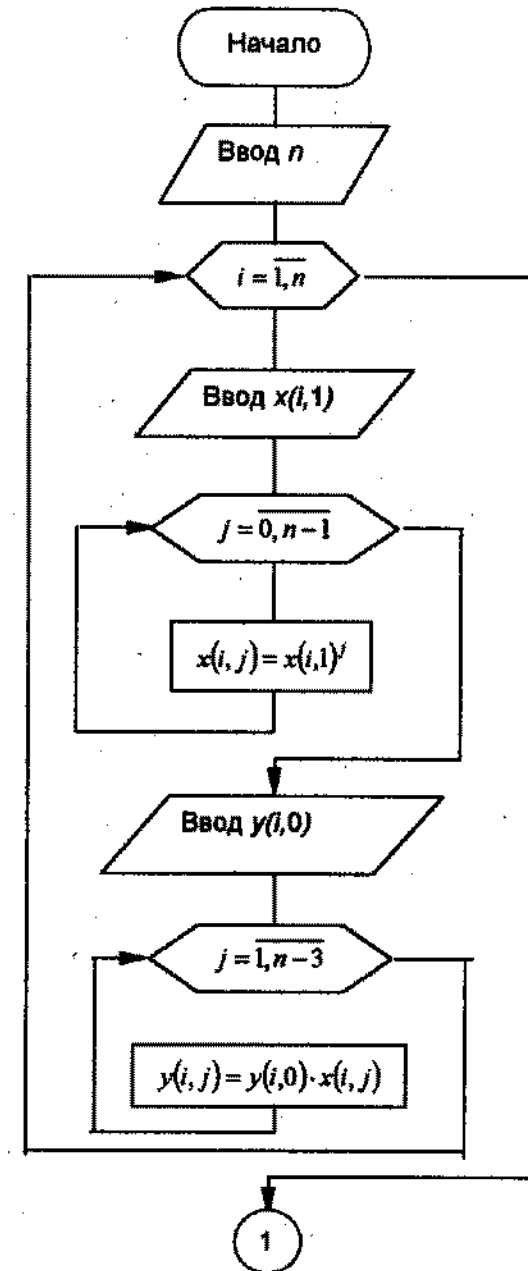


Рис.4.7

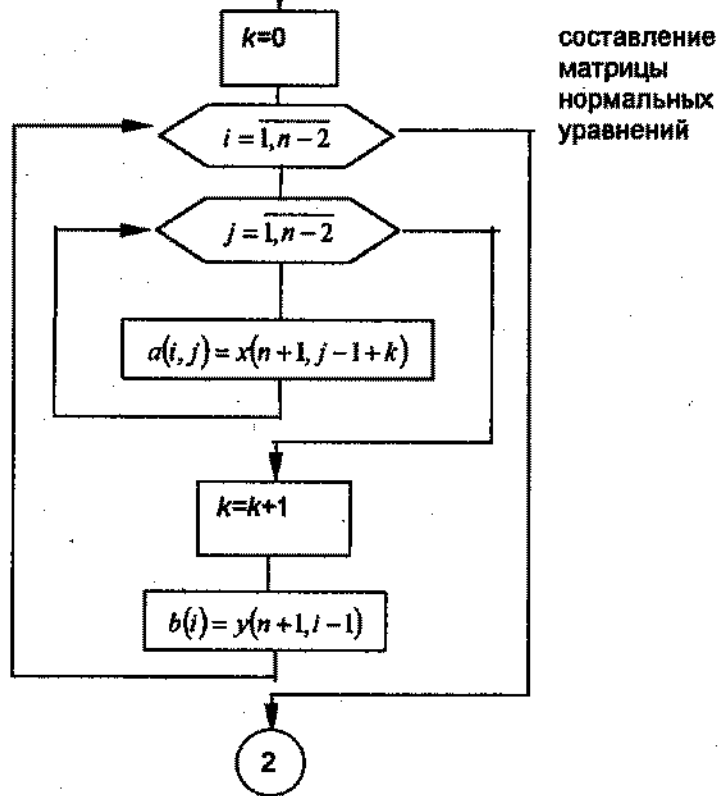
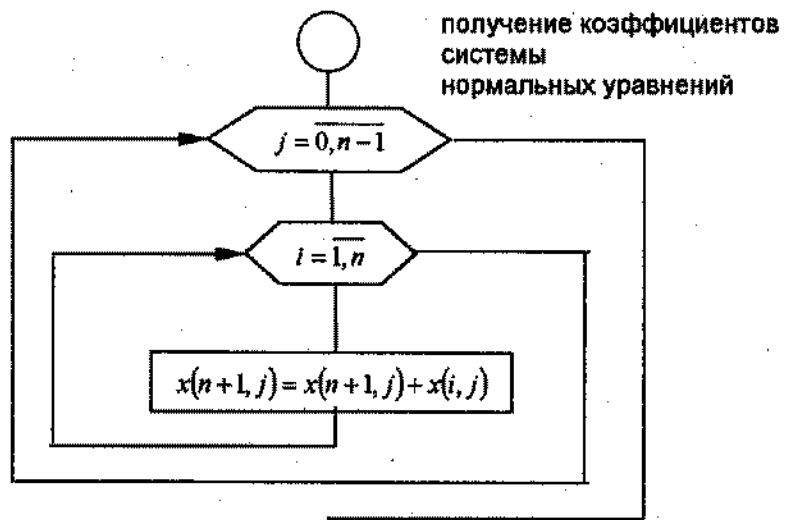


Рис. 4.8

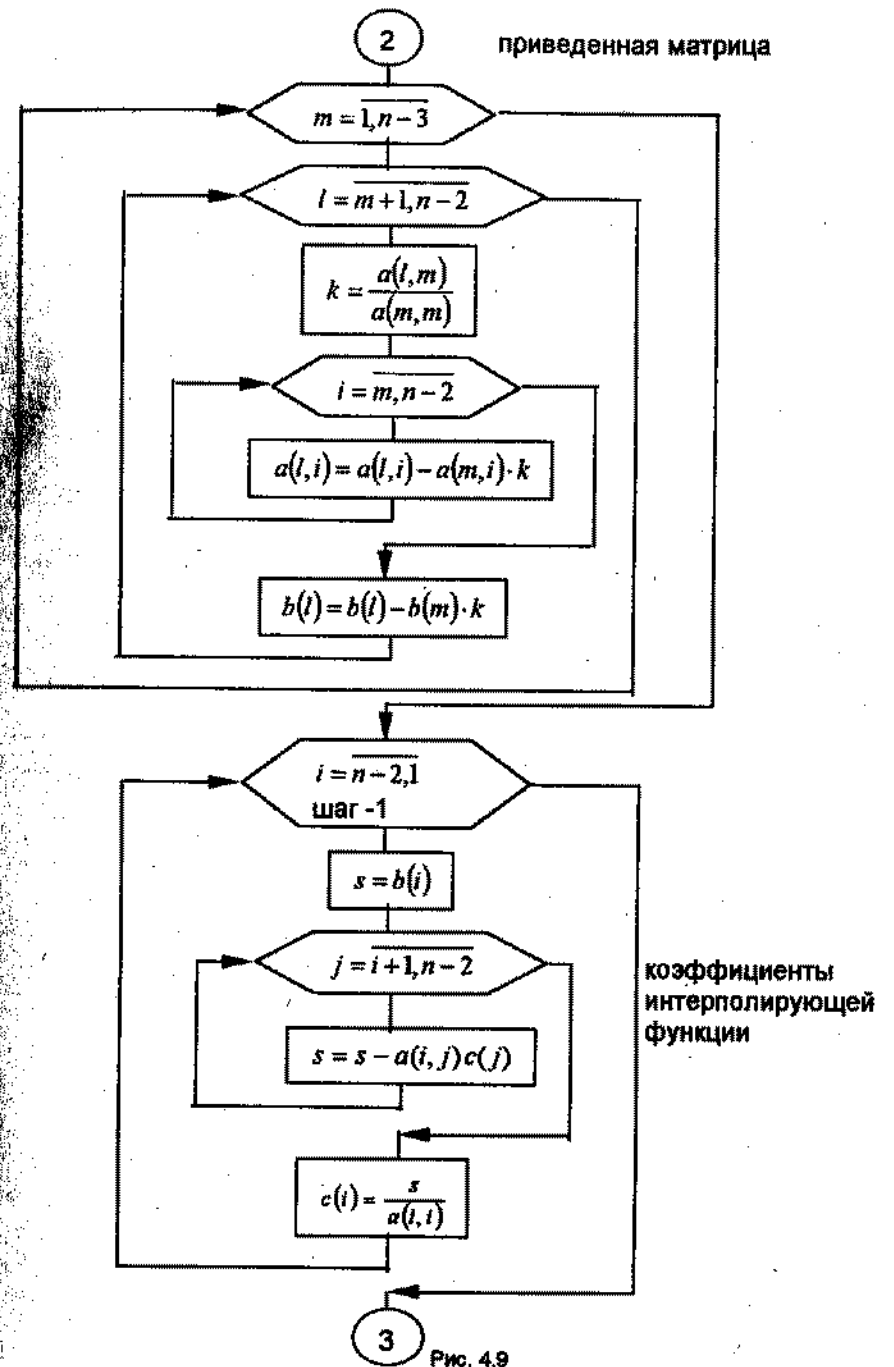


Рис. 4.9

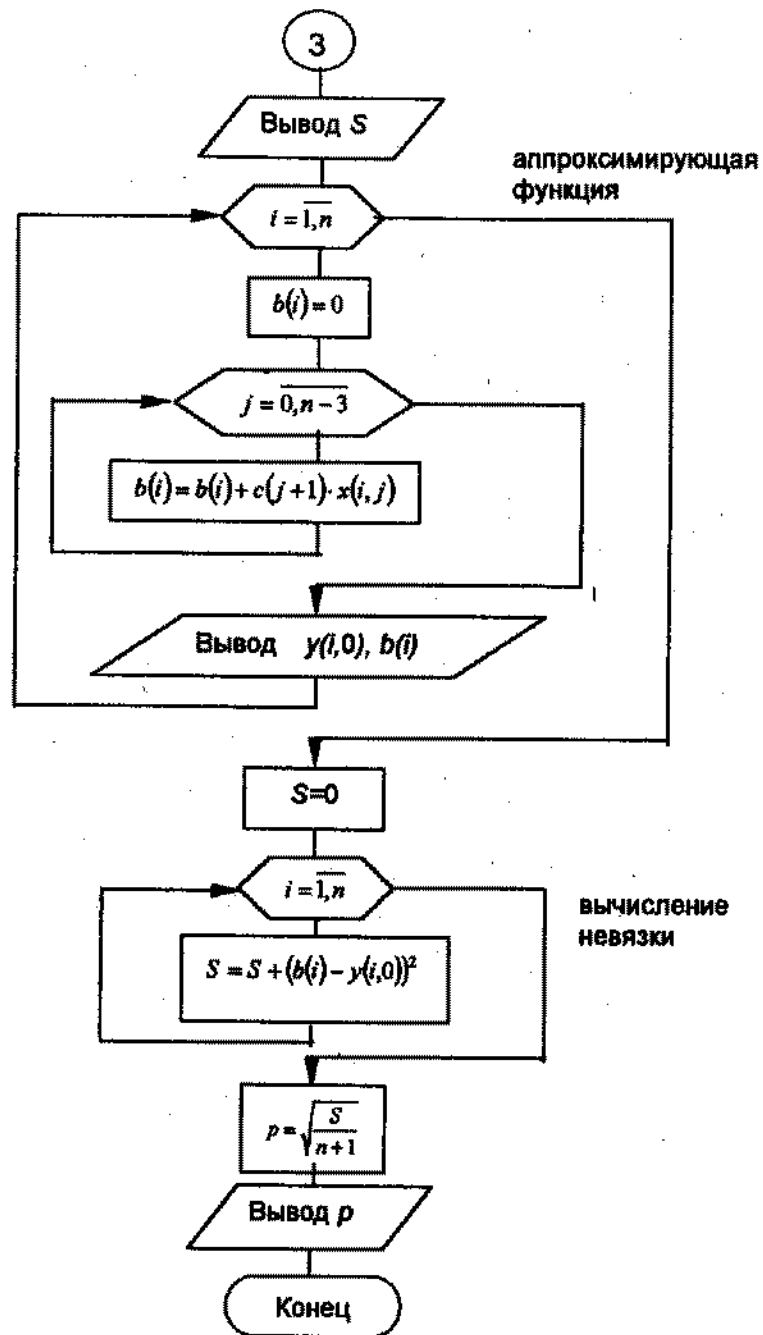


Рис. 4.10

По формальному критерию за счет повышения степени полинома точность должна повышаться, но фактически усиливается «структурный дефект». Кроме того, при увеличении  $m$ , из-за погрешности округления и быстрого возрастания неустраняемой погрешности (погрешность, возникшая за счет неточного математического описания физической задачи, и не точного округления параметров), может возникнуть и вычислительная неустойчивость.

**Пример 8.** [20]

Требуется аппроксимировать динамическую вольт-амперную характеристику диода, описываемую функцией

$$y = \sqrt{1+x}, \quad (-0.9 \leq x \leq 0.9).$$

**Решение.**

Построим полином среднеквадратичного приближения при различных  $m$  и  $n$ .

Результаты приведены в таблицах 1,2,3 (1<sup>й</sup> набор коэффициентов), соответственно для случаев:

$$m = 3, \quad n = 40;$$

$$m = 10, \quad n = 20;$$

$$m = 17, \quad n = 40.$$

Обнаруживается, что при высоких  $n$  коэффициенты при старших степенях  $x$  вовсе не оказываются малыми, как хотелось бы ожидать, а распределяются весьма причудливо. В тоже время невязки относительно невелики: если в первом случае  $\rho \approx 9 \cdot 10^{-3}$ , то во втором случае  $\rho \approx 6 \cdot 10^{-5}$ , в третьем -  $\rho = 4 \cdot 10^{-4}$ .

Наблюдаются также различия у коэффициентов с одинаковыми номерами, причем различия растут с увеличением номера.

Какая из причин в данном случае преобладает - структурный дефект или вычислительная неустойчивость, - пока не ясно.

Проследим на простом примере за картиной в спектральной области. Пусть  $x(t) = 0.9 \sin t$ , используем аппроксимацию при  $v = 17, n = 40$ . Спектры сигналов  $y(t) = \sqrt{1 + 0.9 \sin t}$  и  $\tilde{y}(t) = Q_{17}(0.9 \sin t)$  были с высокой точностью вычислены (вычислительная погрешность не превосходила 0.01%). Оказалось, что высшие гармоники во втором случае превосходят по амплитуде гармоники истинного сигнала соответственно в 6 раз (17-я гармоника), в 18 раз (16-я), в 40 раз (15-я), в 37 раз (14-я), в 8 раз (13-я); у низших гармоник различия меньше.

Таким образом, налицо заметные искажения спектра.

Прием, изложенный в [20], позволяет приближенно оценить, насколько результат искажен при вычислениях. Пусть

$C = \{C_0, C_1, \dots, C_n\}$  - искомый набор коэффициентов, доставляющий невязке (32) точный минимум,  $\tilde{C}$  - приближение, численный результат, полученный в результате квадратичной аппроксимации.

Восстановим функцию в узлах  $x = x_m$  по вычисленному

$$\text{набору } \tilde{C}: \tilde{f}(x_n) = \tilde{\varphi}_m(x_n) = \sum_{i=0}^m \tilde{C}_i x_i.$$

Обычно погрешность восстановления (вычисления) функции много меньше погрешности решения системы. Повторяя вычисление набора коэффициентов для  $\tilde{f}(x_n)$ , получим набор  $\tilde{\tilde{C}}$ , отличный от  $\tilde{C}$ . Если расстояние между  $\tilde{C}$  и  $\tilde{\tilde{C}}$  (в выбранном случае) велико, то качество вычислений  $\tilde{C}$  надо считать неудовлетворительным. Если же  $\tilde{C}$  и  $\tilde{\tilde{C}}$  близки, то для проверки вычислим еще 1-2 набора. «Тесное» расположение последовательных наборов позволяет предположить удовлетворительное качество вычислений.

Таблица 12

Номер коэффициентов	1-й набор коэффициентов	2-й набор коэффициентов	3-й набор коэффициентов
0	1.0051180	1.0051180	1.0051180
1	0.4875174	0.4875175	0.4875176
2	-0.1751661	-0.1751661	-0.1751661
3	0.1171231	0.1171229	0.1171228

Степень полинома - 3, число узлов - 40.

Таблица 13

Номер коэффициентов	1-й набор коэффициентов	2-й набор коэффициентов	3-й набор коэффициентов
0	1.0000170	1.0000290	1.0000330
1	0.5003389	0.5003294	0.5003298
2	-0.1264650	-0.1271612	-0.1274461
3	0.0540756	0.0542342	0.0542300
4	-0.0193915	-0.0124250	-0.0096288
5	0.0827680	0.0820488	0.0820164
6	-0.1142736	-0.1387242	-0.1484048
7	-0.1195585	-0.1183593	-0.1183724
8	.1736646	0.2086069	0.2223018
9	0.1282029	0.1275413	0.1275452
10	-0.1478974	-0.1652972	-0.1720637

Степень полинома - 10, число узлов - 20.

Таблица 14

Номер коэффициентов	1-й набор коэффициентов	2-й набор коэффициентов	3-й набор коэффициентов
0	1.0003020	0.9983827	1.0079940
1	0.4982780	0.5044467	0.7491205
2	-0.1596688	0.0574859	-1.0330460
3	0.1659968	-0.1826036	1.3681570
4	0.5936997	-3.3669830	16.5928900
5	-1.5553800	3.5789750	-20.0224000
6	-4.1465240	21.9906700	-110.4528000
7	10.0443400	-21.4071900	125.2539000
8	11.4294600	-63.8486700	321.8299000
9	-31.1155200	62.5713700	-375.0067000
10	-10.9273000	73.3259400	-374.0315000
11	48.7922400	-88.2483500	541.1609000
12	-8.3926220	11.3695300	-47.3426500
13	-33.9124700	44.5772400	-284.1445000
14	22.6178800	-86.7570300	482.2936000
15	3.0028310	16.6733900	-94.2605100
16	-11.4592500	47.9148000	-238.7573000
17	4.9230760	-17.8461500	108.6153000

Степень полинома - 17, число узлов - 40.

В табл. 12 - 14 столбцы (2 правых) и представляют собой последовательные наборы, вычисленные для  $y = \sqrt{1+x}$ .

В первом случае ( $m=3, n=40$ ) качество вычислений хорошее, о плохом качестве приближения ( $m=17, n=40$ ) можно судить уже по второму набору.

Другие способы уточнения вычислений при использовании метода наименьших квадратов приведены в [24].

#### 4.8. Ортогональные функции. Ряд Фурье

В случае, если степень аппроксимирующего полинома велика, методы отыскания такого полинома громоздки.

В этих случаях предпочтительнее методы, основанные на идее ортогонализации функции.

• Система интегрируемых функций  $\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x)$  (33) называется ортогональной на  $[a, b]$ , если

$$(\varphi_m, \varphi_n) = \int_a^b \varphi_m(x) \cdot \varphi_n(x) dx = 0 \quad \text{при } m \neq n. \quad (34)$$

• Число

$$\|\varphi_m(x)\| = \sqrt{(\varphi_m, \varphi_m)} = \sqrt{\int_a^b \varphi_m^2(x) dx} \quad (35)$$

называется нормой функции на  $[a, b]$ .

- Система называется ортонормированной, если нормы всех функций (33) равны 1.
- Если система ортогональных функций такова, что для любой непрерывной функции  $f(x)$  справедливо соотношение  $\lim_{m \rightarrow \infty} I_m = 0$ , то эта система называется полной.

Таким образом, для полной ортогональной системы

$$\sum_{i=0}^{\infty} C_i^2 \|\varphi_i(x)\|^2 = \|f(x)\|^2, \quad (36)$$

для ортонормированной системы

$$\sum_{i=0}^{\infty} C_i^2 = \|f(x)\|^2 \quad (37)$$

- равенство Парсеваля или условие полноты.

Для ортонормированной системы выполняется условие

$$\int_a^b \varphi_m(x) \varphi_n(x) dx = \delta_{mn}, \quad (38)$$

$$\delta_{mn} = \begin{cases} 0, & \text{при } m \neq n, \\ 1, & \text{при } m = n. \end{cases}$$

Следуя (28) будем искать полином в виде

$$\varphi(x) = C_0 \varphi_0(x) + C_1 \varphi_1(x) + \dots + C_m \varphi_m(x) = \sum_{i=0}^m C_i \varphi_i(x) \quad (39)$$

вводя дополнительные требования (33).

В случае ортогональной системы из необходимого условия минимума функции

$$I_m = \int_a^b \left[ \sum_{i=0}^m C_i \varphi_i(x) - f(x) \right]^2 dx, \quad (40)$$

следует

$$C_i = \frac{\int_a^b f(x) \varphi_i(x) dx}{\|\varphi_i(x)\|^2}, \quad i = \overline{0, m}. \quad (41)$$

- Коэффициенты  $C_i$  называются коэффициентами Фурье функции  $f(x)$  относительно выбранной ортогональной системы  $\{\varphi_i(x)\}$ .

Для ортонормированной системы

$$C_i = \int_a^b f(x) \varphi_i(x) dx, \quad i = \overline{0, m}. \quad (42)$$

Доказано [24] утверждение: обобщенный полином (39) с коэффициентами Фурье данной функции обладает наименьшим квадратичным отклонением от этой функции.

Свойства полинома:

- При увеличении числа слагаемых  $m$  младшие коэффициенты  $C_i$  остаются наименьшими;
- При увеличении  $m$  квадратичная погрешность (40) монотонно убывает, т. е.  $I_1 \geq I_2 \geq \dots \geq I_m \geq I_{m+1} \geq \dots$

$$\text{Система } 1, \sin x, \cos x, \sin 2x, \cos 2x, \dots, \sin nx, \cos nx, \dots \quad (43)$$

ортогональна на любом отрезке, длиной  $2\pi$ , например  $[-\pi, \pi]$ .

Нормы функций, входящих в систему

$$\begin{aligned} \|1\| &= \sqrt{\int_{-\pi}^{\pi} 1^2 dx} = \sqrt{2\pi}, \\ \|\sin nx\| &= \sqrt{\int_{-\pi}^{\pi} \sin^2 nx dx} = \sqrt{\pi}, \\ \|\cos nx\| &= \sqrt{\int_{-\pi}^{\pi} \cos^2 nx dx} = \sqrt{\pi}. \end{aligned} \quad (44)$$

#### 4.9. Ряд Фурье для периодической функции

Для непрерывной периодической функции с периодом  $2\pi$  можно составить тригонометрический полином

$$\varphi_n = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx). \quad (45)$$

Коэффициенты

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx, \\ b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx, \end{aligned} \quad (k = \overline{0, n}) \quad (46)$$

называются тригонометрическими коэффициентами Фурье функции  $f(x)$ , полином (45) - тригонометрическим интегральным полиномом Фурье.

При  $n \rightarrow \infty$  получаем для функции  $f(x)$  ряд Фурье

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx). \quad (47)$$

Коэффициенты (46) чаще всего вычисляются приближенно численными методами интегрирования.

Укажем без доказательства [53] условия разложимости функции в ряд Фурье.

Функция  $f(x)$  на интервале  $(-\pi, \pi)$  удовлетворяет условиям Дирихле:

1) если она имеет на этом интервале конечное число точек разрыва 1-го рода ( $\lim_{x \rightarrow x_0-0} f(x) \neq \lim_{x \rightarrow x_0+0} f(x)$ ),

2) если интервал  $(-\pi, \pi)$  можно разбить на конечное число частей так, чтобы в каждой из этих частей функция изменялась монотонно, т. е. имела бы конечное число точек экстремума,

3) если существуют конечные предельные значения функции  $f(-\pi+0)$  и  $f(+\pi-0)$ .

В случае выполнения условий Дирихле сумма ряда (47) равна:

1)  $f(x)$  во всех точках интервала  $(-\pi, \pi)$ , где  $f(x)$  непрерывна,

2)  $\frac{1}{2} [f(x-0) + f(x+0)]$  во всех точках разрыва,

3)  $\frac{1}{2} [f(-\pi+0) + f(\pi-0)]$  на концах интервала.

В случае, если  $f(x)$  - четная функция, её ряд Фурье

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx, \quad (48)$$

для нечетной функции ряд Фурье содержит только синусы

$$f(x) = \sum_{n=1}^{\infty} b_n \sin nx. \quad (49)$$

#### Пример 9.

Разложить в ряд Фурье функцию  $f(x) = x$ ,  $(-\pi < x < \pi)$ .

**Решение.**

Условия Дирихле выполнены, поэтому разложение возможно.

На интервале  $(-\pi, \pi)$   $f(x)$  - нечетная, следовательно, разложение будет содержать только синусы:

$$b_n = \frac{2}{\pi} \int_0^{\pi} x \sin nx dx.$$

Интегрируя по частям, получим

$$b_n = (-1)^{n+1} \frac{2}{n}.$$

Итак,  $x = 2 \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\sin nx}{n}$ . Во всех внутренних точках интервала  $(-\pi, \pi)$  равенство верно. Вне этого интервала ряд изображает

периодическое продолжение рассматриваемой функции. В точках разрыва  $(\pm\pi, \pm 3\pi, \dots)$  сумма ряда равна

$$s = \frac{f(\pi-0) + f(-\pi+0)}{2}.$$

Найдем эти пределы, например, в точке  $x = \pi$ .

$$\lim_{x \rightarrow \pi-0} f(x) = \lim_{x \rightarrow \pi-0} x = \pi;$$

$$\lim_{x \rightarrow -\pi+0} f(x) = \lim_{x \rightarrow -\pi+0} x = -\pi.$$

Таким образом,  $s = \frac{\pi - \pi}{2} = 0$ , тогда разложение имеет вид:

$$2 \left( \frac{\sin x}{1} - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \dots \right) = \begin{cases} x, & \text{если } -\pi < x < \pi; \\ 0, & \text{если } x = (2k+1)\pi, k \in \mathbb{Z}. \end{cases}$$

#### 4.10. Ряд Фурье для произвольного периода

Ряд Фурье для функции  $f(x)$  (период  $T = 2l$  ( $l \neq \pi$ )), удовлетворяющей условиям Дирихле:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left( a_n \cos \frac{n\pi x}{l} + b_n \sin \frac{n\pi x}{l} \right), \quad (50)$$

$$a_n = \frac{1}{l} \int_{-l}^l f(x) \cos \frac{n\pi x}{l} dx,$$

$$b_n = \frac{1}{l} \int_{-l}^l f(x) \sin \frac{n\pi x}{l} dx. \quad (51)$$

#### Пример 10.

Разложить в ряд Фурье функцию

$$f(x) = \begin{cases} 0, & \text{при } -2 \leq x \leq 0, \\ x, & \text{при } 0 \leq x \leq 2. \end{cases}$$

**Решение.**

Здесь  $l = 2$ , по формулам (51) вычислим  $a_n$  и  $b_n$ .



$$a_n = \frac{1}{2} \int_{-2}^2 f(x) \cos \frac{\pi x}{2} dx = \frac{1}{2} \left[ \int_{-2}^0 0 \cdot \cos \frac{\pi x}{2} dx + \int_0^2 x \cos \frac{\pi x}{2} dx \right] =$$

$$= \frac{1}{2} \int_0^2 x \cos \frac{\pi x}{2} dx = \frac{2}{\pi^2 n^2} [(-1)^n - 1].$$

$$a_n = \frac{2}{\pi^2 n^2} [(-1)^n - 1], \quad n = 1, 2, 3, \dots \quad a_0 = \frac{1}{2} \int_0^2 x dx = 1.$$

$$b_n = \frac{1}{2} \int_{-2}^2 f(x) \sin \frac{\pi x}{2} dx = \frac{1}{2} \int_0^2 x \sin \frac{\pi x}{2} dx = -\frac{2}{\pi n} (-1)^n.$$

Получено разложение

$$f(x) = \frac{1}{2} + \frac{2}{\pi^2} \sum_{n=1}^{\infty} \frac{(-1)^n - 1}{n^2} \cos \frac{\pi x}{2} - \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin \frac{\pi x}{2}.$$

#### 4.11. Интеграл Фурье

Интеграл Фурье – преобразование разложения функции в ряд Фурье при неограниченном расширении интервала  $(-l; l)$ , т. е. при  $l \rightarrow \infty$ .

Достаточные условия представимости функции интегралом Фурье, формулирует

##### Теорема [53].

Если  $f(x)$  удовлетворяет следующим условиям:

- 1) определена при  $-\infty < x < \infty$ ,
- 2) является кусочно-гладкой на всяком конечном промежутке,
- 3) абсолютно интегрируема в промежутке  $-\infty < x < \infty$ ,

(т. е.  $\int_{-\infty}^{\infty} |f(x)| dx$  сходится), то возможно интегральное представление функции в каждой точке ее непрерывности:

$$f(x) = \int_0^{\infty} [a(\alpha) \cos(\alpha \cdot x) + b(\alpha) \sin(\alpha \cdot x)] d\alpha, \quad (52)$$

$$a(\alpha) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \cos(\alpha \cdot t) dt; \quad b(\alpha) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(x) \sin(\alpha \cdot t) dt. \quad (53)$$

##### Пример 11.

$$\text{Функцию } f(x) = \begin{cases} 0, & \text{при } |x| > 1, \\ 1, & \text{при } 0 < x < 1, \\ -1, & \text{при } -1 < x < 0 \end{cases}$$

представить интегралом Фурье.

*Решение.*

Функция  $f(x)$  является нечетной, поэтому

$$a(\alpha) = 0, \quad b(\alpha) = \frac{2}{\pi} \int_0^{\infty} f(t) \sin(\alpha \cdot t) dt = \frac{2}{\pi} \cdot \frac{1 - \cos \alpha}{\alpha}.$$

Интеграл Фурье для данной функции

$$f(x) = \int_0^{\infty} b(\alpha) \sin(\alpha \cdot x) d\alpha = \frac{2}{\pi} \int_0^{\infty} \frac{1 - \cos \alpha}{\alpha} \sin(\alpha \cdot x) d\alpha = \frac{4}{\pi} \int_0^{\infty} \sin^2 \frac{\alpha}{2} \sin(\alpha \cdot x) \frac{d\alpha}{\alpha}$$

#### 4.12. Комплексная форма записи ряда Фурье

Преобразуем (46) и (47) с помощью формул Эйлера:

$$\cos nx = \frac{e^{inx} + e^{-inx}}{2}, \quad \sin nx = \frac{e^{inx} - e^{-inx}}{2i}.$$

Ряд Фурье для функции  $f(x)$ , записанный в форме

$$\sum_{-\infty}^{\infty} C_n e^{inx}, \quad (54)$$

$$C_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx \quad (55)$$

называется рядом Фурье в комплексной форме.

##### Пример 12.

Записать разложение функции  $f(x) = x$  в интервале  $(-\pi; \pi)$  в ряд Фурье в комплексной форме.

*Решение.*

Вычислим коэффициенты  $C_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} x e^{-inx} dx$ . Интегрируя по частям

$$C_n = \frac{1}{2\pi} \left[ \frac{e^{-inx}}{(-in)^2} \cdot (-inx - 1) \right]_{-\pi}^{\pi} = \frac{1}{2\pi} \left[ \frac{e^{-inx}}{n^2} \cdot (inx + 1) \right]_{-\pi}^{\pi} =$$

$$= \frac{1}{2\pi} \left[ \frac{e^{inx}}{n^2} \cdot (in\pi + 1) - \frac{e^{-inx}}{n^2} \cdot (-in\pi + 1) \right] = \frac{1}{2\pi n^2} \left( 2i\pi n \frac{e^{inx} + e^{-inx}}{2} - 2i \frac{e^{inx} - e^{-inx}}{2i} \right) =$$

$$= \frac{1}{2n^2 \pi} (2i\pi n \cos n\pi - 2i \sin n\pi) = \frac{i}{n} \cos n\pi = \frac{i}{n} (-1)^n \quad \text{при } n \neq 0,$$

$$C_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} x dx = 0,$$

получаем разложение:  $x = \sum_{-\infty}^{\infty} \frac{i(-1)^n}{n} e^{inx}$ .

#### 4.13. Интеграл Фурье в комплексной форме

Комплексная форма интеграла Фурье - результат преобразования (52) и (53) с помощью формул Эйлера [53].

$$f(x) = \int_{-\infty}^{\infty} C(\alpha) e^{i\alpha x} d\alpha, \quad (56)$$

$$C(\alpha) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-i\alpha t} dt \quad (57)$$

или

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t) e^{i\alpha(x-t)} dt d\alpha. \quad (58)$$

Перепишем (58) в виде

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \left[ \int_{-\infty}^{\infty} f(t) e^{-i\alpha t} dt \right] \cdot e^{i\alpha x} d\alpha.$$

Положим

$$\varphi(\alpha) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\alpha t} dt, \text{ или } \varphi(\alpha) = \frac{1}{\sqrt{2\pi}} \left( 2\pi \frac{d(\alpha)}{2} - i b(\alpha) \right), \quad (59)$$

тогда

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(\alpha) e^{i\alpha x} d\alpha. \quad (60)$$

Функция  $\varphi(\alpha)$  называется преобразованием Фурье, ее также называют спектральной характеристикой функции  $f(x)$ .

Модуль спектральной характеристики  $\varphi(\alpha)$  называется спектром функции  $f(x)$ .

Спектральные методы расчета (гармонический синтез и анализ) являются мощным средством вычисления количественных характеристик процессов, устройств и систем связи. Теоретической базой методов является дискретное преобразование Фурье (ряд Фурье).

Спектром временной зависимости  $y(t)$  называется совокупность ее гармонических составляющих (гармоник), образующих ряд Фурье.

Запишем ряд Фурье для  $y(t)$  в виде:

$$y(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} M_n \cos(n\Omega t + \varphi_n), \quad (61)$$

$$M_n = \sqrt{a_n^2 + b_n^2} \text{ - амплитуда,} \quad (62)$$

$$\varphi_n = -\arctg \frac{b_n}{a_n} \text{ - фаза гармоник,} \quad (63)$$

$$\Omega = \frac{2\pi}{T}, \quad T \text{ - период функции } y(t), \quad \omega = n\Omega.$$

Коэффициенты Фурье (51) примут вид:

$$a_n = \frac{2}{T} \int_0^T y(t) \cos(n\Omega t) dt, \quad (64)$$

$$b_n = \frac{2}{T} \int_0^T y(t) \sin(n\Omega t) dt.$$

Спектральный анализ периодической функции заключается в вычислении коэффициентов  $a_n, b_n$  ряда Фурье.

Известно, что периодический сигнал  $y(x)$  имеет дискретный (линейный) частотный спектр

$$C_n = a_n - i b_n = \frac{2}{T} \int_0^T y(t) e^{-in\Omega t} dt, \quad (65)$$

$C_n$  - комплексная амплитуда спектральной составляющей с частотой  $\omega = n\Omega$ .

Спектральный анализ непериодических функций, т. е. функций, полностью определенных на отрезке  $[0; t_0]$ , заключается в вычислении составляющих их комплексной спектральной плотности

$$s(i\omega) = s_1(\omega) + i s_2(\omega) = s(\omega) e^{i\varphi(\omega)}, \quad (66)$$

где модуль спектральной плотности

$$s(\omega) = |s(i\omega)| = \sqrt{s_1(\omega)^2 + s_2(\omega)^2}, \quad (67)$$

фаза на частоте  $\omega$

$$\varphi(\omega) = \arctg(s(i\omega)) = -\arctg \frac{s_2(\omega)}{s_1(\omega)}. \quad (68)$$

При этом

$$s_1(\omega) = \int_0^{t_0} y(t) \cos(\omega t) dt, \quad s_2(\omega) = \int_0^{t_0} y(t) \sin(\omega t) dt. \quad (69)$$

В случае непериодического неделимого сигнала составляющие спектральной плотности определяются несобственными интегралами (с бесконечными пределами).

Численный спектральный анализ заключается в нахождении коэффициентов  $a_0, b_0, a_1, b_1, \dots, a_k, b_k$  (или  $M_1, M_2, \dots, M_k, \varphi_1, \dots, \varphi_k$ ) для периодической функции  $y(x)$ , заданной на отрезке  $[0; T]$  дискретными отсчетами с шагом  $\Delta t = \frac{T}{N}$ . При этом интегралы в выра-

жениях для коэффициентов Фурье (64) берутся численными методами.

Интерпретируя периодические функции как частный случай непериодических, можно вычисление параметров  $a_n$ ,  $b_n$  (64) и параметров  $s_1$ ,  $s_2$  (69) объединить в одну задачу. Интегралы в формулах (64) и (69) вычисляются методами численного интегрирования, например, по формулам прямоугольников. Тогда (64) принимают вид

$$a_n = \frac{2}{N} \sum_{i=1}^{N-1} y_i \cos(2\pi n f_1 i \Delta t), \quad b_n = \frac{2}{N} \sum_{i=1}^{N-1} y_i \sin(2\pi n f_1 i \Delta t), \quad (70)$$

где

$$\Delta t = \frac{T}{N} - \text{шаг отсчетов } y(x),$$

$N$  - количество отсчетов на одном периоде,

$$f_1 = \frac{1}{T} - \text{частота первой гармоники,}$$

$n$  - номер гармоники.

Аналогично, формулы (69) запишутся так:

$$s_1 = \Delta t \sum_{i=1}^{N-1} y_i \cos(2\pi f \Delta t i), \quad s_2 = \Delta t \sum_{i=1}^{N-1} y_i \sin(2\pi f \Delta t i). \quad (71)$$

Перепишем (70), (71) с учетом  $t_0 = T$ ,  $f = n f_1$ .

$$A_{n0} = \frac{s_1}{\Delta t} = \frac{a_n N}{2} = \sum_{i=0}^{N-1} y_i \cos(2\pi f i \Delta t), \quad (72)$$

$$B_{n0} = \frac{s_2}{\Delta t} = \frac{b_n N}{2} = \sum_{i=0}^{N-1} y_i \sin(2\pi f i \Delta t).$$

Рассматриваются две схемы спектрального анализа: последовательный и параллельный спектральный анализ. В последнем случае параллельно вычисляются  $N$  гармоник, в первом - одна гармоника на шаге.

Алгоритм параллельного спектрального анализа приведен в [26]. Детальный анализ указанных схем [71] не отдает предпочтения ни одной из них, выбор схемы определяется конкретными задачами.

При большом количестве отсчетов и значений спектра обе

схемы мало эффективны (большой объем вычислений, велико время обработки).

Разработана наиболее экономичная вычислительная процедура - «быстрое преобразование Фурье».

Практическая реализация метода сопряжена с довольно значительными трудностями, однако существуют программы «быстрой схемы» для  $k = N+1 < 2^{11}$ . [3]

Обсудим некоторые аспекты использования тригонометрической аппроксимации [20].

Обычные методы численного интегрирования пригодны уменьшать  $\Delta t$ . Эмпирический способ оценки  $\Delta t$  изложен в [21].

- Выбор дискретных значений частоты.

Если не существует (или не указаны) каких-то физических соображений относительно выбора дискретных значений частоты, то выбирают равноотстоящие значения с некоторым шагом  $\Delta \Omega$ .

Согласно теории В. А. Котельникова, спектр функции, имеющей ограниченную длительность  $T$ , может быть с любой точностью воспроизведен при помощи чисел, следующих через интервалы

$$\Delta \Omega = \frac{2\pi}{T}. \quad (73)$$

- Вычисление больших массивов синуса и косинуса.

Время вычисления синуса и косинуса по стандартным программам (их количество равно  $k \cdot (N+1)$ ) существенно влияет на общее время решения задачи. Кроме того, функции косинуса и синуса периодичны по индексам  $k$  и  $N$ . Если пользоваться рекуррентными формулами

$$\sin(u + \Delta u) = \sin u \cos \Delta u + \sin \Delta u \cos u, \quad (74)$$

$$\cos(u + \Delta u) = \cos u \cos \Delta u - \sin u \sin \Delta u,$$

время вычисления уменьшается (в 5-7 раз).

При вычислении длинных последовательностей тригонометрических функций можно через какое-то число итераций проверять вычисленные по (74) значения либо стандартной программой, либо проверяя условие

$$|\cos^2 u + \sin^2 u - 1| \leq \delta, \quad \delta - \text{точность.}$$

Вычисление разрывных спектров - отдельная задача. Методы решения этой задачи обсуждены в [20].

**Пример 13.**

Приведем в качестве примера проблему расчета частотных характеристик по известным воздействию и реакции.

**Решение.**

Пусть  $x(t)$  и  $y(t)$  - соответственно входной и выходной сигналы,  $S(i\omega)$  и  $Q(i\omega)$  - их спектры. Амплитудно- и фазо-частотные характеристики представляют собой модуль и аргумент комплексного коэффициента передачи, связывающего спектры воздействия и реакции:

$$k(i\omega) = k(\omega)e^{i\varphi(\omega)} = \frac{Q(i\omega)}{S(i\omega)}$$

Вещественные и мнимые составляющие спектра вычисляются, как указано выше (69), затем вычисляются

$$k(\omega_m) = \frac{|Q(i\omega_m)|}{|S(i\omega_m)|}, \quad \varphi(\omega_m) = \text{arctg} \frac{Q(i\omega_m)}{S(i\omega_m)} - \omega_m t_0, \quad (75)$$

$t_0$  - время запаздывания  $y(t)$  относительно  $x(t)$ , если эта величина неизвестна, принимаем ее равной нулю и получаем по формулам (75) значения криволинейной составляющей фазо-частотной характеристики.

**Пример 14.**

Одним из важнейших практических приложений спектрального анализа является расчет амплитудно-частотной (АЧХ) и фазо-частотной (ФЧХ) характеристик линейной системы по ее переходной характеристике (ПХ)  $a(t)$ .

$$A(i\omega) = a(0) + A_1(\omega) + i \cdot A_2(\omega),$$

$$A_1(\omega) = \int_0^{t_0} a'(t) \cos(\omega \cdot t) dt,$$

$$A_2(\omega) = \int_0^{t_0} a'(t) \sin(\omega \cdot t) dt.$$

Амплитудно-частотная характеристика  $\left( f = \frac{\omega}{2\pi} \right)$

$$A(f) = \sqrt{A_1^2(\omega) + A_2^2(\omega)},$$

фазо-частотная характеристика

$$\varphi(f) = -\text{arctg} \frac{A_2(f)}{A_1(f)},$$

$a'(t)$  - производная ПХ  $a(t)$ ,  $a(0)$  - значение  $a(t)$  при  $t = 0$

**Схема последовательного спектрального анализа**

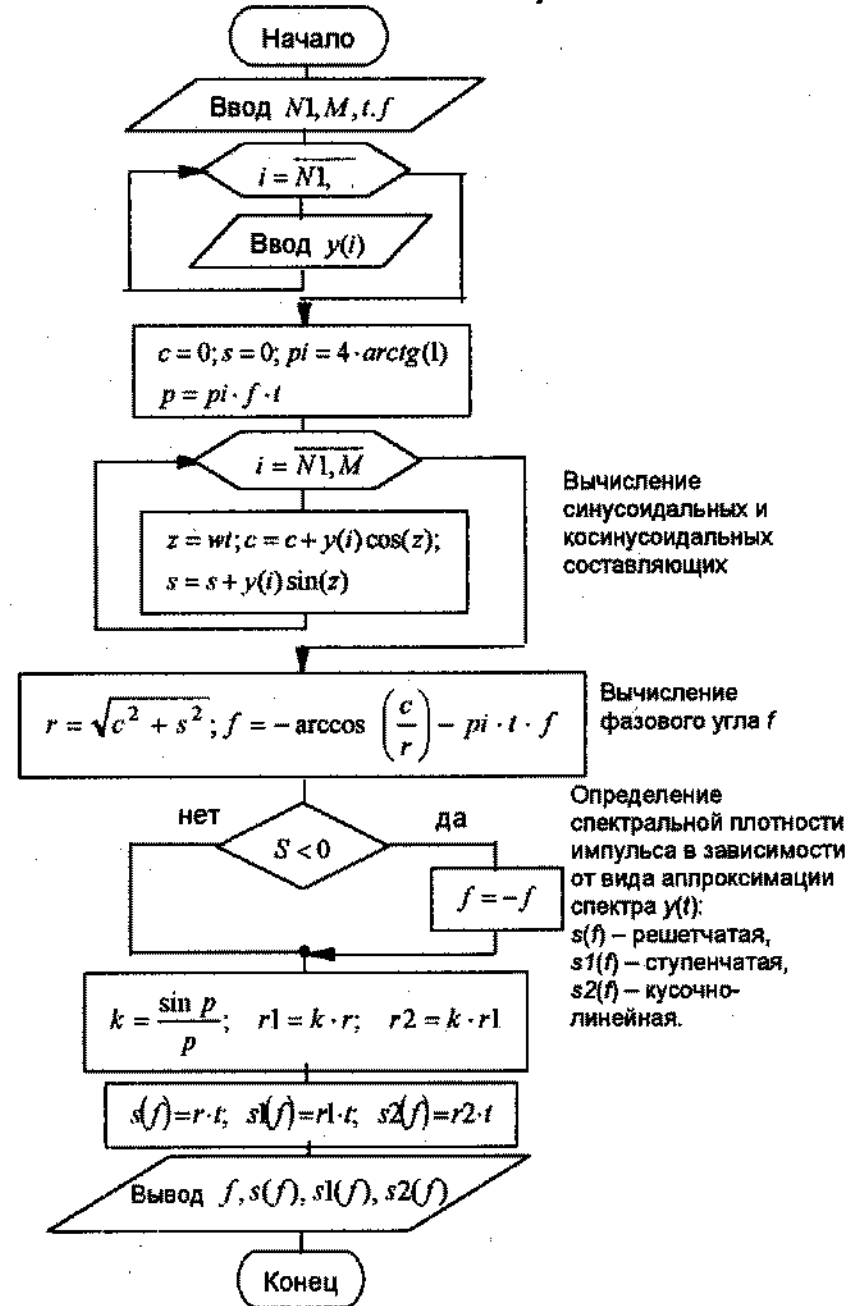


Рис. 4.11

Численный метод расчета АЧХ и ФЧХ по ПХ основывается на аппроксимации  $a(t)$  в интервале между ее отсчетами.

$$a'(t) \approx \frac{\Delta a(t)}{\Delta t} = \frac{a_k - a_{k-1}}{\Delta t}$$

Для вычисления  $A_1(\omega)$  и  $A_2(\omega)$  применяются методы численного интегрирования с использованием отсчетов ПХ  $a_k = a(t_k)$  для заданных  $f$  и  $\Delta t$ . Затем вычисляются амплитудно-частотная характеристика  $A(f)$  и фазо-частотная характеристика  $\varphi(f)$ .

#### 4.14. Многомерная интерполяция. Интерполяционные формулы Эрмита

В технических и физических приложениях встречаются задачи, приводящие к многомерной интерполяции. Не останавливаясь подробно, укажем лишь, что достаточно широко распространены двумерные таблицы, несколько реже – трехмерные. Свойства и методы решения указанных задач приведены в [32], [47].

Сформируем задачу, называемую задачей интерполирования с кратными узлами.

Пусть для функции  $y = f(x)$  известны значения в узлах  $x_1, x_2, \dots, x_m$  самой функции и всех ее производных до порядка  $t_j - 1$ .

$$f(x_j), f'(x_j), \dots, f^{(t_j-1)}(x_j), \quad j = 1, 2, \dots, m, \quad (76)$$

$t_j$  – кратность узла  $x_j$ ,  $t_1 + t_2 + \dots + t_m = n + 1$ .

Задача построения многочлена  $P_n(x)$  степени не выше  $n$ , удовлетворяющего условиям

$$P_n^{(s)}(x_j) = f^{(s)}(x_j), \quad j = \overline{1, m}, \quad s = \overline{0, t_j - 1} \quad (77)$$

называется интерполированием Эрмита и рассмотрена в [50], [8], [32].

Частный случай задачи (77).

Пусть имеется единственный узел  $x_1$ , т.е.  $m = 1$ .

Известны

$$f(x_1), f'(x_1), \dots, f^{(q-1)}(x_1), \quad q = n + 1.$$

Тогда

$$P_n(x) = \sum_{k=0}^n \frac{f^{(k)}(x_1)}{k!} (x - x_1)^k. \quad (78)$$

$P_n(x)$  – ряд Тейлора функции  $f(x)$ .

## 5. ЧИСЛЕННЫЕ МЕТОДЫ ИНТЕГРИРОВАНИЯ

Задача численного интегрирования заключается в вычислении интеграла  $\int_a^b f(x) dx$  посредством ряда значений подынтегральной функции  $y = f(x)$ .

Численные методы условно можно сгруппировать по способу аппроксимации подынтегральной функции.

Методы Ньютона - Котеса основаны на аппроксимации функции  $f(x)$  полиномом степени  $n$ . Алгоритмы этого класса отличаются только степенью полинома. Как правило, узлы аппроксимирующего полинома – равноотстоящие.

Методы сплайн-интегрирования базируются на аппроксимации функции  $f(x)$  сплайном – кусочным полиномом.

В методах наивысшей алгебраической точности (метод Гаусса) используются специально выбранные неравноотстоящие узлы, обеспечивающие минимальную погрешность интегрирования при заданном (выбранном) количестве узлов.

Методы Монте-Карло используются чаще всего при вычислении кратных интегралов, узлы выбираются случайным образом, ответ носит вероятностный характер.

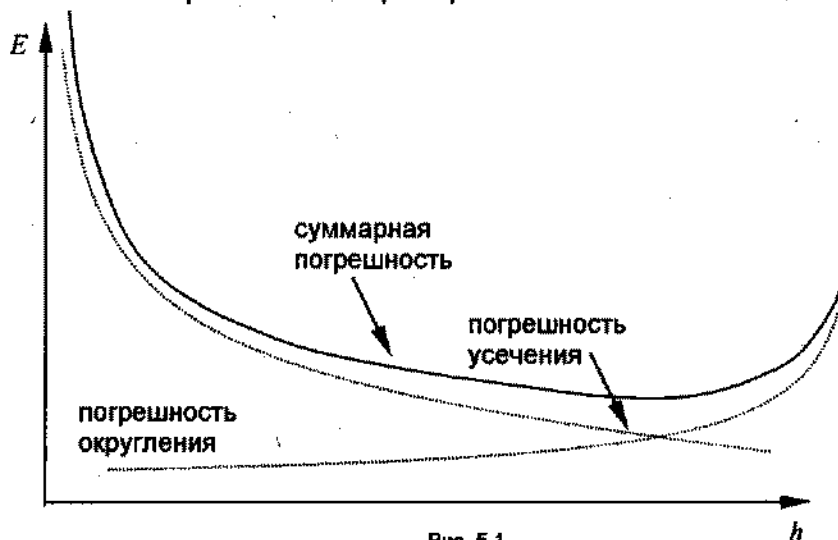


Рис. 5.1

Независимо от выбранного метода в процессе численного интегрирования необходимо вычислить приближенное значение интеграла и оценить погрешность. Погрешность уменьшается при увеличении  $n$  - количества разбиений отрезка  $[a, b]$ . Однако при этом возрастает погрешность округления за счет суммирования значений интегралов, вычисленных на частичных отрезках.

Погрешность усечения зависит от свойств подынтегральной функции и длины  $h$  частичного отрезка.

### 5.1. Методы Ньютона - Котеса

Пусть  $y = f(x)$  - непрерывная функция на отрезке  $[a, b]$ . Требуется вычислить  $\int_a^b f(x) dx$ . Разобьем отрезок интегрирования на  $n$  равных частей точками  $x_0 = a < x_1 < x_2 < \dots < x_n = b$  так, что  $x_{i+1} - x_i = \frac{b-a}{n} = h$ , ( $i = 0, n$ ) и пусть  $y_i = f(x_i)$  - значения  $f(x)$  в точках деления.

#### 5.1.1. Формулы прямоугольников

На отрезке  $[x_i, x_{i+1}]$  возьмем произвольную точку  $C$ . Интерполяционным многочленом в случае одного узла является  $f(C)$ .

$$\int_{x_i}^{x_{i+1}} f(x) dx = h \cdot f(C), \quad h = x_{i+1} - x_i. \quad (1)$$

Геометрическая интерпретация формулы прямоугольников: (1) - площадь, ограниченная графиком функции  $y = f(x)$  на отрезке  $[x_i, x_{i+1}]$ , осью  $Ox$ , ординатами  $f(x_i)$  и  $f(x_{i+1})$  заменяется площадью прямоугольника с высотой, равной  $f(C)$ .

Общий вид формулы прямоугольников:

$$\int_a^b f(x) dx = h \cdot \sum_{k=0}^{n-1} f(a + kh). \quad (2)$$

Отметим частные случаи формулы (2).

1) Если узел  $a$  совпадает с левым концом частичного промежутка интегрирования, формулу называют формулой левых прямоугольников ( $\alpha = a$ ).

2) Если  $\alpha = x_{i+1}$  - формулу называют формулой правых прямоугольников ( $\alpha = a + h$ ).

3) Если  $\alpha = \frac{x_i + x_{i+1}}{2}$  - середина отрезка. Тогда формула называется формулой средних прямоугольников ( $\alpha = a + \frac{h}{2}$ ).

Погрешности методов левых, правых и средних прямоугольников:

$$R_n(f, a) = \frac{(b-a)^2}{2n} f'(x^*), \quad x^* \in [a, b], \quad (3)$$

$$R_n(f, a+h) = -\frac{(b-a)^2}{2n} f'(x^*),$$

$$R_n(f, a + \frac{h}{2}) = \frac{(b-a)^3}{24n^2} f''(x^*).$$

$$f'(x^*) = \max f'(x^*), \quad f''(x^*) = \max f''(x), \quad x, x^* \in [a, b].$$

(Метод средних прямоугольников - метод второго порядка.)

#### 5.1.2. Формула трапеций

Заменяем на отрезке  $[x_i, x_{i+1}]$  подынтегральную функцию интерполяционным многочленом первой степени (формула (13), раздел 4) и проинтегрируем этот многочлен.

$$\text{Тогда } \int_{x_i}^{x_{i+1}} f(x) dx = \int_{x_i}^{x_{i+1}} (y_i + \frac{\Delta y_i}{2} \frac{x - x_i}{h}) dx = \frac{h}{2} (y_i + y_{i+1}).$$

Общая формула в случае линейной интерполяции называется формулой трапеций и имеет вид:

$$\int_a^b f(x) dx = \frac{h}{2} (y_0 + 2(y_1 + y_2 + \dots + y_{n-1}) + y_n). \quad (4)$$

Погрешность формулы трапеций (погрешность усечения)

$$R = -\frac{(b-a)^3}{12n^2} f''(x^*), \quad f''(x^*) = \max f''(x), \quad x, x^* \in [a, b]. \quad (5)$$

Метод трапеций - метод второго порядка.

Структурная схема алгоритма метода трапеций - на рис.3.

На элементарном отрезке  $[x_i, x_{i+1}]$  площадь под кривой  $y = f(x)$  полагают равной площади трапеции с основаниями  $y_i, y_{i+1}$  и высотой  $h = x_{i+1} - x_i$ .

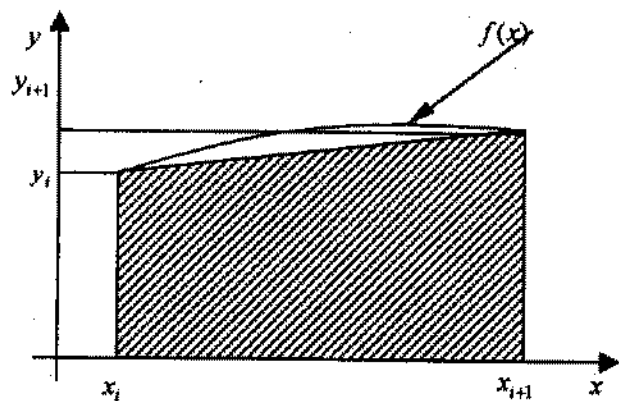


Рис. 5.2

### 5.1.3. Формула Симпсона

Если для каждой пары отрезков  $[x_i; x_{i+2}]$  построить многочлен второй степени, затем проинтегрировать его и воспользоваться свойством аддитивности интеграла, то получим формулу Симпсона:

$$\int_a^b f(x) dx = \frac{h}{3} [y_0 + y_{2n} + 4(y_1 + y_3 + \dots + y_{2n-1}) + 2(y_2 + y_4 + \dots + y_{2n-2})], \quad (6)$$

$$h = \frac{b-a}{2n}. \quad (7)$$

Погрешность формулы Симпсона - метода четвертого порядка

$$R = -\frac{(b-a)^5}{180(2n)^4} f^{(IV)}(x^*), \quad x^* \in [a; b] \quad (8)$$

Метод Симпсона позволяет получить высокую точность, если  $f^{(IV)}(x)$  не слишком велика. В противном случае метод второго порядка может дать большую точность.

Например, для функции  $f(x) = -25x^4 + 45x^2 - 7$  формула трапеции при  $n=2$  для  $\int_{-1}^1 f(x) dx$  дает точный результат  $\int_{-1}^1 f(x) dx = 6$ ,

тогда как по формуле Симпсона получаем  $\int_{-1}^1 f(x) dx = -\frac{2}{3}$  [26].

Геометрическая иллюстрация метода Симпсона (рис. 4).

На отрезке  $[x_i; x_{i+2}]$  длиной  $2h$  строится парабола, проходящая через три точки  $(x_i; y_i)$ ,  $(x_{i+1}; y_{i+1})$ ,  $(x_{i+2}; y_{i+2})$ . Площадь под параболой, заключенная между осью  $Ox$  и прямыми  $x = x_i$ ,  $x = x_{i+2}$ , принимают равной интегралу  $\int_{x_i}^{x_{i+2}} f(x) dx$ .

Особенностью применения формулы Симпсона является тот факт, что число разбиений отрезка интегрирования - четное.

Если же количество отрезков разбиения - нечетное, то для первых трех отрезков следует применить формулу, использующую параболу третьей степени, проходящую через четыре первые точки, для аппроксимации подынтегральной функции.

$$I_3 = \frac{3}{8} h (y_0 + 3y_1 + 3y_2 + y_3). \quad (9)$$

Это формула Симпсона «трех восьмых».

Для произвольного отрезка интегрирования  $[a; b]$  формула (9) может быть «продолжена»; при этом число частичных отрезков должно быть кратно трем  $((3m+1)$  точек).

$$I = \int_a^b f(x) dx = \frac{3}{8} h (y_0 + 3y_1 + 3y_2 + 2y_3 + 3y_4 + 3y_5 + 2y_6 + \dots + y_{3m}).$$

$$I = \frac{3}{8} h \left( y_0 + y_{3m} + 2 \sum_{i=1}^k y_{3i} + 3 \sum_{\substack{j=1 \\ j \neq 3l}}^m y_j \right), \quad m = 2, 3, \dots \quad (10)$$

$$k = \left\lfloor \frac{3m}{4} \right\rfloor - \text{целая часть.}$$

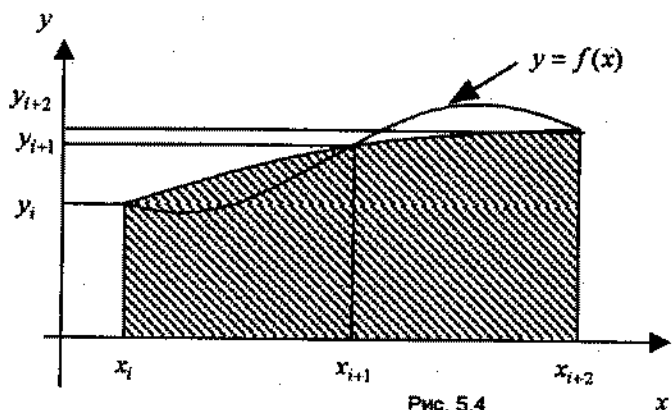


Рис. 5.4

### Структурная схема алгоритма интегрирования

Входные параметры:

A, B – границы интервала интегрирования,

N – начальное число участков разбиения интервала интегрирования,

H – шаг.

Выходные параметры:

S – значение интеграла.

**методом трапеций**

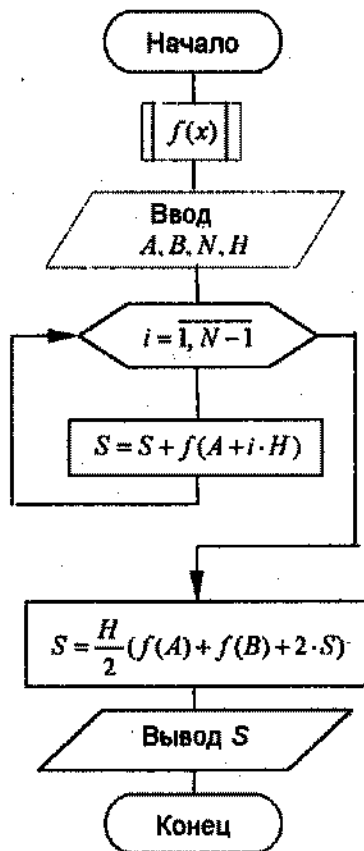


Рис. 5.3

**методом Симпсона**

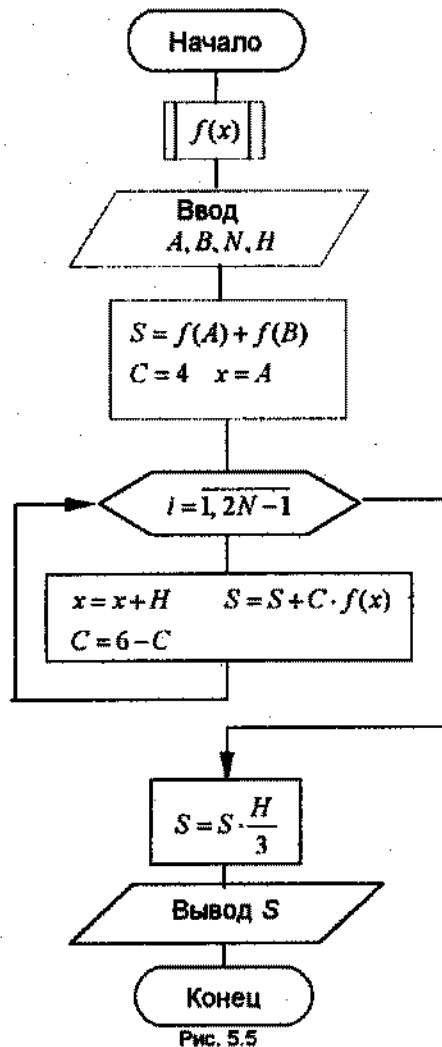


Рис. 5.5

$$I = \int_a^b f(x) dx = \frac{3}{8} h (y_0 + 3y_1 + 3y_2 + 2y_3 + 3y_4 + 3y_5 + 2y_6 + \dots + y_{3m}).$$

$$I = \frac{3}{8} h \left( y_0 + y_{3m} + 2 \sum_{i=1}^k y_{3i} + 3 \sum_{\substack{j=1 \\ j \neq 3i}}^m y_j \right), \quad m = 2, 3, \dots \quad (10)$$

$$k = \left\lfloor \frac{3m}{4} \right\rfloor - \text{целая часть.}$$

Можно получить формулы Ньютона-Котеса старших порядков

$$\int_a^b f(x) dx = c_0 h \sum_{i=0}^n a_i f_i + c_1 h^{k+1} f^{(k)}(x^*), \quad (11)$$

$n$  - количество отрезков разбиения,

$k$  - степень используемого полинома,

$f^{(k)}(x^*)$  - производная  $k$ -го порядка в точке  $x^* \in [a, b]$ ,

$h$  - шаг разбиения.

В таблице 1 выписаны коэффициенты  $C_0, a_i$ .

Каждая строка соответствует одному набору  $k$  промежутков с  $(k+1)$  узлами для построения многочлена  $k$ -ой степени. Чтобы воспользоваться этой схемой для большего количества наборов (например, при  $k=2$  и  $n=6$ ), нужно «продолжить» коэффициенты, а затем сложить их.

Таблица 1

$k$	$c_0$	$a_0$	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	
2	1/3	1	4	1					
				1	4	1			
						1	4	1	
		1	4	2	4	2	4	1	$\Sigma$

$$\int_a^b f(x) dx = \frac{h}{3} (y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + 4y_5 + y_6).$$

Априорные оценки погрешностей формул трапеций и Симпсона можно записать в виде:

$$R_0 = Ah^p, \quad (12)$$

где  $A$  - коэффициент, зависящий от метода интегрирования и свойств подынтегральной функции,

$h$  - шаг интегрирования,

$p$  - порядок метода.



Правило Рунге применяют для вычисления погрешности путем двойного просчета интеграла с шагами  $h$  и  $kh$

$$R_0 = \frac{I_h - I_{kh}}{k^p - 1}, \quad (13)$$

(13) - апостериорная оценка.

Тогда

$$I_{\text{уточн}} = I_h + R_0 \quad (14)$$

$I_{\text{уточн}}$  - уточненное значение интеграла  $I = \int_a^b f(x) dx$ .

Если порядок метода неизвестен, необходимо вычислить  $I$  в третий раз с шагом  $k^2h$ , т.е.

$$I = I_{k^2h} + A(k^2h)^p,$$

Из системы трех уравнений

$$\begin{cases} I = I_h + Ah^p, \\ I = I_{kh} + A(kh)^p, \\ I = I_{k^2h} + A(k^2h)^p \end{cases}$$

с неизвестными  $I$ ,  $A$  и  $p$  получаем

$$k^p = \frac{I_{kh} - I_{k^2h}}{I_h - I_{kh}} \quad (15)$$

$$\text{Из (15) следует } p = \frac{\ln\left(\frac{I_{kh} - I_{k^2h}}{I_h - I_{kh}}\right)}{\ln k} \quad (16)$$

Таким образом, метод двойного просчета, использованный необходимое число раз, позволяет вычислить интеграл с заданной степенью точности. Выбор необходимого числа разбиений осуществляется автоматически. Можно при этом использовать многократное обращение к подпрограммам соответствующих методов интегрирования, не изменяя алгоритмов этих методов. Однако для методов, использующих равноотстоящие узлы, удастся модифицировать алгоритмы и уменьшить вдвое количество вычислений подынтегральной функции за счет использования интегральных сумм, накопленных при предыдущих кратных разбиениях интервала интегрирования [49].

Два приближенных значения интеграла  $I = \int_a^b f(x) dx$   $I_k$  и  $I_{k-1}$ , вычисляемые по методу трапеций с шагами  $h_k$  и  $h_{k-1}$ , связаны соотношением:

$$I_k = \frac{I_{k-1}}{2} + h_k \sum f(a + (2i-1)h_k), \quad i = \overline{1, 2^{k-1}} \quad (17)$$

$$h_k = \frac{h_{k-1}}{2} = 2^{-k} h_0, \quad h_0 = b - a, \quad I_0 = \frac{f(a) + f(b)}{2} h_0, \quad k = 1, 2, \dots$$

Схема алгоритма приведена на рис. 7.

Аналогично, для интегралов, вычисленных по формуле Симпсона с шагами  $h_k$  и  $h_{k-1}$ , справедливы соотношения:

$$I_k = (S_k + A) \frac{h_k}{3}, \quad k = 0, 1, \dots, \quad A = f(a) + f(b), \quad (18)$$

$$S_k = S_{k1} + S_{k2},$$

$$S_{k1} = \frac{S_{k-1,1} + S_{k-1,2}}{2}, \quad S_{01} = f(a) + f(b),$$

$$S_{k2} = 2 \sum_i f(a - h_k + ih_{k-1}), \quad i = \overline{1, 2^k}, \quad S_{02} = 2f(a + h_0),$$

$$h_k = \frac{h_{k-1}}{2}, \quad h_0 = \frac{b-a}{2}, \quad k = 1, 2, \dots$$

Схема алгоритма приведена на рис. 8.

Дальнейшее развитие приведенных алгоритмов позволяет перейти к адаптивным алгоритмам [70], в которых за счет выбора различного шага интегрирования в разных частях отрезка интегрирования в зависимости от свойств  $f(x)$  уменьшается количество вычислений подынтегральной функции.

Другая схема уточнения значений интеграла - процесс Эйткена. Производится вычисление интеграла с шагами  $h_1$ ,  $h_2$ ,  $h_3$ , причем

$$\frac{h_2}{h_1} = \frac{h_3}{h_2}. \quad \text{Вычисленные значения } I(h_1) = I_1, I(h_2) = I_2, I(h_3) = I_3.$$

Тогда

$$\int_a^b f(x) dx = I_1 - \frac{(I_1 - I_2)^2}{(I_1 - 2I_2 + I_3)}. \quad (19)$$

Наряду с оценками (5) и (8) используются практические оценки погрешности [49] элементарных формул численного интегрирования. Для формулы трапеций

$$R(f) = \frac{b-a}{3} \left( f(b) - 2f\left(\frac{a+b}{2}\right) + f(a) \right). \quad (20)$$

За меру погрешности метода Симпсона принимаем величину (сильно завышенная оценка)

$$R(f) = \frac{4(b-a)}{81} \left( f(b) - 2f\left(\frac{a+b}{2}\right) + f(a) \right). \quad (21)$$

**Структурная схема алгоритма интегрирования методом Симпсона с использованием правила Рунге**

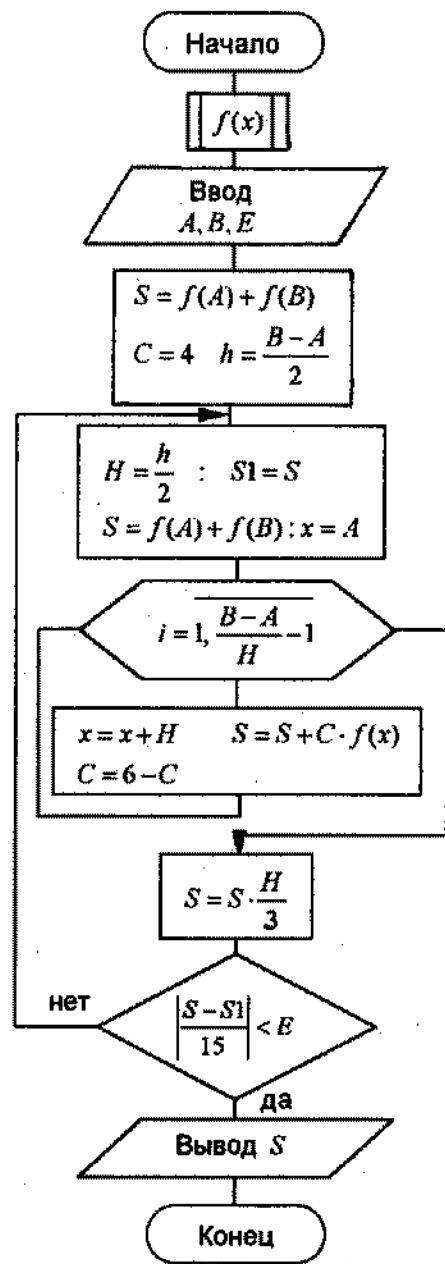


Рис. 5.6.

Входные параметры:  
A, B – границы отрезка интегрирования,  
E – погрешность.  
Выходные параметры:  
S – значение интеграла.

**Пример 1.**

Вычислить интеграл  $\int_{0.7}^{1.3} \frac{dx}{\sqrt{2x^2 + 0.3}}$  по формуле трапеций, определить погрешность, если  $h = 0.15$ .

Составим таблицу.

Таблица 2

$i$	$x_i$	$y_0, y_4$	$y_1, y_2, y_3$
0	0.7	0.88386	
1	0.85		0.75700
2	1		0.65937
3	1.15		0.58272
4	1.3	0.52129	

$$I_T \approx h \left( \frac{y_0 + y_4}{2} + y_1 + y_2 + y_3 \right) = 0.4044$$

Оценим погрешность по формуле (5)  $|R| = \frac{(b-a)^3}{12n^2} M_2$ .

$$M_2 = \max |f''(x)| = \max \left| \frac{8x^2 - 0.6}{\sqrt{(2x^2 + 0.3)^3}} \right| \leq \frac{8 \cdot 1.3^2 - 0.6}{\sqrt{(2 \cdot 0.7^2 + 0.3)^3}} \approx 6.98,$$

$$|R| = \frac{(1.3 - 0.7)^2}{12 \cdot 4^2} \cdot 6.98 = 0.008 \text{ для } x \in [0.7; 1.3].$$

**Пример 2.**

Вычислить  $\int_0^{0.8} f(x) dx$  по формуле Симпсона, если  $f(x)$  задана таблицей. Оценить погрешность.

Таблица 3

$x$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
$f(x)$	1	0.995	0.98	0.955	0.921	0.878	0.825	0.765	0.697

**Решение.**

Вычислим по формуле (6) при  $h = 0.1$  и  $h = 0.2$  интеграл

$$I_S = \int_0^{0.8} f(x) dx = \frac{1}{30} (1 + 0.697 + 4(0.995 + 0.955 + 0.878 + 0.765) + 2(0.980 + 0.921 + 0.825)) = 0.71735,$$

$$I_4 = \int_0^{0.8} f(x) dx = \frac{2}{30} (1 + 0.697 + 2 \cdot 0.921 + 4(0.980 + 0.825)) = 0.71833.$$

По правилу Рунге получаем  $R \approx 6.53 \cdot 10^{-5}$ .

Принимаем  $I = 0.71735 + 6.53 \cdot 10^{-5} \approx 0.7174$ .

Интегралы, вычисленные с шагом  $h = 0.2$  по формуле средних прямоугольников и формуле трапеций, соответственно равны:

$$I_{4np} = 0.7186, \quad I_{4mp} = 0.7149.$$

## 5.2. Сплайн-квадратура

Пусть требуется вычислить  $I = \int_a^b f(x) dx$ .

Разобьем  $[a; b]$  на отрезки длиной

$$h_j = x_j - x_{j-1}, \quad j = \overline{1, n}, \quad x_0 = a, \quad x_n = b.$$

На каждом из отрезков заменим подынтегральную функцию кубическим сплайном  $\varphi_j(x)$

$$\varphi_j(x) = a_j + b_j(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3, \quad x \in [x_{j-1}; x_j]. \quad (22)$$

$$I = \int_a^b f(x) dx = \sum_{j=1}^n \int_{x_{j-1}}^{x_j} \varphi_j(x) dx.$$

Коэффициенты сплайнов  $a_j, b_j, d_j$  выражаются через коэффициенты  $c_j$  и значения подынтегральной функции  $f_i(x)$ :

$$\begin{aligned} a_j &= f_{j-1}, \\ b_j &= \frac{f_j - f_{j-1}}{h_j} - \frac{(c_{j+1} + 2c_j)h_j}{3}, \\ d_j &= \frac{c_{j+1} - c_j}{3h_j}, \end{aligned} \quad (23)$$

из условий сплайн-интерполяции и граничных условий [49].

Полагаем, что граничные условия для свободных концов сплайнов  $c_1 = 0$  и  $c_{n+1} = 0$  ( $n$  - количество сплайнов).

Полученную систему линейных алгебраических уравнений относительно коэффициентов  $c_j$  удобнее решать методом прогонки.

Тогда интеграл

$$I \approx \sum_{j=1}^n h_j \frac{(f_j + f_{j-1})}{2} - \sum_{j=1}^n h_j^3 \frac{(c_j + c_{j+1})}{12}. \quad (24)$$

## Структурная схема алгоритма численного интегрирования методом трапеций (с использованием рекуррентных соотношений)

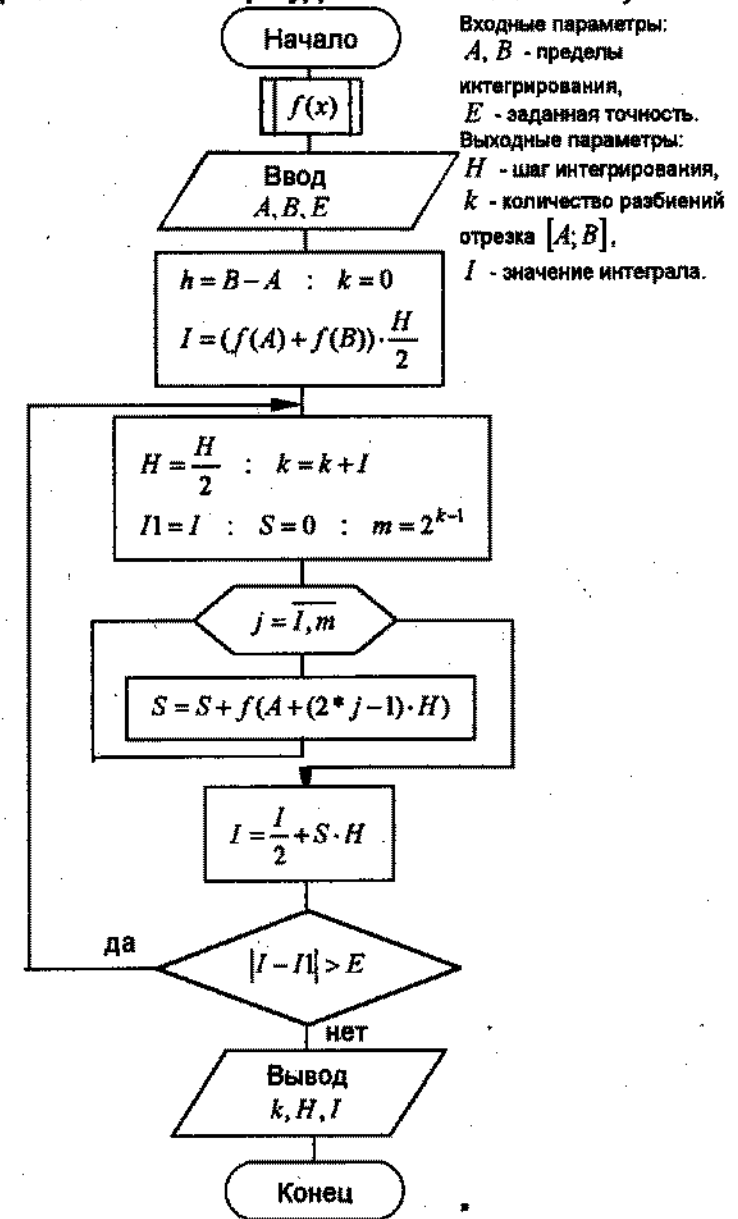


Рис. 5.7

**Схема алгоритма интегрирования по методу Симпсона  
(с использованием рекуррентных соотношений)**

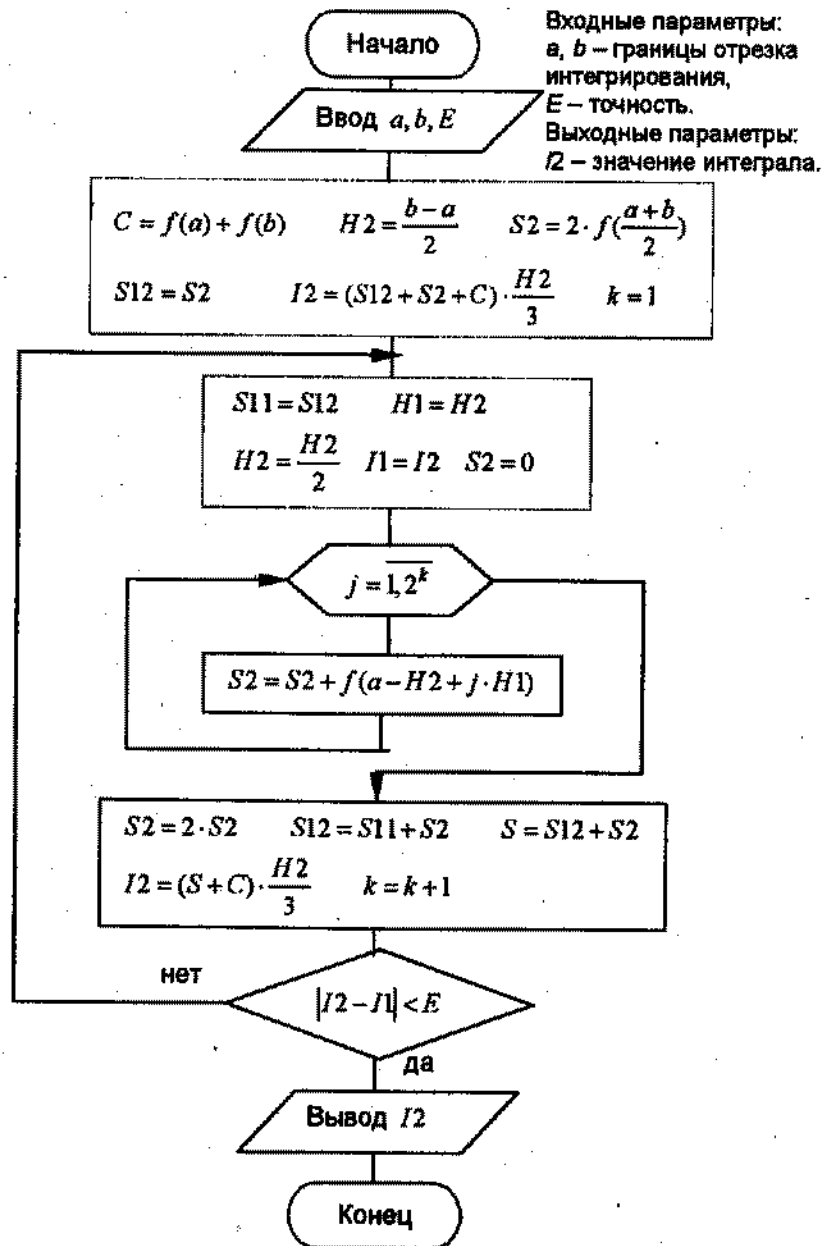


Рис. 5.8.

Алгоритм сплайн-квадратуры сложнее алгоритмов методов трапеций и Симпсона за счет необходимости решения системы линейных уравнений для определения коэффициентов сплайнов  $c_j$  (рис. 9-12).

Метод лучше применять для комплексной обработки данных, сглаживания зависимостей и т.п. Доказано, что погрешность сплайн-квадратуры меньше погрешности метода трапеций.

**5.3. Квадратуры Гаусса**

Квадратуры Гаусса используют, если интегрируемая функция задана аналитически. Подынтегральную функцию аппроксимируют полиномами различных степеней, узлы интерполяции выбирают из условия обеспечения минимальной погрешности интегрирования.

Эта задача была решена Гауссом для интегралов вида

$I = \int_a^b f(x) dx$ , для интегралов  $I_1 = \int_a^b p(x) f(x) dx$  с произвольной весовой функцией  $p(x)$  - Кристоффелем.

Квадратурная формула Гаусса-Кристоффеля при  $n$  узлах содержит  $2n$  параметров и имеет вид

$$I_1 = \sum_{k=1}^n c_k f(x_k).$$

Заменой  $x = \frac{b-a}{2}t + \frac{a+b}{2}$  интеграл преобразуется к виду

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

Аппроксимация полиномом степени  $n$  дает выражение

$$\int_a^b f(x) dx = \frac{b-a}{2} \sum_{i=1}^n A_i f(x_i), \quad x_i = \frac{b-a}{2}t_i + \frac{a+b}{2}. \quad (25)$$

Для обеспечения равенства (25) необходимо и достаточно, чтобы оно было верным при  $f(t) = 1, t, t^2, \dots, t^{2n-1}$ .

Доказано, что для достижения наивысшей  $(2n-1)$  точности квадратурной формулы (25) в качестве  $t_i$  достаточно взять нули соответствующего полинома Лежандра.

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n (x^2 - 1)^n}{dx^n}, \quad n = 0, 1, \dots \quad (26)$$

### Алгоритм вычисления интеграла методом сплайнов

Входные параметры:  
 a, b – границы отрезка интегрирования,  
 N – количество сплайнов.  
 Выходные параметры:  
 S – значение сплайн-квадратуры.

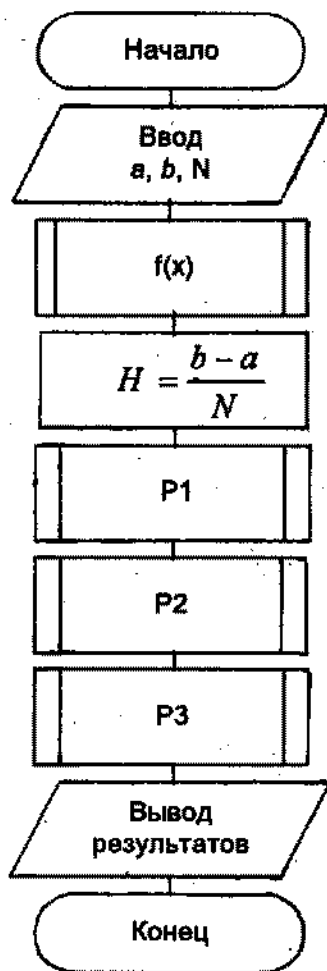


Рис. 5.9

### Процедура вычисления сплайн-квадратуры

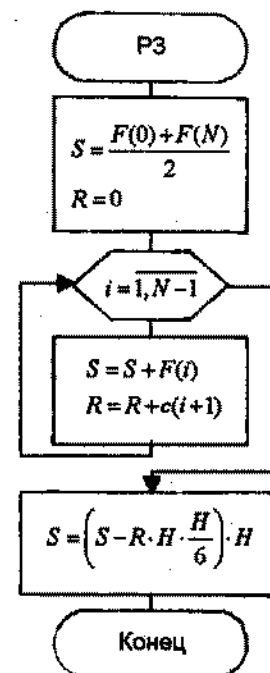


Рис.5.10.

### Процедура формирования таблицы

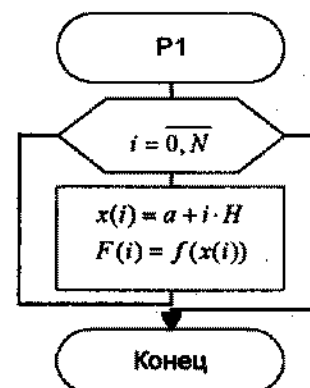


Рис. 5.11

Процедура формирования таблицы коэффициентов сплайнов и прогоночных коэффициентов

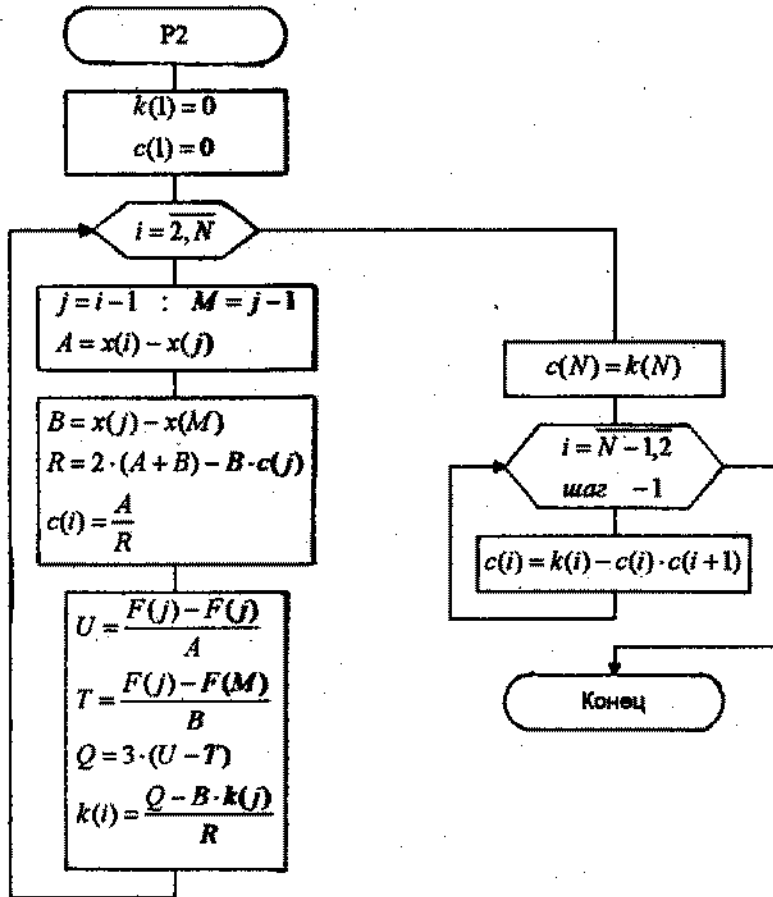


Рис. 5.12

Важнейшие свойства полиномов Лежандра:

•  $P_n(1) = 1, P_n(-1) = (-1)^n, n = 0, 1, \dots$

•  $\int_{-1}^1 P_n(x) \cdot Q_k(x) dx = 0, (k < n),$

$Q_k(x)$  - любой полином степени  $k$ , меньшей  $n$ .

• Полином Лежандра  $P_n(x)$  имеет  $n$  различных и действительных корней, которые располагаются на интервале  $(-1; 1)$ .

Первые пять полиномов Лежандра:

$$P_0(x) = 1,$$

$$P_1(x) = x,$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1),$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x),$$

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3).$$

Приближенные значения абсцисс  $t_i$  и коэффициентов  $A_i$

Таблица 4

$n$	$i$	$t_i$	$A_i$
1	1	0	2
2	1; 2	$\mp 0.5735027$	1
3	1; 3 2	$\mp 0.77459667$ 0	0.55555556 0.88888889
4	1; 4 2; 3	$\mp 0.86113631$ $\mp 0.33998104$	0.34785484 0.65214516
5	1; 5 2; 4 3	$\mp 0.90617985$ $\mp 0.53849631$ 0	0.23692688 0.47862867 0.56888889

Остаточный член формулы Гаусса выражается следующим образом:

$$R_n = \frac{2^{2n+1} (n!)^4 f^{(2n)}(\xi)}{((2n)!)^3 (2n+1)}, \quad -1 \leq \xi \leq 1.$$

**Замечание.**

Для функций, имеющих достаточное количество производных, формула Гаусса значительно точнее, чем формулы трапеций и Симпсона даже при малом количестве ординат. Для функций с разрывом производных это утверждение не всегда справедливо.

**Пример 3.**

По формуле Гаусса при  $n = 3$  вычислить интеграл

$$I = \int_0^1 \sqrt{1+x^2} dx.$$

**Решение.**

$a = 0, b = 1$ . Из таблицы (4) при  $n = 3$

$$t_1 = -0.77459667, \quad t_2 = 0, \quad t_3 = 0.77459667,$$

$$A_1 = A_3 = 0.55555556, \quad A_2 = 0.88888889.$$

$$x_1 = \frac{1}{2} + \frac{1}{2}t_1 = 0.11270, \quad f(x_1) = 1.100633,$$

$$x_2 = 0.50000, \quad f(x_2) = 1.11803,$$

$$x_3 = 0.88730, \quad f(x_3) = 1.33690.$$

$$\int_0^1 \sqrt{1+x^2} dx = \frac{1}{2} \sum_{i=1}^3 A_i f(x_i) = 1.14780.$$

Наиболее употребительные весовые функции  $p(x)$  интеграла Гаусса-Кристоффеля  $I_1 = \int_a^b p(x)f(x)dx$  связаны с так называемыми классическими ортогональными многочленами: многочленами Якоби (частные случаи – многочлены Лежандра и Чебышева), Лягерра и Эрмита.

1) Весовая функция многочлена Якоби

$$p(x) = (1-x)^\alpha (1+x)^\beta; \quad \alpha, \beta > -1. \quad (27)$$

При  $\alpha = \beta = -\frac{1}{2}$   $p(x) = \frac{1}{\sqrt{1-x^2}}$  используют многочлены Че-

бышева первого рода  $T_n(x)$ .

Первые пять многочленов имеют вид:

$$T_0(x) = 1,$$

$$T_1(x) = x,$$

$$T_2(x) = 2x^2 - 1,$$

$$T_3(x) = 4x^3 - 3x,$$

$$T_4(x) = 8x^4 - 8x^2 + 1.$$

На отрезке  $[-1; 1]$  узлы многочлена и коэффициенты:

$$t_i^{(n)} = \cos\left(\pi\left(i - \frac{1}{2}\right)/n\right), \quad A_i^{(n)} = \frac{\pi}{n}, \quad 1 \leq i \leq n.$$

$$\int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx = \sum_{i=1}^n A_i f(t_i) = \frac{\pi}{n} \sum_{i=1}^n f(t_i).$$

Формула для оценки погрешности:

$$R_n = \frac{\pi}{2^{2n-1}} \frac{f^{(2n)}(\xi)}{(2n)!}, \quad -1 \leq \xi \leq 1.$$

При  $\alpha = \beta = \frac{1}{2}$   $p(x) = \sqrt{1-x^2}$  применяют ортогональные

многочлены Чебышева второго рода  $U_n(x)$ :

$$U_0(x) = 1,$$

$$U_1(x) = 2x,$$

$$U_2(x) = 4x^2 - 1,$$

$$U_3(x) = 8x^3 - 4x,$$

$$U_4(x) = 16x^4 - 12x^2 + 1.$$

Узлы многочленов  $t_m^{(n)} = \cos \frac{m\pi}{n+1}$ ,  $1 \leq m \leq n$  и коэффициенты

в таблице 5.

Таблица 5

$n$	$i$	$A_i$
1	1	$\frac{\pi}{2}$
2	1; 2	$\frac{\pi}{4}$
3	1; 3	$\frac{\pi}{8}$
	2	$\frac{\pi}{4}$

$$\int_{-1}^1 \sqrt{1-x^2} f(x) dx = \frac{\pi}{n+1} \sum_{m=1}^n \sin^2 \frac{m\pi}{n+1} f(t_m).$$

$$R_n = \frac{\pi}{2^{2n}} \frac{f^{(2n)}(\xi)}{(2n)!}, \quad -1 \leq \xi \leq 1.$$

2) При интегрировании на полуоси  $[0; \infty)$  с весом  $p(x) = x^\alpha e^{-x}$  используют многочлены Лягерра  $L_n^{(\alpha)}(x)$ .

При  $\alpha = 0$

$$L_0^{(0)}(x) = 1,$$

$$L_1^{(0)}(x) = -x + 1,$$

$$L_2^{(0)}(x) = x^2 - 4x + 2,$$

$$L_3^{(0)}(x) = -x^3 + 9x^2 - 18x + 6,$$

$$L_4^{(0)}(x) = x^4 - 16x^3 + 72x^2 - 96x + 24.$$

Таблица 6

$n$	$i$	$x_i$	$A_i$
1	1	1.6768750	1
2	1	1.6379870	0.5042845
	2	0.2234308	0.4957155
3	1	2.5271443	0.2387737
	2	0.7114176	0.4964939
	3	0.1097820	0.2647324

$$\int_0^{\infty} e^{-x} f(x) dx = \sum_{i=1}^n A_i f(x_i).$$

3) Многочлены Эрмита  $H_n(x)$  используются при интегрировании на интервале  $(-\infty; +\infty)$ ,  $p(x) = e^{-x^2}$ .

$$H_0(x) = 1,$$

$$H_1(x) = 2x,$$

$$H_2(x) = 4x^2 - 2,$$

$$H_3(x) = 8x^3 - 12x,$$

$$H_4(x) = 16x^4 - 48x^2 + 12.$$

(30)

Таблица 7

$n$	$i$	$x_i$	$A_i$
1	1	0	$\sqrt{\pi}$
2	1;2	$\mp \frac{1}{\sqrt{2}}$	$\frac{\sqrt{\pi}}{2}$
	1;3	$\mp \frac{\sqrt{3}}{\sqrt{2}}$	
2		0	

$$\int_{-\infty}^{+\infty} e^{-x^2} f(x) dx = \sum_{i=1}^n A_i f(x_i).$$

Если подынтегральная функция интеграла  $I(f) = \int_0^{\infty} f(x) dx$  хорошо приближается тригонометрическими многочленами с периодом  $\omega$ , то целесообразно применять квадратуру, являющуюся аналогом квадратуры Гаусса [7].

### Структурная схема алгоритма вычисления квадратур Гаусса



Рис. 5.13

$$I(f) = S_n(f) = \frac{\omega}{N} \sum_{j=0}^{N-1} f\left(\frac{j\omega}{N}\right). \quad (31)$$

Эта квадратура точна для функции  $\cos\left(\frac{2\pi m x}{\omega}\right)$  при  $m=0$  или не целом значении  $\frac{m}{N}$  и для всех  $\sin\left(\frac{2\pi m x}{\omega}\right)$ , то есть квадратура точна для любого тригонометрического многочлена.

$$t_N(x) = a_0 + \sum_{m=1}^{N-1} \left( a_m \cos\left(2\pi m \frac{x}{\omega}\right) + b_m \sin\left(2\pi m \frac{x}{\omega}\right) \right) + b_N \sin\left(2\pi N \frac{x}{\omega}\right). \quad (32)$$

Оценка такова

$$|R_N(f)| \leq 2\omega \cdot \inf_{t_N \in [0, \omega]} \max |f(x) - t_N(x)|.$$



#### 5.4. Интегрирование быстро осциллирующих функций

Задача интегрирования быстро осциллирующих функций - одна из самых трудных задач численных методов - имеет большое прикладное значение в радиотехнике связи. Примерами могут служить задачи спектрального анализа, построение диаграмм направленности антенн и другие.

$$\text{Пусть требуется вычислить } \int_a^b f(x) \exp(i\omega x) dx, \quad (33)$$

где  $f(x)$  - гладкая функция,  $\omega(b-a) \gg 1$ .

Поскольку  $\operatorname{Re}(f(x)\exp(i\omega x))$  и  $\operatorname{Im}(f(x)\exp(i\omega x))$  имеют на  $[a, b]$  примерно  $\frac{\omega(b-a)}{\pi}$  нулей, то интеграл предпочтительно вычислять методом Гаусса-Кристоффеля с весовой функцией  $\exp(i\omega x)$ .

Взяв  $n$  узлов интерполяции

$$x_j = \frac{b+a}{2} + \frac{b-a}{2} t_j, \quad j = \overline{1, n}, \quad (34)$$

заменяем  $f(x)$  многочленом, в результате интегрирования получаем

$$\int_a^b f(x) \exp(i\omega x) dx = \frac{b-a}{2} \exp\left(i\omega \frac{a+b}{2}\right) \sum_{j=1}^n D_j \left(\omega \frac{b-a}{2}\right) f(x_j), \quad (35)$$

$$D_j(p) = \int_{-1}^1 \left( \prod_{k \neq j} \frac{t-d_k}{d_j-d_k} \right) \exp(ip t) dt. \quad (36)$$

Оценка остаточного члена

$$|R_n(f)| \leq D(d_1, d_2, \dots, d_n) \cdot \max_{[a,b]} |f^{(n)}(x)| \cdot \left(\frac{b-a}{2}\right)^{n+1}. \quad (37)$$

Как видно из (37) оценка погрешности требует вычисления производных высоких порядков. Поэтому часто пользуются эмпирическим методом. При малом шаге интегрирования квадратурные формулы назовем широкополосными, в при большом шаге - узкополосными. Так как одной из целей гармонического анализа сигнала является определение эффективной ширины спектра сигнала, то вначале выбирается заведомо широкополосная формула. Иногда применяют такой прием. При некотором начальном шаге  $\Delta x_0$  вычисляют спектр и проверяют относительную величину и скорость убывания спектральных составляющих по мере приближения к границе «полосы пропускания» формулы. Если составляющие малы и

стремятся к убыванию, то принимают, что качество интегрирования достаточное, в противном случае вычисляют спектр при  $\Delta x_1 < \Delta x_0$ .

Для некоторых частных случаев функций, быстро осциллирующих в промежутке интегрирования, существуют методы, позволяющие учесть осцилляцию функции в самой квадратурной формуле. Один из методов - метод Филона, применяемый для интегрирования функций вида  $f(x)\cos(\omega x)$ ,  $f(x)\sin(\omega x)$ .

Метод Филона предусматривает аппроксимацию внутри каждого сдвоенного интервала параболой. Шаг интегрирования нужно выбирать из требования обеспечения заданной точности интегрирования не всего подынтегрального выражения, а только медленно меняющейся и не зависящей от  $\omega$  функции  $f(x)$ .

Формулы Филона:

$$\begin{aligned} \int_a^b f(x) \sin \omega x dx &= \\ &= \frac{b-a}{2n} \sum_{k=0}^{n-1} \left[ -\alpha (f(x_{k+1}) \cos \omega x_{k+1} - f(x_k) \cos \omega x_k) + \right. \\ &\left. + \frac{\beta}{2} (f(x_{k+1}) \sin \omega x_{k+1} + f(x_k) \sin \omega x_k) + \gamma f\left(\frac{x_k + x_{k+1}}{2}\right) \sin \frac{\omega}{2} (x_k + x_{k+1}) \right], \end{aligned} \quad (38)$$

$$\begin{aligned} \int_a^b f(x) \cos \omega x dx &= \\ &= \frac{b-a}{2n} \sum_{k=0}^{n-1} \left[ \alpha (f(x_{k+1}) \sin \omega x_{k+1} - f(x_k) \sin \omega x_k) + \right. \\ &\left. + \frac{\beta}{2} (f(x_{k+1}) \cos \omega x_{k+1} + f(x_k) \cos \omega x_k) + \gamma f\left(\frac{x_k + x_{k+1}}{2}\right) \cos \frac{\omega}{2} (x_k + x_{k+1}) \right], \end{aligned} \quad (39)$$

$$\begin{aligned} \alpha &= \frac{1}{\theta^3} (\theta^2 + \theta \sin \theta \cos \theta - 2 \sin^2 \theta), \\ \beta &= \frac{2}{\theta^3} (\theta (1 + \cos^2 \theta) - 2 \sin \theta \cos \theta), \\ \gamma &= \frac{4}{\theta^3} (\sin \theta - \theta \cos \theta), \\ \theta &= \frac{\omega}{2n} (b-a). \end{aligned} \quad (40)$$

В случае, если для аппроксимации используется кусочно-постоянная или кусочно-линейная функция, получают формулы вида

$$\int_a^b f(x) \cos \omega x dx = I(\omega) = \Delta x \sum_{k=1}^N f(k\Delta x) \cos(\omega k\Delta x) \cdot \frac{\sin \frac{\omega \Delta x}{2}}{\frac{\omega \Delta x}{2}}, \quad (41)$$

$$\int_a^b f(x) \cos \omega x dx = \Delta x \sum_{k=1}^N f(k\Delta x) \cos(\omega k\Delta x) \left( \frac{\sin \frac{\omega \Delta x}{2}}{\frac{\omega \Delta x}{2}} \right)^2, \quad (42)$$

$$\Delta x = \frac{b-a}{N} \text{ - шаг интегрирования.}$$

Для шага интегрирования используются [44] оценки

$$\Delta x \approx \frac{2\sqrt{6E}}{\omega_{\max}}, \quad (43)$$

или

$$\Delta x \approx \frac{2\sqrt{6E(E+1)}}{\omega_{\max}}, \quad (44)$$

где  $E$  - относительная погрешность аппроксимации в диапазоне частот  $[0, \omega_{\max}]$ .

Стандартные программы вычисления интегралов от быстро осциллирующих функций используют формулы (38), (39) при  $n=3$  (формулы Филона) или  $n=5$  [21], [44].

$$n=3; d_1=-1; d_2=0; d_3=1,$$

$$n=5; d_1=-1; d_2=-0.5; d_3=0; d_4=0.5; d_5=1.$$

Эти программы должны иметь блок, предусматривающий возможность изменения расчетных формул в зависимости от величины  $p$ , чтобы избежать при малых  $p$  существенного влияния вычислительной погрешности.

Если  $|p| > p_n$  то вычисления следует проводить по формулам (36), (35), если  $|p| \leq p_n$ , то исходный интеграл вычисляют по формуле трапеции с подынтегральной функцией  $f(x) \exp(iax)$ . Величину  $p_n$  подбирают экспериментально.

#### Пример 5.

Вычислить интеграл  $I = \int_0^{\pi} e^x \sin 10x dx$  с точностью  $E = 10^{-6}$  по формулам Филона.

Решение.

Воспользуемся формулой (8) для теоретического определения величины шага  $h$ , обеспечивающее выполнение условия

$$|R| \leq E \text{ для интеграла } \int_0^{\pi} e^x dx. \text{ Получим } |R| = \frac{\pi^5}{180(2n)^4} e^{\pi}, \quad n > 12,$$

$$h < 0.23.$$

Результаты вычислений по формулам (38) сведены в таблицу 8

Таблица 8

$n$	$h$	$I$
2	1.57	-2.195669
4	0.785	-2.197889
6	0.5233334	-2.194502
8	0.3925	-2.191976
10	0.314	-2.191867
12	0.2616667	-2.191855
14	0.2242857	-2.191853

### 5.5. Интегрирование периодических функций

Требуется вычислить интеграл

$$\int_0^{2\pi} p(x) f(x) dx, \quad (45)$$

где  $p(x)$  - весовая функция, произвольная интегрируемая функция, не равная тождественно нулю. Функция  $f(x)$  - любая периодическая функция, имеющая непрерывные производные до некоторого порядка.

Пусть  $f(0) = f(2\pi)$ . Интеграл вычисляется по отрезку, равному периоду функции  $f(x)$ . При интерполяции функции  $f(x)$  используются формулы интерполирования тригонометрическими многочленами, исходя из системы функций  $1, \cos kx, \sin kx, k=1,2,3...$

$$f(x) \approx T(x) = \sum_{k=1}^n l_k f(x_k), \quad (46)$$

$$\int_0^{2\pi} p(x) f(x) dx \approx \sum_{k=1}^n A_k f(x_k), \quad (47)$$

$$A_k = \int_0^{2\pi} p(x) \chi_k(x) dx.$$

Структурная схема алгоритма интегрирования  
методом Филона

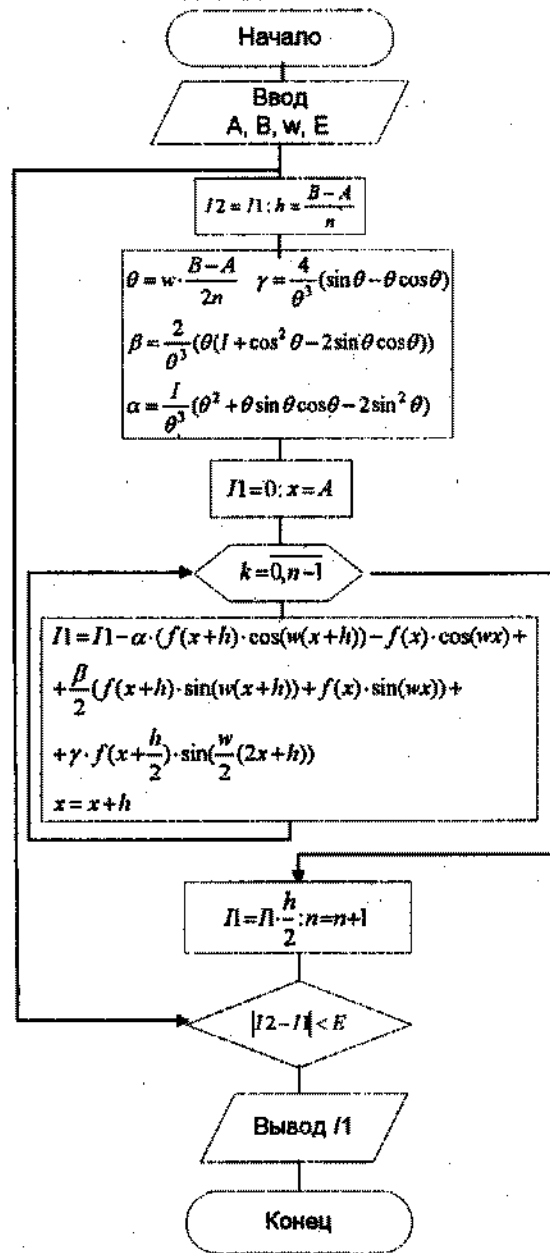


Рис. 5.14

При  $n = 2m + 1$

$$T(x) = a_0 + \sum_{i=1}^m (a_i \cos ix + b_i \sin ix), \quad (48)$$

$$l_k(x) = \frac{\omega(x)}{\omega'(x) 2 \sin \frac{1}{2}(x-x_k)}, \quad \omega(x) = \prod_{j=1}^n \sin \frac{1}{2}(x-x_j), \quad (49)$$

В этом случае правило (47) будет точным для любых тригонометрических многочленов степени  $m$ .

При  $n = 2m$

$$T(x) = T_{m-1}(x) + a_m \cos mx, \quad (50)$$

$$l_k(x) = \frac{\omega(x) \cos \frac{1}{2}(x-x_k+2\delta)}{\omega'(x) \cdot 2 \sin \frac{1}{2}(x-x_k) \cos \delta}, \quad (51)$$

$$2\delta = \sum_{j=1}^n x_j, \quad (52)$$

$$\delta = \left(i - \frac{1}{2}\right)\pi, \quad i = 1, 2, 3, \dots \quad (53)$$

При указанных ограничениях правило интегрирования (47) будет точным, если  $f(x)$  есть многочлен вида (50).

Для случая  $p(x) = 1$

$$\int_0^{2\pi} f(x) dx \approx \sum_{k=1}^n A_k f(x_k). \quad (54)$$

Наивысшую тригонометрическую степень точности  $(n-1)$  имеет правило прямоугольников с равноотстоящими абсциссами и равными коэффициентами. Разделим отрезок  $[0, 2\pi]$  на  $n$  равных частей и выберем произвольную точку  $\alpha$ ,  $0 \leq \alpha \leq \frac{2\pi}{n}$ .

При любом  $\alpha$

$$\int_0^{2\pi} f(x) dx \approx \frac{2\pi}{n} \sum f\left(\alpha + \frac{2\pi k}{n}\right). \quad (55)$$

Равенство (55) будет верным для всех тригонометрических многочленов степени  $n-1$ . Если положить  $\alpha = \frac{\pi}{2n}$  или  $\alpha = \frac{3\pi}{2n}$ , то (55) будет точным для многочленов вида (50).

Вопрос о сходимости квадратурных формул рассмотрен в [44].

## 5.6. Несобственные интегралы

• Для сходящихся несобственных интегралов с бесконечными пределами есть несколько приемов вычисления.

1) Метод замены переменных, позволяющий перейти к конечным пределам.

$$\int_a^{\infty} f(x) dx \text{ заменой } x = \frac{a}{1-t} \text{ приводится к интегралу } \int_0^1 \varphi(t) dt.$$

2) Замена исходного интеграла суммой

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

Интеграл  $\int_b^{\infty} f(x) dx$  оценивают приближенно численным методом или аналитически.

Если  $\left| \int_b^{\infty} f(x) dx \right| < E$ , где  $E$  - точность, то последний интеграл

вычисляют обычными методами интегрирования.

3) Использование для вычисления квадратурных формул Гаусса-Кристоффеля.

Подынтегральную функцию  $f(x)$  записывают в виде произведения  $p(x)F(x)$ , где  $p(x)$  - весовая функция.

В таблицах [44] приведены виды весовых функций, веса квадратурных формул, количество узлов и заранее вычисленные узлы.

### Пример 6.

$$\text{Вычислить } \int_{-\infty}^{\infty} \cos x \cdot e^{-x^2} dx.$$

Решение.

Для  $S = \int_{-\infty}^{\infty} \cos x \cdot e^{-x^2} dx$  весовой функцией может

быть  $p(x) = e^{-x^2}$ , узлами квадратурной формулы будут корни многочлена Эрмита  $H_n(x)$ .

Так, при  $n = 5$  узлы и весовые коэффициенты таковы:

$$x_1 = 0.958572, \quad x_2 = 2.020183, \quad x_3 = 0, \quad x_4 = -x_1, \quad x_5 = -x_2,$$

$$c_1 = 0.393619, \quad c_2 = 0.019953, \quad c_3 = 0.945300, \quad c_4 = c_1, \quad c_5 = c_2.$$

Вычисление  $S$  проводится по формуле

$$S = \sum_{k=1}^5 c_k \cos(x_k), \quad S = 1.3804.$$

• Несобственные интегралы с конечными пределами от разрывных функций вычисляются методами аддитивного или мультипликативного выделения особенностей, а также построением нестандартных квадратурных формул.

При аддитивном способе выделения особенностей из подынтегральной функции  $f(x)$  выделяют некоторую функцию  $g(x)$ , имеющую ту же особенность, что и  $f(x)$ , элементарно интегрируемую и такую, чтобы  $f(x) - g(x)$  была достаточно гладкой. Тогда [8] будем иметь:

$$\int_a^b f(x) dx = \int_a^b g(x) dx + \int_a^b (f(x) - g(x)) dx. \quad (56)$$

Выбор  $g(x)$  производится различным образом.

Для одного часто встречающегося класса интегралов от функции

$$f(x) = (x-c)^{\alpha} \varphi(x), \quad a \leq c \leq b, \quad (57)$$

где  $-1 < \alpha < 0$ , а  $\varphi(x)$  непрерывна и имеет достаточное количество производных,  $f(x)$  представляют в виде

$$f(x) = \left( \varphi(c)(x-c)^{\alpha} + \frac{\varphi'(c)}{1!}(x-c)^{\alpha+1} + \dots + \frac{\varphi^{(k)}(c)}{k!}(x-c)^{\alpha+k} \right) + (x-c)^{\alpha} \left( \varphi(x) - \varphi(c) - \frac{\varphi'(c)}{1!}(x-c) - \dots - \frac{\varphi^{(k)}(c)}{k!}(x-c)^k \right), \quad (58)$$

$$g(x) = \varphi(c)(x-c)^{\alpha} + \frac{\varphi'(c)}{1!}(x-c)^{\alpha+1} + \dots + \frac{\varphi^{(k)}(c)}{k!}(x-c)^{\alpha+k}.$$

Интеграл от  $g(x)$  берется непосредственно.

Выражение

$$\left( \varphi(x) - \varphi(c) - \frac{\varphi'(c)}{1!}(x-c) - \dots - \frac{\varphi^{(k)}(c)}{k!}(x-c)^k \right) \quad (59)$$

обращается в нуль в точке  $x=c$  вместе со своими производными до порядка  $k$  включительно. Произведение этого выражения на многочлен  $(x-c)^{\alpha}$  будет функцией, непрерывной вместе с производными до порядка  $(k-1)$ . Поэтому интеграл от этой функции - собственный, то есть можно применить известные квадратурные формулы.

**Пример 7. [38]**

Вычислить приближенно интеграл

$$I = \int_0^{0.5} \frac{dx}{\sqrt{x(1-x)}}.$$

Решение.

Функция  $f(x) = \frac{1}{\sqrt{x(1-x)}} = x^{-\frac{1}{2}}(1-x)^{-\frac{1}{2}}$  имеет разрыв при  $x=0$ .

Таким образом,  $\alpha = -\frac{1}{2}$ ,  $c=0$ ,  $\varphi(x) = (1-x)^{-\frac{1}{2}}$  (следуя (57)).

По формуле Тейлора имеем

$$\varphi(x) = 1 + \frac{1}{2}x + \frac{3}{8}x^2 + \frac{5}{16}x^3 + \frac{35}{128}x^4 + R_4(x).$$

Тогда

$$f(x) = \left( x^{-\frac{1}{2}} + \frac{1}{2}x^{\frac{1}{2}} + \frac{3}{8}x^{\frac{3}{2}} + \frac{5}{16}x^{\frac{5}{2}} + \frac{35}{128}x^{\frac{7}{2}} \right) + \frac{\psi(x)}{\sqrt{x}},$$

где

$$\psi(x) = \frac{1}{\sqrt{1-x}} - \left( 1 + \frac{1}{2}x + \frac{3}{8}x^2 + \frac{5}{16}x^3 + \frac{35}{128}x^4 \right),$$

причем  $\psi(0) = 0$ .

Отсюда

$$I = \int_0^{0.5} \left( x^{-\frac{1}{2}} + \frac{1}{2}x^{\frac{1}{2}} + \frac{3}{8}x^{\frac{3}{2}} + \frac{5}{16}x^{\frac{5}{2}} + \frac{35}{128}x^{\frac{7}{2}} \right) dx + I_1 = 1.5691585 + I_1,$$

$$I_1 = \int_0^{0.5} \frac{\psi(x)}{\sqrt{x}} dx, \quad I_1 \text{ вычислен по формуле Симпсона.}$$

При  $n=10$   $I_1 = 0.0016385$ . Окончательно

$$I = \int_0^{0.5} \frac{dx}{\sqrt{x(1-x)}} = 1.570797.$$

Для мультипликативного способа  $f(x)$  представляется в виде произведения

$$f(x) = p(x) \cdot \varphi(x), \quad (60)$$

$\varphi(x)$  - ограничена,  $p(x) > 0$  и интегрируема на  $[a; b]$ . Тогда можно применить квадратурные формулы Гаусса-Кристоффеля, которые требуют вычисления лишь в узлах функции  $\varphi(x)$ ,  $p(x)$  - весовая функция.

**Пример 8.**

Вычислить приближенно интеграл

$$I = \int_0^1 \frac{dx}{(1+x)\sqrt{x}}.$$

Решение.

$f(x) = \frac{1}{(1+x)\sqrt{x}}$  имеет разрыв в точке  $x=0$ .

Представим  $f(x) = \frac{1}{\sqrt{x}} \cdot \frac{1}{1+x}$ , обозначим

$$p(x) = \frac{1}{\sqrt{x}}, \quad \varphi(x) = \frac{1}{1+x}, \quad I = \sum c_k^{(n)} \varphi(x_k).$$

При  $n=3$  для весовой функции  $p(x) = \frac{1}{\sqrt{x}}$

$$x_1 = 0.056919, \quad x_2 = 0.437198, \quad x_3 = 0.869499,$$

$$c_1^{(3)} = 0.935828, \quad c_2^{(3)} = 0.342649, \quad c_3^{(3)} = 0.342649.$$

$$I = 0.935828 \cdot \frac{1}{1+0.056939} + 0.72123 \cdot \frac{1}{1+0.437198} + 0.342649 \cdot \frac{1}{1+0.869499} = 1.570724.$$

• Для приближенного численного интегрирования сходящегося несобственного интеграла, где  $c$  - точка разрыва второго рода - внутренняя точка отрезка интегрирования  $[a; b]$ , выбирают  $\delta_1, \delta_2$  настолько малыми, чтобы

$$\left| \int_{c-\delta_1}^{c+\delta_1} f(x) dx \right| < \frac{\epsilon}{2}.$$

Затем вычисляют собственные интегралы  $\int_a^{c-\delta_1} f(x) dx$  и  $\int_{c+\delta_2}^b f(x) dx$ .

• Если  $c$  - точка разрыва I рода, то есть существуют  $\lim_{x \rightarrow c-0} f(x) = f(c-0)$  и  $\lim_{x \rightarrow c+0} f(x) = f(c+0)$ ,

то полагают

$$\int_a^b f(x) dx = \int_a^c f_1(x) dx + \int_c^b f_2(x) dx, \quad (61)$$

$f_1, f_2$  - непрерывны на отрезках  $[a; c]$  и  $[c; b]$  и равны

$$f_1(x) = \begin{cases} f(x), & \text{если } a \leq x \leq c; \\ f(c-0), & \text{если } x = c; \end{cases} \quad (62)$$

$$f_2(x) = \begin{cases} f(c+0), & \text{если } x = c; \\ f(x), & \text{если } c < x \leq b. \end{cases}$$

### 5.7. Кратные интегралы

Формулы численных кубатур предназначены для вычисления интегралов вида  $\iint_{\Omega} f(x) dx dy$ .

Пусть область интегрирования  $\Omega: \{y = \varphi(x), y = \psi(x), (\varphi(x) \leq \psi(x)), a \leq x \leq b\}$ .

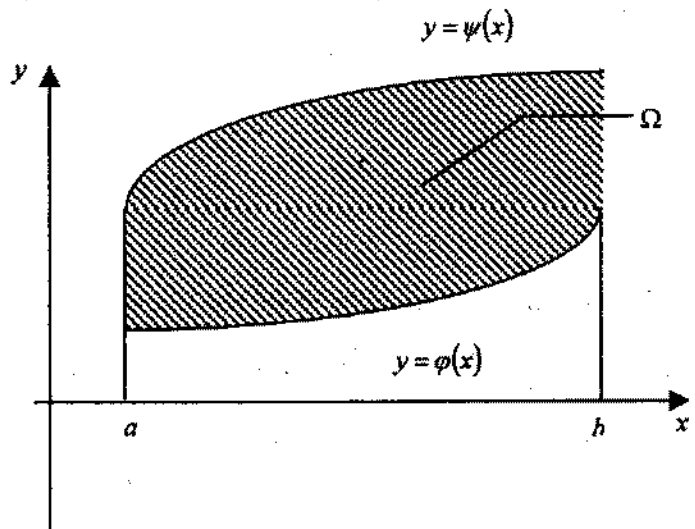


Рис. 5.15

$$I = \iint_{\Omega} f(x, y) dx dy = \int_a^b dx \int_{\varphi(x)}^{\psi(x)} f(x, y) dy. \quad (63)$$

Применяя к внутреннему интегралу одну из квадратурных формул, получим:

$$I = \sum_{i=1}^n A_i F(x_i) \quad (64)$$

где  $A_i$  - постоянные коэффициенты.

$F(x_i) = \int_{\varphi(x)}^{\psi(x)} f(x_i, y_i) dy$  можно также найти по некоторым квадратурным формулам.

$$F(x_i) = \sum_{j=1}^m B_{ij} f(x_i, y_j), \quad (65)$$

$$\iint_{\Omega} f(x, y) dx dy = \sum_{i=1}^n \sum_{j=1}^m A_i B_{ij} f(x_i, y_j). \quad (66)$$

### 5.8. Кубатурная формула Коровова

Формула Коровова используется при вычислении интегралов, если область  $\Omega: \{0 \leq x \leq 1, 0 \leq y \leq 1\}$ .

$$\iint_{00}^{11} f(x, y) dx dy = \frac{1}{U_m} \sum \left\{ \left\{ \frac{k}{U_m} \right\}, \left\{ \frac{kU_{m-1}}{U_m} \right\} \right\} + R, \quad (67)$$

$\{a\}$  - дробная часть числа  $a$ ,

$U_m$  - общий член последовательности Фибоначчи, описываемой рекуррентной формулой

$$U_m = U_{m-1} + U_{m-2}, \quad m = 2, 3, \dots \quad (68)$$

$$U_0 = 1, \quad U_1 = 1.$$

Погрешность формулы Коровова

$$R \approx \frac{\ln U_m}{U_m}. \quad (69)$$

При использовании формулы (67) число членов рекомендуется брать достаточно большим.

#### Пример 9.

Пользуясь формулой Коровова вычислить  $\iint_{00}^{11} xy^2 dx dy$  и оценить погрешность.

Решение.

Выберем число узлов квадратурной формулы  $m = 10$ ,  $U_{10} = 89$ .

$$\text{Тогда } \iint_{00}^{11} xy^2 dx dy = \frac{1}{89} \sum_{k=1}^{89} \left\{ \frac{k}{89} \right\} \cdot \left\{ \frac{55k}{89} \right\}^2 = 0.164.$$

Погрешность пропорциональна  $\frac{\ln 89}{89} \approx 0.05$ .

### 5.9. Кубатурная формула Симпсона

Пусть

$$\Omega: \{a \leq x \leq A; b \leq y \leq B\}, \quad (70)$$

$$h = \frac{A-a}{2}; \quad k = \frac{B-b}{2},$$

тогда

$$\iint_{\Omega} f(x,y) dx dy = \frac{hk}{9} (\sigma_0 + 4\sigma_1 + 16\sigma_2). \quad (71)$$

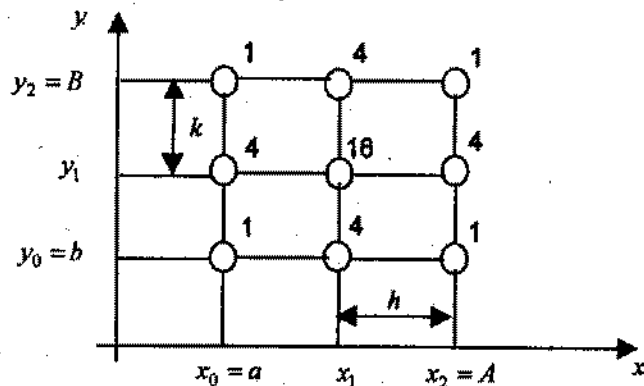


Рис.5.16

$\sigma_0$  - сумма значений подынтегральной функции  $f(x,y)$  в вершинах прямоугольника,

$\sigma_1$  - сумма значений  $f(x,y)$  в серединах сторон прямоугольника,

$\sigma_2$  - значение функции  $f(x,y)$  в центре.

Кратности этих значений - на рис. 16.

При недостаточной точности вычисления интеграла можно воспользоваться «сгущением» сетки, например

$$h = \frac{A-a}{2n}, \quad k = \frac{B-b}{2m},$$

тогда координаты сетки узлов

$$x_i = x_0 + ih, \quad x_0 = a, \quad i = 0, 2n, \quad (72)$$

$$y_j = y_0 + jk, \quad y_0 = b, \quad j = 0, 2m.$$

$$\iint_{\Omega} f(x,y) dx dy = \frac{hk}{9} \sum_{i=0}^{2n} \sum_{j=0}^{2m} \lambda_{ij} f_{ij}. \quad (73)$$

### Схема алгоритма интегрирования по методу кубатур

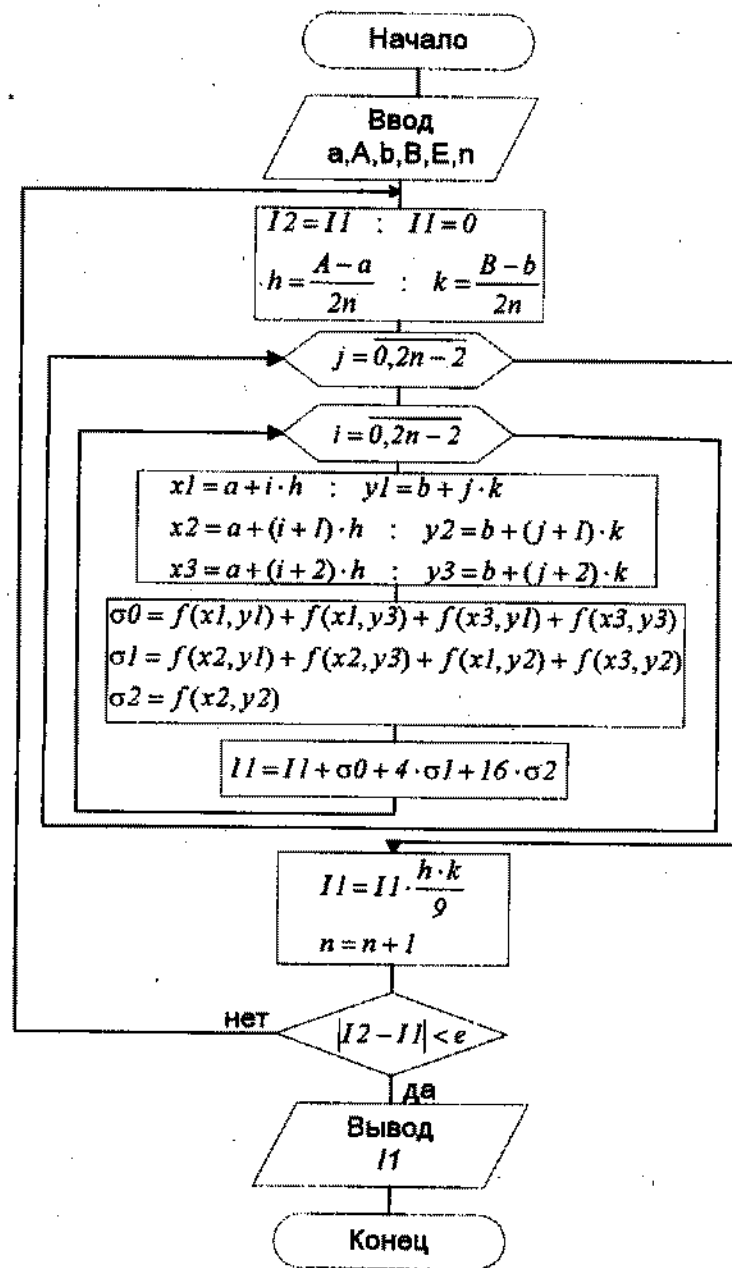


Рис.5.17

Координаты  $\lambda_{ij}$  - элементы матрицы  $\Lambda$ .

$$\Lambda = \begin{pmatrix} 1 & 4 & 2 & 4 & 2 & \dots & 4 & 2 & 4 & 1 \\ 4 & 16 & 8 & 16 & 8 & \dots & 16 & 8 & 16 & 4 \\ 2 & 8 & 4 & 8 & 4 & \dots & 8 & 4 & 8 & 2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 2 & 8 & 4 & 8 & 4 & \dots & 8 & 4 & 8 & 2 \\ 4 & 16 & 8 & 16 & 8 & \dots & 16 & 8 & 16 & 4 \\ 1 & 4 & 2 & 4 & 2 & \dots & 4 & 2 & 4 & 1 \end{pmatrix} \quad (74)$$

Если область  $\Omega$  не прямоугольник, то строим прямоугольник  $R \in \Omega$ , стороны которого параллельны осям координат. Рассмотрим функцию

$$F(x, y) = \begin{cases} f(x, y), & \text{если } x, y \in \Omega, \\ 0, & \text{если } x, y \in R - \Omega \end{cases}$$

$$\text{Тогда } \iint_{\Omega} f(x, y) dx dy = \iint_R F(x, y) dx dy.$$

Формулы типа Симпсона достаточно трудоемки, поэтому предпочтительно интегрировать по формуле Гаусса, которая точна при  $n$  ординатах для многочленов  $(2n-1)$  степени.

Обозначим

$$y_j = \frac{1}{2}(\alpha + \beta) + \frac{1}{2}(\beta - \alpha)t_j, \quad (75)$$

$$c_{ij} = \frac{1}{4}(b-a)(\beta-a)\gamma_i\gamma_j, \quad 1 \leq i, j \leq n,$$

где  $t, \gamma$  - нули многочленов Лежандра и соответствующие веса,

$$I = \iint_{\alpha a}^{\beta b} f(x, y) dx dy = \sum_i \sum_j c_{ij} f(x_i, y_j). \quad (76)$$

Формулы (75), (76) рассчитаны на функции высокой гладкости и дают для них большую экономию в числе узлов. Например, для  $m$  измерений кубатурная формула Симпсона с  $3^m$  узлами и формулы (75-76) с  $2^m$  узлами дают примерно одинаковую точность, хотя формула Гаусса имеет при  $m=2$  вдвое меньше узлов, при  $m=3$  втрое меньше узлов, чем кубатурная формула Симпсона.

#### Пример 10.

Вычислить  $I = \iint_{-1}^{12} xy^2 dx dy$  пользуясь кубатурной формулой

Симпсона.

Решение.

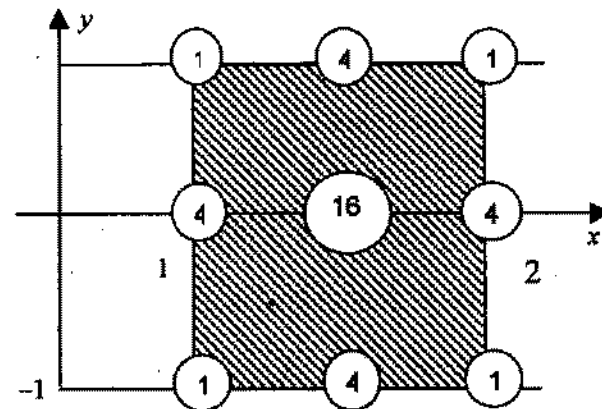


Рис. 5.18

Согласно (70-71)  $h = \frac{1}{2}$ ,  $k=1$ ,

$$I = \frac{1}{18}(\sigma_0 + 4\sigma_1 + 16\sigma_2).$$

$$\sigma_0 = 1+1+2+2=6, \quad \sigma_1 = 0 + \frac{3}{2} + 0 + \frac{3}{2} = 3, \quad \sigma_2 = 0, \quad I = 1.$$

### 5.10. Метод Монте-Карло вычисления двойных интегралов

Пусть требуется вычислить

$$\iint_{\Omega} f(x, y) dx dy, \quad \Omega: \{0 \leq x \leq 1, 0 \leq y \leq 1\}.$$

#### Алгоритм решения задачи

1. Выбрать  $N$  случайных, равномерно распределенных на  $[0,1]$  чисел. Каждые два числа принять за координаты случайной точки  $M(x, y)$ .
2. Определить количество точек  $k$ , принадлежащих области  $\Omega$  - области интегрирования.
3. Вычислить значения подынтегральной функции  $f(x, y)$  в этих точках.
4. Вычислить  $S = \frac{1}{N} \sum_{i=1}^k f(M_i)$ .



При достаточно большом значении  $N$  справедлива формула

$$I = \iint_{\Omega} f(x, y) dx dy = \frac{1}{N} \sum_{i=1}^k f(M_i). \quad (77)$$

**Пример 11.**

Методом Монте-Карло вычислить

$$I = \iint_{\Omega} \sqrt{x^2 - y^2} dx dy$$

по области  $\Omega$  - треугольнику с вершинами  $O(0,0)$ ,  $A(1,0)$ ,  $B(1,1)$ .

*Решение.*

Область интегрирования определяется неравенствами  $0 \leq x \leq 1$ ,  $0 \leq y \leq 1$ . (78)

принадлежит единичному квадрату:  $0 \leq x \leq 1$ ,  $0 \leq y \leq 1$ .

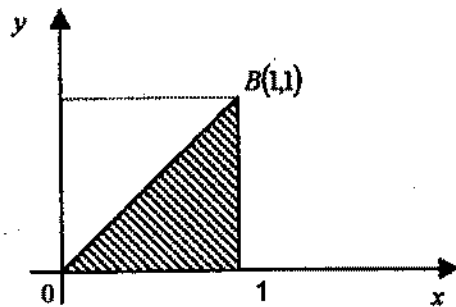


Рис.5.19

При выполнении п.2 алгоритма следует определять количество случайных точек, удовлетворяющих условиям (78).

Задача решена при  $N=1000$ , области  $\Omega$  принадлежит  $k=496$  точек. По формуле (77) значение интеграла  $I=0.2662857$ .

Структурная схема алгоритма интегрирования по методу Монте-Карло приведена на рис. 20.

**5.11. Вычисление неопределенного интеграла**

Рассматривается задача [44] вычисления функции

$$y(x) = y_0 + \int_{x_0}^x f(t) dt, \quad (x_0 \leq x \leq X) \quad (79)$$

на сетке равноотстоящих значений аргумента

$$x_k = x_0 + hk, \quad k = 0, 1, \dots, N, \quad N = \left[ \frac{X - x_0}{h} \right], \quad h > 0. \quad (80)$$

**Структурная схема алгоритма интегрирования методом Монте-Карло**

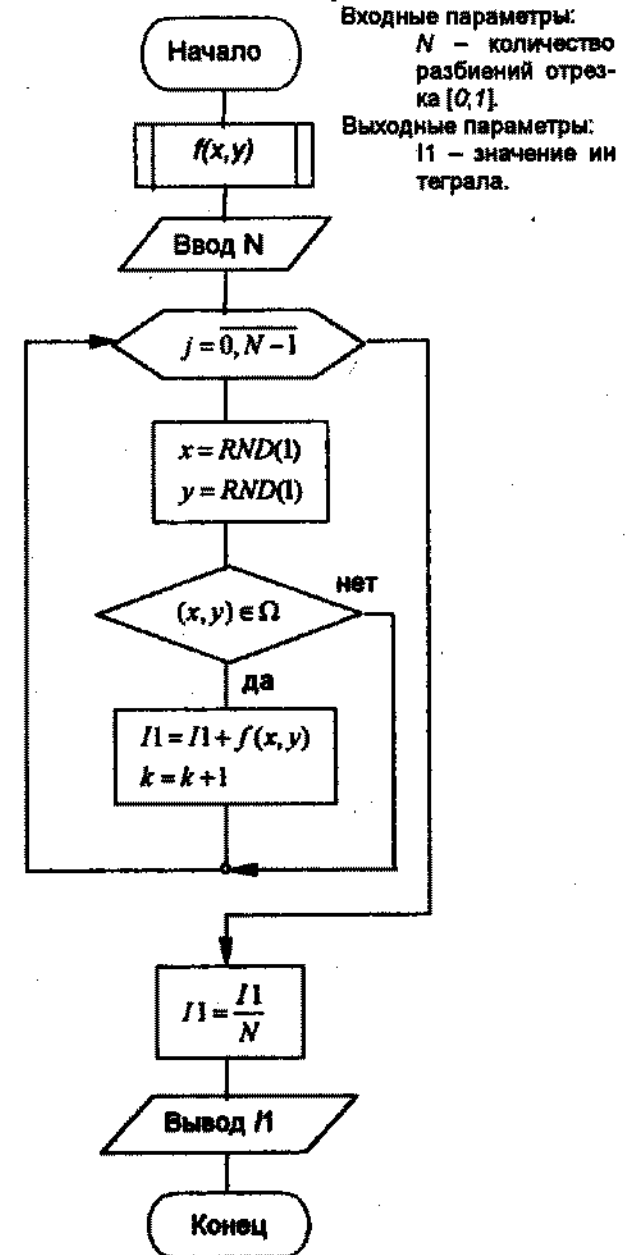


Рис. 5.20

Если  $f(x)$  задана таблично, то для вычисления  $y_{n+1} = y(x_{n+1})$  можно пользоваться любыми, ранее определенными значениями  $y_k = y(x_k)$ ,  $k = 0, 1, \dots, N$ .

• Вычисление интеграла в точках вблизи начала  $x_0$  таблицы.

а)

$$y(x_1) = y(x_0) + h \left[ \frac{f_0 + f_1}{2} - \frac{1}{12} \Delta^2 f_0 + \frac{1}{24} \Delta^3 f_0 - \frac{19}{720} \Delta^4 f_0 + \frac{1}{160} \Delta^5 f_0 + \dots + C_k \Delta^k f_0 \right] + R_{nk} \quad (81)$$

$k \geq 2$ ,

$$C_k = \frac{1}{k!} \int_0^1 u(u-1) \dots (u-k+1) du,$$

$$R_{nk} = h^{k+2} \frac{f^{(k+1)}(\xi)}{(k+1)!} \int_0^1 u(u-1) \dots (u-k) du, \\ x_0 \leq \xi \leq x_k.$$

б)

$$y(x_2) = y(x_1) + h \left[ \frac{f_1 + f_2}{2} - \frac{1}{12} \frac{\Delta^2 f_0 + \Delta^2 f_1}{2} + \frac{11}{720} \Delta^4 f_0 - \frac{11}{1440} \Delta^5 f_0 + \frac{271}{60480} \Delta^6 f_0 + \dots + D_k \Delta^k f_0 \right] + R_{nk} \quad (82)$$

$k \geq 4$ ,

в)

$$y(x_3) = y(x_2) + h \left[ \frac{f_2 + f_3}{2} - \frac{1}{12} \frac{\Delta^2 f_1 + \Delta^2 f_2}{2} + \frac{11}{720} \frac{\Delta^4 f_0 + \Delta^4 f_1}{2} - \frac{191}{60480} \Delta^6 f_0 + \frac{191}{120960} \Delta^7 f_0 + \dots + E_k \Delta^k f_0 \right] + R_{nk} \quad (83)$$

$k \geq 6$ ,

$$E_k = \frac{1}{k!} \int_0^1 (u+2)(u+1) \dots (u-k+3) du,$$

$$R_{nk} = h^{k+2} \frac{f^{(k+1)}(\xi)}{(k+1)!} \int_0^1 (u+2)(u+1) \dots (u-k+2) du$$

$x_0 \leq \xi \leq x_k$ .

• Вычисление интеграла в точках вблизи конца таблицы  $x_N$  производится по формулам, аналогичным (81-83) путем переноса начала отсчета в  $x_N$  и изменения направления оси  $Ox$ .

а)

$$y(x_N) = y(x_{N-1}) + h \left[ \frac{f_N + f_{N-1}}{2} - \frac{1}{12} \Delta^2 f_{N-2} + \frac{1}{24} \Delta^3 f_{N-3} - \frac{19}{720} \Delta^4 f_{N-4} + \frac{1}{160} \Delta^5 f_{N-5} - \dots - (-1)^{k-1} C_k \Delta^k f_{N-k} \right] + R'_{nk} \quad (84)$$

$k \geq 2$ ,

б)

$$y(x_{N-1}) = y(x_{N-2}) + h \left[ \frac{f_{N-2} + f_{N-1}}{2} - \frac{1}{12} \frac{\Delta^2 f_{N-2} + \Delta^2 f_{N-3}}{2} + \frac{11}{720} \Delta^4 f_{N-4} - \frac{11}{1440} \Delta^5 f_{N-5} + \frac{271}{60480} \Delta^6 f_{N-6} + \dots + (-1)^k D_k \Delta^k f_{N-k} \right] + R'_{nk} \quad (85)$$

$k \geq 4$ ,

в)

$$y(x_{N-2}) = y(x_{N-1}) + h \left[ \frac{f_{N-3} + f_{N-2}}{2} - \frac{1}{12} \frac{\Delta^2 f_{N-4} + \Delta^2 f_{N-3}}{2} + \frac{11}{720} \frac{\Delta^4 f_{N-5} + \Delta^4 f_{N-4}}{2} - \frac{191}{60480} \Delta^6 f_{N-6} + \frac{191}{120960} \Delta^7 f_{N-7} - \dots - (-1)^{k-1} E_k \Delta^k f_{N-k} \right] + R'_{nk} \quad (86)$$

$k \geq 6$ .

$$R'_{nk} = (-1)^{k+1} R_{nk}, \quad x_{N-k} \leq \xi \leq x_N.$$

• Вычисления в точках, не близких к концу и началу таблицы.

Для вычисления  $y(x_{n+1}) = y(x_n) + \int_{x_n}^{x_{n+1}} f(x) dx$  интерполирование

$f(x)$  на отрезке  $[x_n, x_{n+1}]$  может быть выполнено по формуле Бесселя, что приводит к правилу

$$y(x_{n+1}) = y(x_n) + h \left[ \frac{f_n + f_{n+1}}{2} - \frac{1}{12} \frac{\Delta^2 f_{n-1} + \Delta^2 f_n}{2} + \frac{11}{720} \frac{\Delta^4 f_{n-2} + \Delta^4 f_{n-1}}{2} - \frac{191}{60480} \frac{\Delta^6 f_{n-3} + \Delta^6 f_{n-2}}{2} + \dots + C_k \frac{\Delta^{2k} f_{n-2} + \Delta^{2k} f_{n-k+1}}{2} \right] + R_{nk}, \quad (87)$$

$$R_{nk} = h^{2k+3} \frac{f^{(2k+2)}(\xi)}{(2k+2)!} \int_0^1 (u+k)(u+k-1) \dots (u-k-1) du,$$

$$\xi \in [x_n - kh, x_n + (k+1)h].$$

Пусть теперь функция  $f(x)$  может быть вычислена в любых точках отрезка  $[x_0; X]$ .

При неопределенном интегрировании вычисления выполняются для многих значений верхней границы. Доказано, что правила интегрирования можно строить так, чтобы каждое вычисленное значение функции  $f(x)$  применялось для нахождения не одного, а нескольких значений интеграла  $y(x)$ . Другой подход – правила вычислений, использующие несколько предшествующих значений интеграла [44].

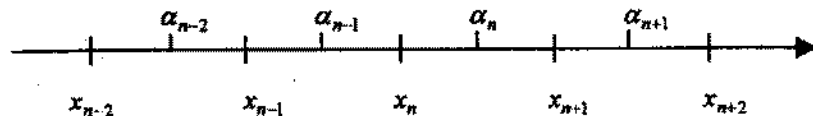
Укажем лишь идею предложенного подхода к решению задачи вычисления неопределенных интегралов.

Точки  $a + hk$ , отвечающие целым значениям  $k$  называются сходственными.

Пусть для нахождения  $y_{n+1}$ , используется лишь одно значение  $y_n$ , тогда

$$y(x_{n+1}) = y(x_n) + \int_{x_n}^{x_{n+1}} f(x) dx, \quad (88)$$

причем на шаге  $[x_n; x_{n+1}]$  берется одна точка  $\alpha_n$ . Кроме того, привлечены по  $k$  ближайших точек слева и справа от  $\alpha_n$ .



Если  $\alpha_n = x_n + \frac{1}{2}h$ , то наивысшая степень точности равна  $2k+1$

Расчетная формула

$$y(x_{n+1}) = y(x_n) + h \left[ f(\alpha_n) + \frac{1}{24} \Delta^2 f(\alpha_n - h) - \frac{17}{5760} \Delta^4 f(\alpha_n - 2h) + \frac{367}{967680} \Delta^6 f(\alpha_n - 3h) + \dots + C_k \Delta^{2k} f(\alpha_n - kh) \right] + R, \quad (89)$$

$$C_k = \frac{1}{(2k)!} \int_{-0.5}^{0.5} t^2 (t^2 - 1^2) \dots (t^2 - (k-1)^2) dt,$$

$$R = h^{2k+3} \frac{f^{(2k+3)}(\xi)}{(2k+3)!} \int_{-0.5}^{0.5} t^2 (t^2 - 1^2) \dots (t^2 - k^2) dt.$$

Существуют правила вычисления с двумя, тремя узлами на шаге. Вид частных правил вычисления зависит от числа дополнительных узлов (сходственных точек) [44].

## 6. ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ

Значительная часть инженерных задач, возникающих в самых разных практических областях, сводится к решению обыкновенных дифференциальных уравнений. Такого рода уравнениями можно описать как траекторию движения небесных тел, так и модели нелинейной автоколебательной цепи с двухполусником, имеющим N-образную вольт-амперную характеристику.

Алгоритм расчета ключевых электронных устройств методом переменных состояний заключается в составлении и решении системы дифференциальных уравнений состояния, описывающей работу моделируемой схемы.

Ряд задач, описываемых уравнениями в частных производных, также сводятся к решению обыкновенных дифференциальных уравнений. Например, задачи нахождения собственных колебаний мембраны, определение спектра собственных значений энергии частицы и другие.

Различают три основных типа задач для обыкновенных дифференциальных уравнений: задачи Коши, краевые задачи и задачи на собственные значения.

### 6.1. Методы Рунге-Кутты

Задача Коши для дифференциального уравнения первого порядка заключается в нахождении решения уравнения

$$y' = f(x, y), \quad (1)$$

удовлетворяющего при  $x = x_0$  начальному условию

$$y(x_0) = y_0. \quad (2)$$

Остановимся на методах Рунге-Кутты решения задачи (1) – (2).

Все методы Рунге-Кутты являются одношаговыми методами: для получения решения  $y_{i+1}$  при  $x_{i+1}$  надо знать решение в одной точке  $(x_i, y_i)$ . Это позволяет начать решение используя начальные условия. Указанная особенность допускает изменение шага интегрирования в любой точке в процессе счета, что дает возможность строить численные алгоритмы с автоматическим выбором шага.

На каждом шаге интегрирования методы Рунге-Кутты требуют многократного вычисления правой части уравнения (1).

Решения методами Рунге-Кутты согласуются с решениями, получающимися при разложении в ряд Тейлора, в котором сохра-

няются члены, содержащие  $h^p$  включительно. Целое число  $p$  называют порядком метода.

Погрешность методов Рунге-Кутты

$$|R| < c h^{p+1},$$

где  $c = \text{const}$ ,  $p$  - порядок метода.

К недостаткам методов Рунге-Кутты можно отнести отсутствие аналитических оценок погрешности.

Интегрируя (1) в пределах от  $x$  до  $x+h$  ( $h = \text{const}$ ), получим

$$y(x+h) = y(x) + \int_x^{x+h} f(x, y) dx \quad (3)$$

В зависимости от вида функции, аппроксимирующей подынтегральную функцию, (константа; прямая, проходящая через концы отрезка; парабола и т. д.) получают различные формулы численного интегрирования дифференциального уравнения.

- Метод Рунге-Кутты первого порядка точности (метод Эйлера)

$$y_{i+1} = y_i + h f(x_i, y_i), \quad i = 0, 1, \dots \quad (4)$$

- Методы Рунге-Кутты второго порядка точности (исправленный метод Эйлера и модифицированный метод Эйлера).

$$1) \begin{cases} \tilde{y}_{i+1} = y_i + h f(x_i, y_i); \\ y_{i+1} = y_i + \frac{h}{2} (f(x_i, y_i) + f(x_i + h, \tilde{y}_{i+1})); \end{cases} \quad (5)$$

$$2) y_{i+1} = y_i + h f(x_i + \frac{h}{2}, y_i + \frac{h}{2} f(x_i, y_i)), \quad i = 0, 1, \dots \quad (6)$$

- Метод Рунге-Кутты третьего порядка точности

$$\begin{cases} y_{i+1} = y_i + \frac{1}{4} (h f(x_i, y_i) + 3h f(x_i + \frac{2}{3}h, y_i + \frac{2}{3}\bar{y}_i)); \\ \bar{y}_i = h f(x_i + \frac{h}{3}, y_i + \frac{h}{3} f(x_i, y_i)). \end{cases} \quad (7)$$
$$i = 0, 1, \dots$$

- Методы Рунге-Кутты четвертого порядка точности

$$1) y_{i+1} = y_i + \frac{1}{6} (k_0 + 2k_1 + 2k_2 + k_3)$$

$$k_0 = h f(x_i, y_i), \quad i = 0, 1, \dots,$$

$$k_1 = h f(x_i + \frac{h}{2}, y_i + \frac{k_0}{2}), \quad (8)$$

$$k_2 = h f(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}),$$

$$k_3 = h f(x_i + h, y_i + k_2),$$

$$2) \quad y_{i+1} = y_i + \frac{1}{8}(k_0 + 3k_1 + 3k_2 + k_3)$$

$$k_0 = h f(x_{i+1}, y_i), \quad i = 0, 1, \dots, \quad (9)$$

$$k_1 = h f(x_i + \frac{h}{3}, y_i + \frac{k_0}{3}), \quad k_2 = h f(x_i + \frac{2}{3}h, y_i - \frac{k_0}{3} + k_1),$$

$$k_3 = h f(x_i + h, y_i + k_0 - k_1 + k_2).$$

Для контроля правильности выбора шага метода Рунге-Кутты 4-го порядка рекомендуется [73] вычислять дробь

$$\Theta = \frac{|k_1^{(j)} - k_2^{(j)}|}{|k_0^{(j)} - k_1^{(j)}|}.$$

Величина  $\Theta$  не должна превышать нескольких сотых. В противном случае шаг следует уменьшать.

Оценку погрешности решения методами Рунге-Кутты проводят следующим образом (метод двойного просчета, правило Рунге).

Пусть  $E$  - требуемая точность,

$y_h$  - решение, полученное с шагом  $h$ ,

$y_{h/2}$  - решение, полученное с шагом  $\frac{h}{2}$ ,

$p$  - порядок метода.

$$\text{Если } \frac{|y_h - y_{h/2}|}{2^p - 1} < E, \quad (10)$$

то можно продолжить решение с шагом  $\frac{h}{2}$ , полагая

$$y(x) = y_{h/2}(x) + \frac{y_{h/2}(x) - y_h(x)}{2^p - 1}, \quad (11)$$

иначе - уменьшать шаг интегрирования дифференциального уравнения.

## 6.1.1. Геометрическая иллюстрация методов Рунге-Кутта

### 1) Метод Эйлера (4)

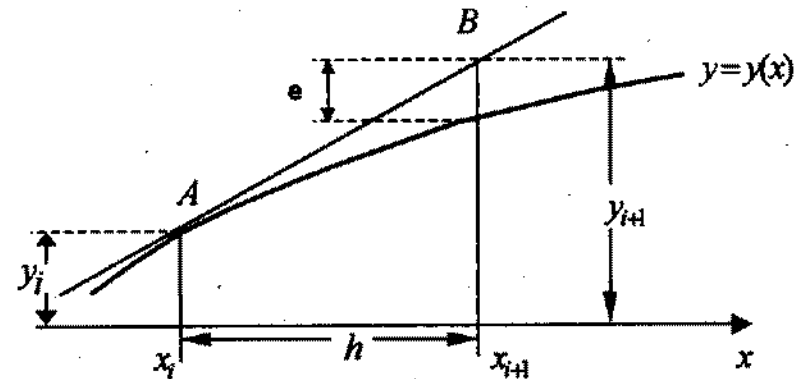


Рис. 6.1

AB - прямая, проведенная через точку  $(x_i, y_i)$  под углом  $\alpha$ ,  $\operatorname{tg} \alpha = f(x_i, y_i)$

### 2) Исправленный метод Эйлера (5)

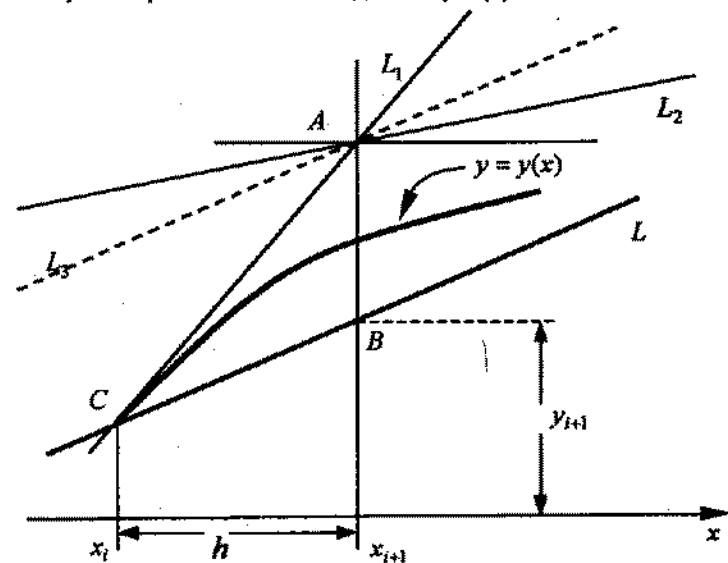


Рис. 6.2

а) прямая  $L_1$ , точка  $A(x_i + h, \bar{y}_{i+1})$  находится по методу Эйлера,  $tg\alpha_1 = f(x_i, y_i)$ ,

б) в точке  $A$  по методу Эйлера вычисляется тангенс угла наклона прямой  $L_2$  к оси  $OX$   $tg\alpha_2 = f(x_{i+1}, \bar{y}_{i+1})$ ,

в) тангенс угла наклона прямой  $L_3$  равен  $(\frac{tg\alpha_1 + tg\alpha_2}{2})$ ,

г) прямая  $L$  проходит через точку  $C(x_i, y_i)$  параллельно прямой  $L_3$ ,

$B(x_{i+1}, y_{i+1})$  - решение при  $x = x_{i+1}$  по формулам (5).

3) Модифицированный метод Эйлера (6)

а) прямая  $L_1$  - через точку  $C(x_i, y_i)$  (по формуле Эйлера),

$$A(x_i + \frac{h}{2}, y_i + \frac{h}{2} f(x_i, y_i)),$$

б)  $L_2$  - прямая, проведенная под углом  $\alpha$ , определяемым из условия  $tg\alpha = f(x_i + \frac{h}{2}, y_i + \frac{h}{2} f(x_i, y_i))$ ,

в) прямая  $L$  проходит через точку  $C(x_i, y_i)$  параллельно прямой  $L_2$ ,

$B(x_{i+1}, y_{i+1})$  - решение при  $x = x_{i+1}$ , вычисленное по формулам (6).

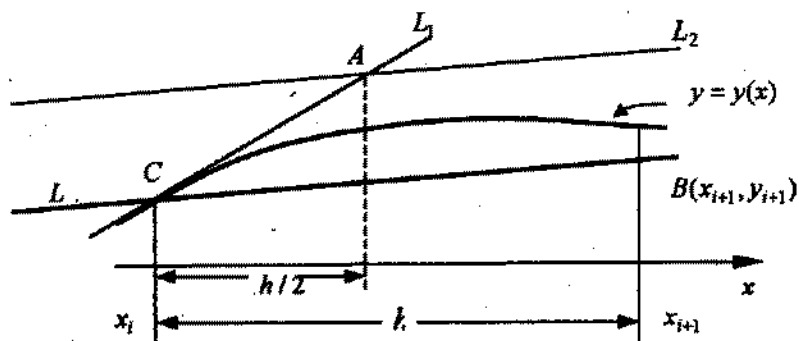
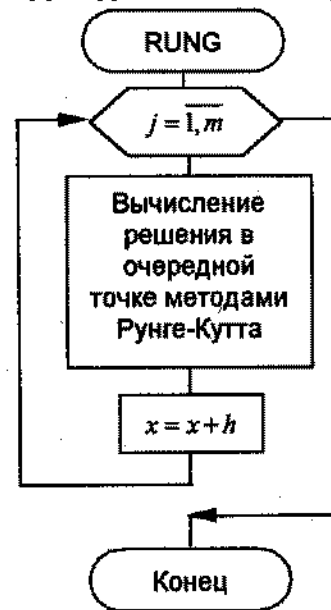


Рис. 6.3

### Структурная схема алгоритма метода Рунге-Кутты



Входные параметры:  
 $x, y$  - начальная точка;  
 $h$  - шаг;  
 $m$  - количество дроблений;  
 начального шага  $h_0$ .  
 Выходные параметры:  
 $y$  - решение.

Рис. 6.4

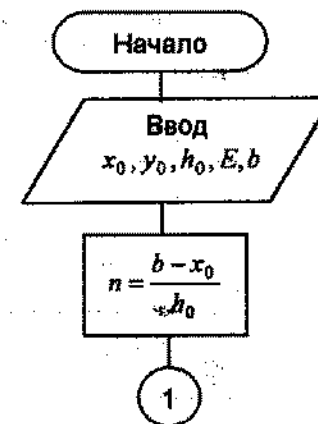
Входные параметры:

$(x_0, y_0)$  - начальные условия;

$h_0$  - начальный шаг;

$E$  - погрешность;

$x_0, b$  - границы отрезка интегрирования.



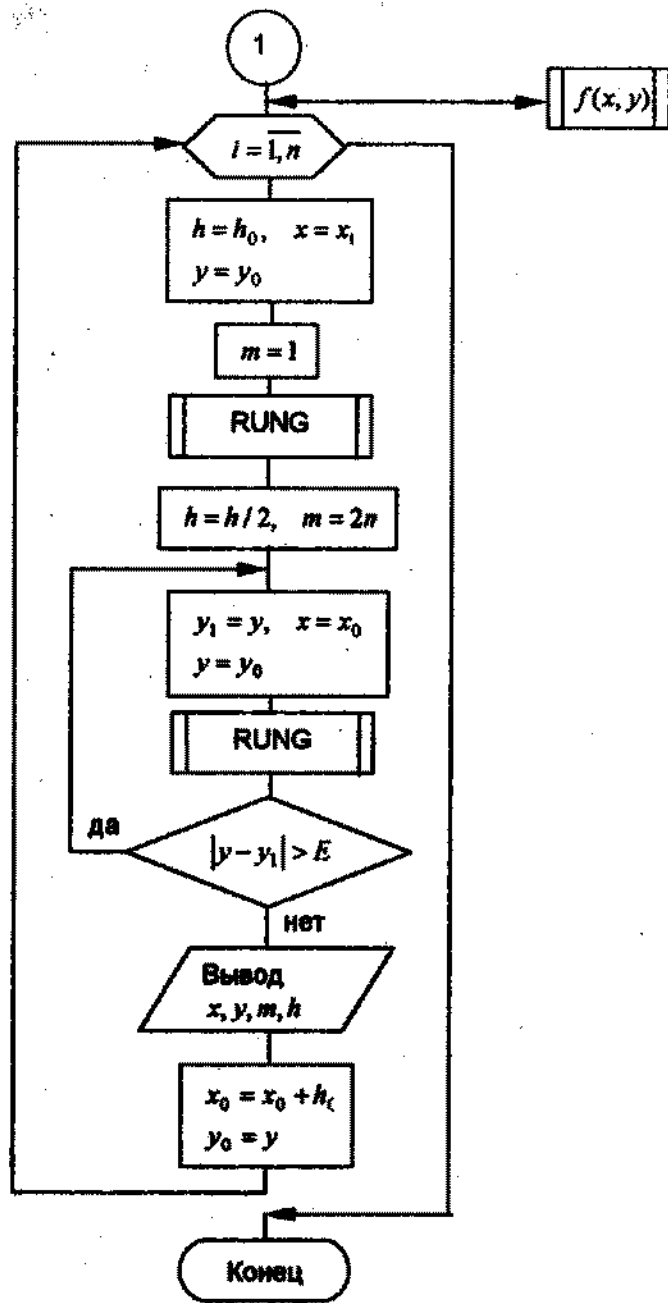


Рис. 6.5

**Пример 1. [38]**

Методом Рунге-Кутты 4-го порядка найти с точностью  $E = 0.5 \cdot 10^{-6}$  решение дифференциального уравнения

$$y' = \frac{sh(0.5y+x)}{1.5} + 0.5y, \quad y_0(0) = 0, \text{ на отрезке } [0; 0.2].$$

*Решение.*

Для выбора шага вычислим решение в точке  $x = 0.1$  с  $h = 0.1$  и  $h = 0.05$ .

При вычислении с  $h = 0.1$  по формулам (8) последовательно имеем

$$k_0^{(0)} = 0,$$

$$k_1^{(0)} = 0.1 \cdot \frac{sh0.05}{1.5} = 0.003335,$$

$$k_2^{(0)} = 0.1 \left( \frac{sh(0.5 \cdot 0.001667 + 0.05)}{1.5} + 0.5 \cdot 0.001667 \right) = 0.003475,$$

$$k_3^{(0)} = 0.1 \left( \frac{sh(0.5 \cdot 0.003475 + 0.1)}{1.5} + 0.5 \cdot 0.003475 \right) = 0.006969,$$

и, следовательно,  $y_1 = y_0 + \Delta y_0 = 0.003432$ .

Аналогично вычисляем  $y_1$  с шагом 0.05.  $y_1 = 0.003432$ .

Полученные результаты совпадают в пределах заданной точности, поэтому продолжим вычисления с шагом  $h = 0.1$  и  $h = 0.2$  (результаты в таблице). Решение с  $h = 0.1$ .

Таблица 1

x	y	f(x,y)	k = hf(x,y)	Δ y
0.1	0.003432	0.06965	0.006965	0.010726
0.15	0.006914	0.10617	0.010617	
0.15	0.008740	0.10770	0.010770	
0.2	0.014202	0.14615	0.014615	
0.2	0.014158			

Решение с  $h = 0.2$

Таблица 2

x	y	f(x,y)	k = hf(x,y)	Δ y
0'	0	0	0	0.014155
0.1	0	0.06678	0.013356	
0.1	0.006678	0.07235	0.014470	
0.2	0.014470	0.14638	0.029276	
0.2	0.014155			

Сравнение расчетов показывает, что с  $E=0.5 \cdot 10^{-6}$  можно принять  $y(0.2) \cong 0.014158$  и что в дальнейшем шаг расчета следовало бы снова удвоить.

### Пример 2.

Решить уравнение

$$y' = 2x^2 + 2y, \quad y(0) = 1, \quad h = 0.1, \quad 0 \leq x \leq 1.$$

Решение.

Точное решение уравнения -  $y(x) = 1.5e^{2x} - (x^2 + x + 0.5)$ .

Результаты расчета приведены в таблице, из которой отчетливо видны преимущества метода Рунге-Кутта 4-го порядка перед остальными

Таблица 3

x	Метод Эйлера	Модифицированный метод Эйлера	Метод Рунге-Кутта 4-го порядка	Точное решение
0.0	1.0000	1.0000	1.0000	1.0000
0.1	1.2000	1.2210	1.2221	1.2221
0.2	1.4420	1.4923	1.4977	1.4977
0.3	1.7384	1.8284	1.8432	1.8432
0.4	2.1041	2.2466	2.2783	2.2783
0.5	2.5569	2.7680	2.8274	2.8274
0.6	3.1183	3.4176	3.5201	3.5202
0.7	3.8139	4.2257	4.3927	4.3928
0.8	4.6747	5.2288	5.4894	5.4895
0.9	5.7376	6.4704	6.8643	6.8645
1.0	7.0472	8.0032	8.5834	8.5836

## 6.2. Метод Рунге - Кутта - Мерсона

Метод Рунге - Кутта - Мерсона - модификация метода Рунге-Кутта 4-го порядка, позволяющий оценивать погрешность на каждом шаге и принимать решение об изменении величины шага интегрирования. Использование схемы Мерсона [75] требует вычисления правой части уравнения в пяти точках.

$$y(x_0 + h) = y_0 + (k_3 + k_4) / 2, \quad (12)$$

где

$$k_0 = h_3 f(x_0, y_0), \quad h_3 = h/3,$$

$$k_1 = h_3 f(x_0 + h_3, y_0 + k_0),$$

$$k_2 = h_3 f(x_0 + h_3, y_0 + (k_0 + k_1) / 2),$$

$$k_3 = k_1 + 4h_3 f(x_0 + h/2, y_0 + 0.375(k_0 + k_2)),$$

$$k_4 = h_3 f(x_0 + h, y_0 + 1.5(k_3 - k_2)).$$

На каждом шаге можно вычислить погрешность

$$R = (2k_3 - 3k_2 - k_4) / 10. \quad (13)$$

Если  $|R| \geq E$ , то шаг уменьшается и вычисления повторяются с точки  $(x_0, y_0)$  по формулам (12).

Если  $|R| < \frac{E}{32}$ , то  $h$  можно увеличить. Как правило, кратность шагов принимается равной 2.

## 6.3. Метод А. Н. Крылова последовательных сближений

Рассмотрим способ построения решения дифференциального уравнения  $y' = f(x, y)$  с начальными условиями  $y(x_0) = y_0$ , применяемом, в том числе, и при табличном задании  $f(x, y)$ .

Заменяем производную искомого интеграла дифференциального уравнения  $y' = f(x, y)$  интерполяционным многочленом Ньютона, ограничиваясь конечными разностями второго порядка. После ряда преобразований получим, вводя обозначения:

$$q_k = hf(x_k, y_k), \quad (14)$$

$$\Delta y_0 = y_1 - y_0 = q_0 + \frac{1}{2} \Delta q_0 - \frac{1}{12} \Delta^2 q_0, \quad (15)$$

$$\Delta y_1 = y_2 - y_1 = q_1 + \frac{1}{2} \Delta q_0 + \frac{5}{12} \Delta^2 q_0, \quad (16)$$

$$\Delta y_2 = y_3 - y_2 = q_2 + \frac{1}{2} \Delta q_1 + \frac{5}{12} \Delta^2 q_0. \quad (17)$$

Запишем последовательность вычислений для уравнения  $y' = f(x, y)$  с начальными условиями  $(x = x_0, y = y_0)$ .

$$q_0 = hf(x_0, y_0).$$

Первое сближение. (верхний индекс - номер сближения - единица)

Ограничиваясь одним слагаемым в формуле (15), вычислим



$$\Delta y_0^{(1)} = y_1^{(1)} - y_0 = q_0,$$

$$y_1^{(1)} = y_0 + q_0.$$

Затем вычислим

$$q_1^{(1)} = hf(x_1, y_1^{(1)}), \quad \Delta q_0^{(1)} = q_1^{(1)} - q_0. \quad (18)$$

Второе сближение проводим по двучленным формулам (15), (16):

$$\Delta y_0^{(2)} = y_1^{(2)} - y_0 = q_0 + \frac{1}{2} \Delta q_0^{(1)},$$

$$\Delta y_1^{(2)} = y_2^{(2)} - y_1^{(2)} = q_1^{(1)} + \frac{1}{2} \Delta q_0^{(1)}.$$

Отсюда получаем

$$y_1^{(2)} = y_0 + \Delta y_0^{(2)} \text{ и } y_2^{(2)} = y_1^{(2)} + \Delta y_1^{(2)}.$$

По значениям  $y_1^{(2)}$  и  $y_2^{(2)}$  находим

$$q_1^{(2)} = hf(x_1, y_1^{(2)}), \quad q_2^{(2)} = hf(x_2, y_2^{(2)}), \quad (19)$$

Затем

$$\Delta q_1^{(2)} = q_2^{(2)} - q_1^{(2)}, \quad \Delta q_0^{(2)} = q_1^{(2)} - q_0,$$

$$\Delta^2 q_0^{(2)} = \Delta q_1^{(2)} - \Delta q_0^{(2)}.$$

Третье сближение проводится по полным формулам (15)-(17)

$$\Delta y_0^{(3)} = q_0 + \frac{1}{2} \Delta q_0^{(2)} - \frac{1}{12} \Delta^2 q_0^{(2)},$$

$$\Delta y_1^{(3)} = q_1^{(2)} + \frac{1}{2} \Delta q_0^{(2)} + \frac{5}{12} \Delta^2 q_0^{(2)}, \quad (20)$$

$$\Delta y_2^{(3)} = q_2^{(2)} + \frac{1}{2} \Delta q_1^{(2)} + \frac{5}{12} \Delta^2 q_0^{(2)},$$

откуда

$$y_1^{(3)} = y_0 + \Delta y_0^{(3)}, \quad y_2^{(3)} = y_1^{(3)} + \Delta y_1^{(3)}, \quad y_3^{(3)} = y_2^{(3)} + \Delta y_2^{(3)}.$$

### Пример 3.

Методом последовательных сближений найти решение уравнения

$$y' = y + x, \quad y(0) = 1, \quad h = 0.1 \text{ на отрезке } [0; 0.3].$$

Решение.

$$q = 0.1(x + y).$$

Первое сближение (по формулам (18)).

$$q_0 = 0.1,$$

$$\Delta y_0^{(1)} = 0.1, \quad y_1^{(1)} = 1 + 0.1 = 1.1,$$

$$q_1^{(1)} = 0.1(0.1 + 1.1) = 0.12, \quad \Delta q_0^{(1)} = 0.12 - 0.1 = 0.02.$$

Второе сближение (по формулам (19)).

$$\Delta y_0^{(2)} = 0.1 + \frac{1}{2} \cdot 0.02 = 0.11, \quad \Delta y_1^{(2)} = 0.12 + \frac{1}{2} \cdot 0.02 = 0.13,$$

$$y_1^{(2)} = 1 + 0.11 = 1.11, \quad y_2^{(2)} = 1.11 + 0.13 = 1.24,$$

$$q_1^{(2)} = 0.1(0.1 + 1.11) = 0.121, \quad q_2^{(2)} = 0.1(0.2 + 1.24) = 0.144,$$

$$\Delta q_1^{(2)} = 0.144 - 0.121 = 0.023, \quad \Delta q_0^{(2)} = 0.121 - 0.1 = 0.021,$$

$$\Delta^2 q_0^{(2)} = 0.002.$$

Третье сближение. Вычисляем

$$\Delta y_0^{(3)} = 0.1 + \frac{1}{2} \cdot 0.021 - \frac{1}{12} \cdot 0.002 = 0.1103,$$

$$\Delta y_1^{(3)} = 0.121 + \frac{1}{2} \cdot 0.021 + \frac{5}{12} \cdot 0.002 = 0.132,$$

$$\Delta y_2^{(3)} = 0.144 + \frac{1}{2} \cdot 0.023 + \frac{5}{12} \cdot 0.002 = 0.1563.$$

$$y_1^{(3)} = 1.1103, \quad y_2^{(3)} = 1.2426, \quad y_3^{(3)} = 1.3989.$$

## 6.4. Метод Адамса

Методы Рунге-Кутты и Крылова требуют многократного вычисления правых частей уравнения.

Рассмотрим метод Адамса - многоточечный метод интегрирования дифференциального уравнения.

Метод Адамса основан на аппроксимации функции  $f(x, y(x))$  (правой части уравнения  $y' = f(x, y)$ ) интерполяционным многочленом.

В предположении, что известны решения в точках  $(x_0, y_0)$ ,  $(x_1, y_1)$ ,  $(x_2, y_2)$ ,  $(x_3, y_3)$ , в результате интерполяции и разложения в ряд Тейлора получим формулу метода Адамса [76]

$$y_4 \approx y_3 + (55 f_3 - 59 f_2 + 37 f_1 - 9 f_0) \frac{h}{24}, \quad (21)$$

$f_0, f_1, f_2, f_3$  - значения  $f(x, y)$  в указанных выше точках.

Формула (21) имеет погрешность  $(251/750)h^5 f^{(5)}$ . Для уменьшения погрешности строится формула коррекции

$$y_4 = y_3 + (9f_4 + 19f_3 - 5f_2 + f_1) \frac{h}{24} \quad (22)$$

Нелинейное относительно  $y_4$  уравнение (22) можно решать, например, методом итерации.

Чаще поступают по следующей схеме [70]. Выражение (21) принимают как прогноз, затем вычисляют  $f_4 = f(x_4, y_4)$  и вычисляют  $y_4$  по формуле (22).

На каждом шаге (от  $x_i$  до от  $x_{i+1}$ ) интегрирования осуществляется одна коррекция.

Схема алгоритма приведена на рис. 6.

Заметим, что после вычисления по формуле (22) нужно осуществить сдвиг узловых значений аппроксимирующей функции для подготовки к выполнению следующего шага.

#### Пример 4.

Решить уравнение  $y' = x + y$ ,  $y(0) = 1$ ,  $h = 0.1$  на отрезке  $[0; 0.6]$ .

Решение.

В примере 3 методом Крылова найдены значения  $y_1(x_1) = 1.1103$ ,  $y_2(x_2) = 1.2426$ ,  $y_3(x_3) = 1.3989$ .

Для нахождения  $y_4$  воспользуемся формулой (21).

Получаем  $y_4^{пр} = 1.5827$ , коррекция по формуле (22) дает

$$y_4^{корр} = 1.5828.$$

Дальнейшие вычисления сведены в таблицу 4.

Таблица 4

$x$	$y^{пр}$	$y' = f(x, y)$	$y^{корр}$
0	1	1	
0.1	1.1103	1.2103	
0.2	1.2426	1.4426	
0.3	1.3989	1.6989	
0.4	1.5827	1.9827	1.5828
0.5	1.7971	2.2971	1.7974
0.6	2.0440	2.6440	2.0442

#### Структурная схема алгоритма метода Адамса

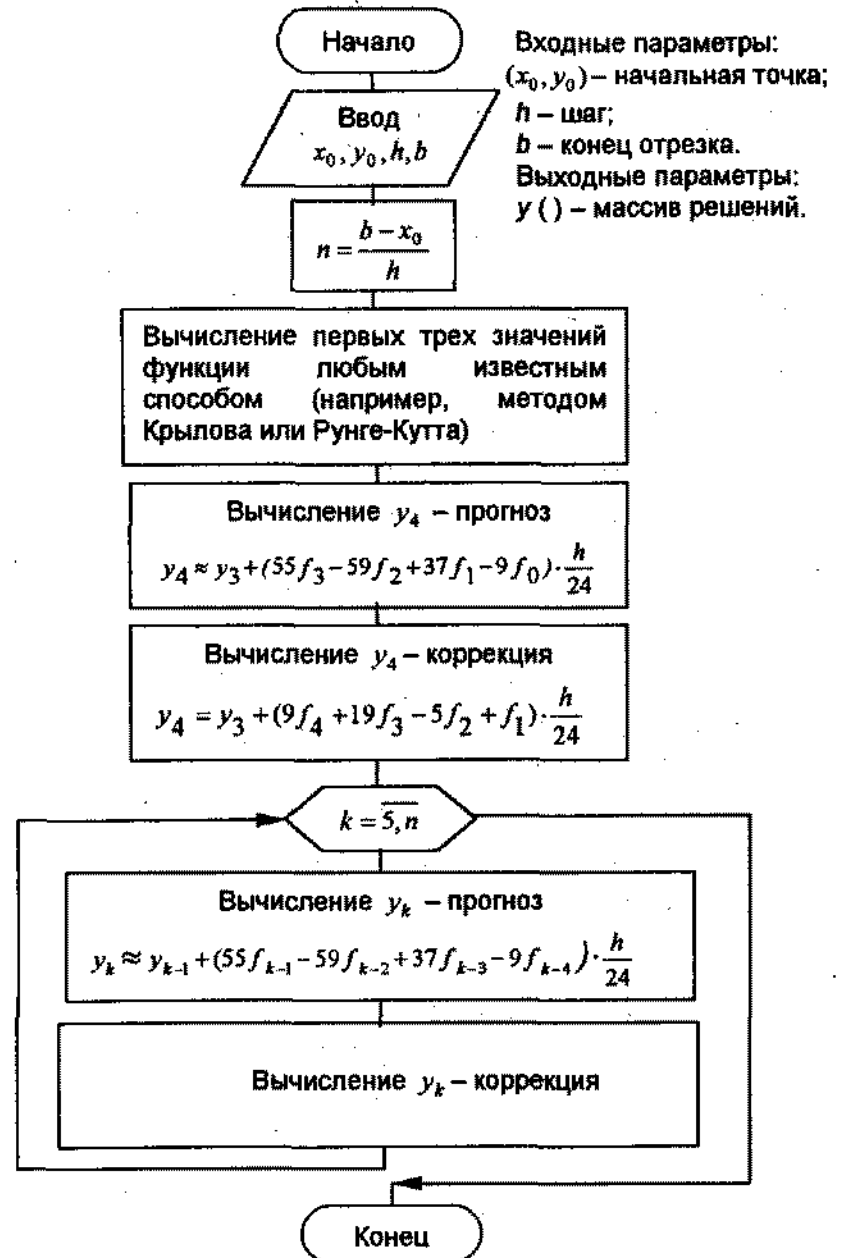


Рис. 6.6

## 6.5. Метод Гира

Для решения так называемых "жестких" уравнений, особенностью которых является медленное изменение их решений при наличии быстро затухающих возмущений, используется метод Гира. "Жесткими" уравнениями моделируются переходные процессы в нелинейных электронных схемах, в задачах управления и т. д.

Метод Гира, также как и метод Адамса, применяется для продолжения решения, когда известны решения в точках  $x_1, x_2, x_3$  ( $x_0$  известно из начальных условий).

Согласно [76] решение в точке  $x_4$

$$y_4 = (3(4hf(x_4, y_4) - y_0) + 16y_1 - 36y_2 + 48y_3) / 25 \quad (23)$$

- неявное уравнение относительно  $y_4$ . Можно решить его любым методом решения нелинейных уравнений. Однако в случае "жестких" уравнений лучше всего использовать метод Ньютона.

За начальное приближение можно взять [49]

$$y_4 = 4hf(x_3, y_3) + (y_0 - 10y_3) / 3 - 2y_1 + 6y_2. \quad (24)$$

### Алгоритм метода Гира

1. По известным  $x_{i-1}$  и  $y_{i-4}, y_{i-3}, y_{i-2}, y_{i-1}$  вычислить  $f(x_{i-1}, y_{i-1})$ .

2. По формуле (24) вычислить прогнозируемое значение  $y_i^{пр} = 4hf(x_{i-1}, y_{i-1}) + (y_{i-4} - 10y_{i-3}) / 3 - 2y_{i-2} + 6y_{i-1}$ .

3. Определить  $f(x, y_i^{пр})$ .

4. Скорректировать значение  $y_i$  - по формуле (23)

$$y_i^{корр} = (3(4hf(x_i, y_i^{пр}) - y_{i-4}) + 16y_{i-3} - 36y_{i-2} + 48y_{i-1}) / 25;$$

5. Искомое решение при  $x = x_i$  положить равным  $y_i^{корр}$ .

П.п. 1-5 повторяют при  $i = 4, 5, 6, \dots$

Использование рассмотренного метода ускоряет на несколько порядков время интегрирования уравнения.

Для заключительной проверки правильности решения дифференциального уравнения (1) - (2) можно указать такой прием.

Имеем

$$y(x_n) = y_0 + \int_{x_0}^{x_n} f(x, y(x)) dx. \quad (25)$$

Используя найденные значения  $\tilde{y}(x_k)$  ( $k = 0, 1, \dots, n$ ) и применяя к правой части равенства (25) одну из квадратурных формул, получим

$$\tilde{y}(x_n) = y_0 + \sum_{k=0}^n A_k f(x_k, \tilde{y}(x_k)),$$

где  $A_k$  - некоторые постоянные коэффициенты.

Величину  $E_n = |\tilde{y}(x_n) - \bar{y}(x_n)|$  можно принять за абсолютную погрешность приближенного решения при  $x_0 \leq x \leq x_n$ .

## 6.6. Системы обыкновенных дифференциальных уравнений

Задача Коши для системы дифференциальных уравнений 1-го порядка

$$\begin{cases} \frac{dy_1}{dx} = F_1(x, y_1, y_2, \dots, y_i, \dots, y_N), \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ \frac{dy_i}{dx} = F_i(x, y_1, y_2, \dots, y_i, \dots, y_N), \\ \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ \frac{dy_N}{dx} = F_N(x, y_1, y_2, \dots, y_i, \dots, y_N) \end{cases} \quad (26)$$

при заданных начальных условиях

$$x = x_0, \quad y_1(x_0) = y_{10}, \quad \dots, \quad y_i(x_0) = y_{i0}, \quad \dots, \quad y_N(x_0) = y_{N0} \quad (27)$$

сводится к нахождению интегральных кривых  $y_1(x), y_2(x), \dots, y_i(x), \dots, y_N(x)$ , проходящих через точки  $(x_0, y_{20}), \dots, (x_0, y_{i0}), \dots, (x_0, y_{N0})$ , задаваемые начальными условиями.

Заметим, что задача Коши для системы уравнений имеет единственное решение, если правые части уравнений непрерывны и удовлетворяют условию Липшица по  $y_i$  ( $i = \overline{1, N}$ ) для значений  $x, y_i, y_i + \Delta y_i$ , принадлежащих некоторой области вблизи начальных условий, причем решение зависит от координат начальной точки:

$$\begin{aligned} |f_i(x, y_1 + \Delta y_1, y_2 + \Delta y_2, \dots, y_N + \Delta y_N) - f_i(x, y_1, y_2, \dots, y_N)| \leq \\ \leq L(|\Delta y_1| + |\Delta y_2| + \dots + |\Delta y_N|). \end{aligned} \quad (28)$$

Методы решения, рассмотренные выше для одного уравнения, распространяются и на системы уравнений первого порядка.

В некоторых случаях алгоритмы решения систем для удобства программной реализации несколько преобразовываются.

К примеру, при интегрировании системы ОДУ методом Рунге-Кутты 4-го порядка рекомендуется [8] следующая схема для нахождения решения в точке  $x_1 = x_0 + h$  ( $i$  - номер уравнения):

$$y_i(x_0 + h) = y_{i0} + (q_{i1} + 2q_{i2} + q_{i3} + q_{i4})/3, \quad (29)$$

$$q_{i1} = h_2 f_i(x_0, y_{i0}), \quad h_2 = h/2, \quad (30)$$

$$q_{i2} = h_2 f_i(x_0 + h_2, y_{i0} + q_{i1}),$$

$$q_{i3} = h f_i(x_0 + h_2, y_{i0} + q_{i2}),$$

$$q_{i4} = h_2 f_i(x_0 + h, y_{i0} + q_{i3}).$$

**Укрупненная схема алгоритма интегрирования системы дифференциальных уравнений**

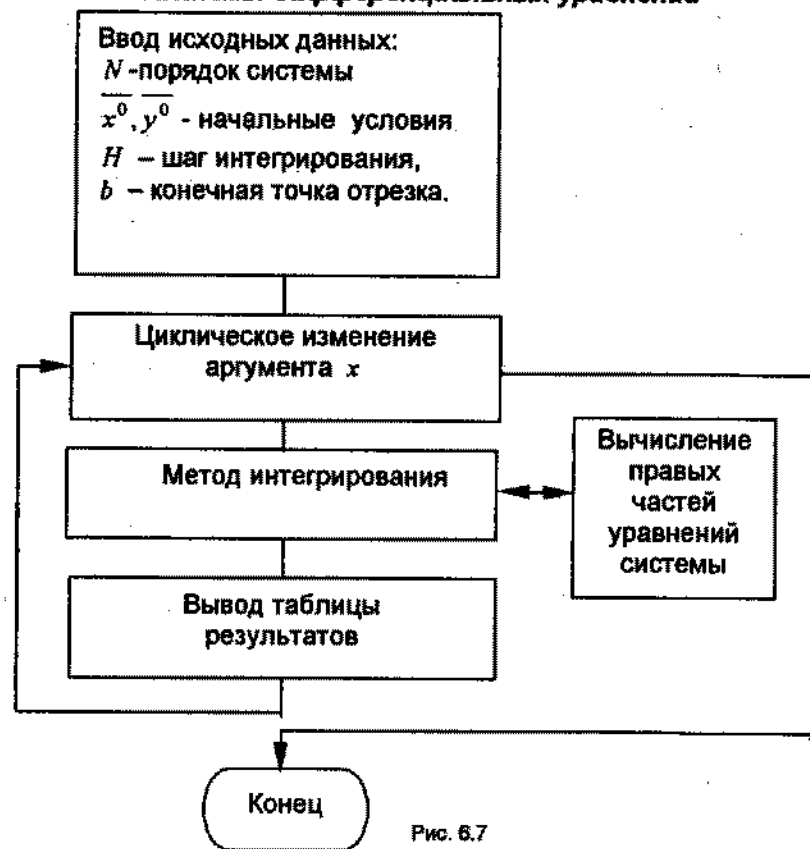


Рис. 6.7

Приведем расчетные формулы [7] и схему алгоритма метода Рунге - Кутта - Мерсона с автоматическим выбором шага, обеспечивающего приближенную оценку погрешности на каждом шаге интегрирования. Погрешность интегрирования  $R \sim h^5$ .

**Алгоритм метода Рунге - Кутта - Мерсона**

1. Задать число уравнений системы  $N$ , погрешность  $E$ , начальный шаг  $h = H$  и начальные условия  $x = x_0, y_{10}, \dots, y_{i0}, \dots, y_{N0}$

2. С помощью пяти циклов с переменной  $j = \overline{1, N}$  вычислить коэффициенты:

$$k_{0j} = h_3 f_j(x_i, y_{ij}), \quad h_3 = \frac{h}{3},$$

$$k_{1j} = h f_j(x_i + h_3, y_{ij} + k_{0j}),$$

$$k_{2j} = h f_j(x_i + h_3, y_{ij} + 0.5(k_{0j} + k_{1j})), \quad (31)$$

$$k_{3j} = k_{0j} + 4h_3 f_j(x_i + \frac{1}{2}h, y_{ij} + 0.375(k_{0j} + k_{2j})),$$

$$k_{4j} = h_3 f_j(x_i + h, y_{ij} + 1.5(k_{3j} - k_{2j})).$$

Вычислить (в последнем цикле) значение

$$y_{j(i+1)} = y_{ij} + (k_{0j} + 4k_{3j} + k_{4j})/6 \quad (32)$$

и погрешность

$$R_{j(i+1)} = (-2k_{0j} + 9k_{2j} - 8k_{3j} + k_{4j})/30. \quad (33)$$

3. Проверить условия

$$|R_{j(i+1)}| \leq E, \quad |R_{j(i+1)}| \geq \frac{E}{30}. \quad (34)$$

Если оба условия выполнены, то  $y_{j(i+1)}$  - значения решения в точке  $x_{i+1} = x_i + h$ .

Если первое из условий (34) не выполняется, следует уменьшить шаг вдвое. Повторить вычисления с п. 2, восстановив начальные значения  $y_{j,i}$ .

Если второе из условий (34) не выполняется, то шаг увеличить вдвое и повторить вычисления с п. 2.

Вычисления повторяют на всем отрезке интегрирования системы.

Часто систему (26-27) записывают в векторном виде

$$\bar{Y}'(x) = \bar{F}(x, \bar{Y}(x)), \quad \bar{Y}'(x) = \{y'_1, y'_2, \dots, y'_N\},$$

$$\bar{Y}(x) = \{y_1, y_2, \dots, y_N\},$$

$$\bar{F}(x, \bar{Y}(x)) = \{F_1(x, y), F_2(x, y), \dots, F_N(x, y)\}$$

**Структурная схема метода Рунге-Кутты  
для решения системы**

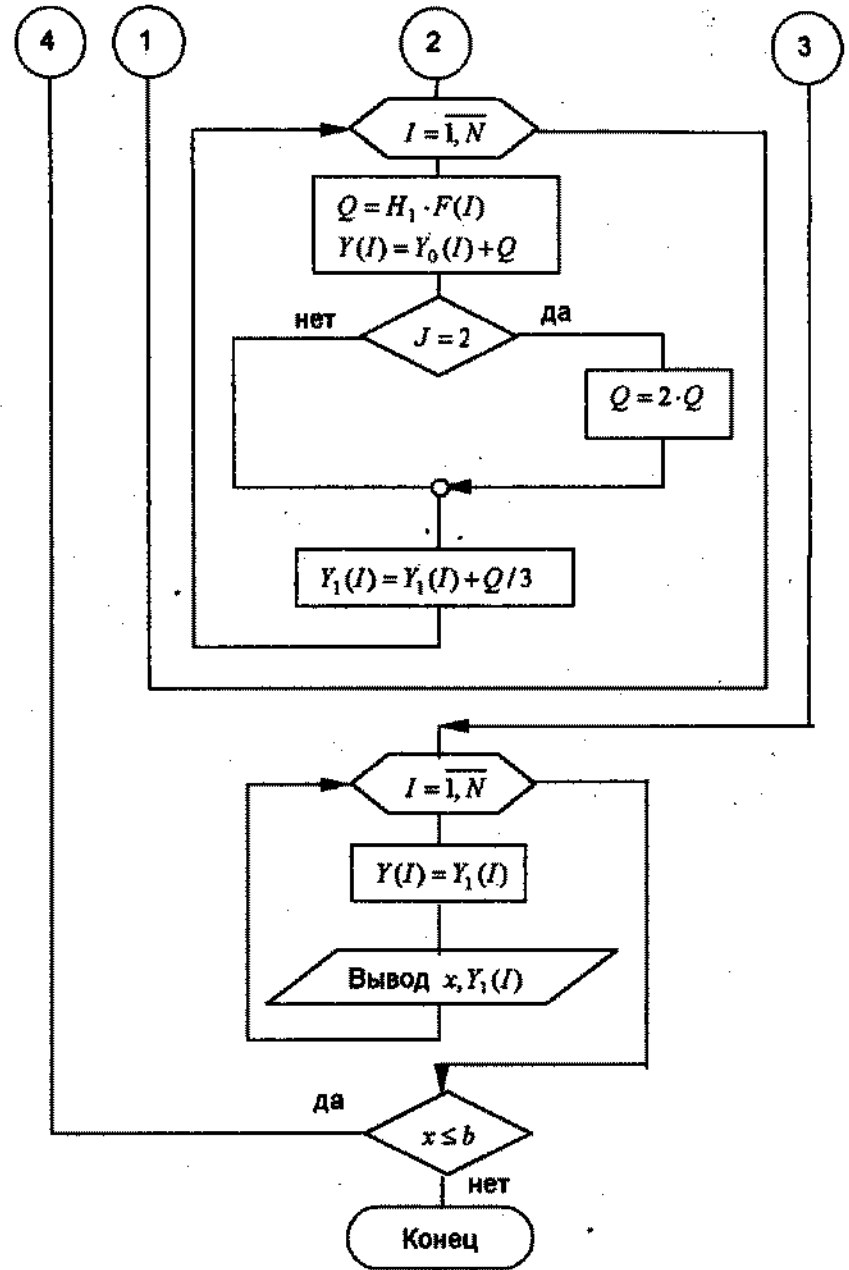
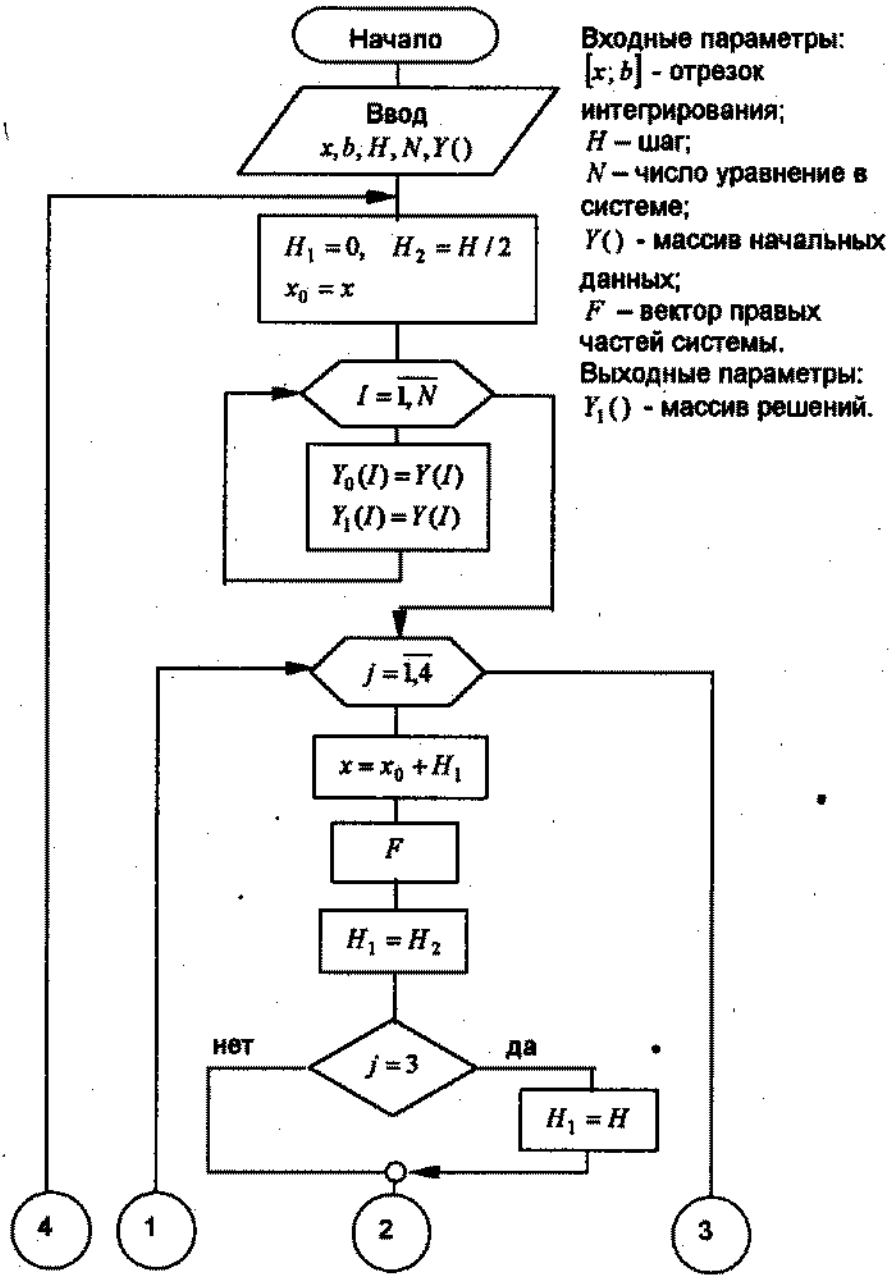


Рис. 6.8

**Пример 5.**

Методом Рунге - Кутты - Мерсона решить систему

$$\begin{cases} F_1 = \frac{dy_1}{dx} = y_1 + y_2 - x^2 + x - 2, \\ F_2 = \frac{dy_2}{dx} = -2y_1 + 4y_2 + 2x^2 - 4x - 7 \end{cases}$$

при  $x_0 = 0, y_{10} = 0, y_{20} = 2, h = 0.02, E = 0.001$ .**Решение.**

Таблица 5

x	y <sub>1</sub>	y <sub>2</sub>
0.02	6.7313*10 <sup>-4</sup>	2.02096
0.04	1.8988*10 <sup>-3</sup>	2.04103
0.08	6.7561*10 <sup>-3</sup>	2.08118
0.16	2.6099*10 <sup>-2</sup>	2.1615
0.32	1.0333*10 <sup>-1</sup>	2.3226
0.64	4.1272*10 <sup>-1</sup>	2.7477
0.96	9.3094*10 <sup>-1</sup>	2.9816
1.28	1.6650	3.3389

Из таблицы видно, что шаг автоматически увеличивается от  $h = 0.02$  до  $h = 0.32$ .

При увеличении числа уравнений в системе метод Рунге-Кутты нуждается в модификации.

Рассмотрим модификацию Гилла [61], которая дает максимально достижимую точность в смысле влияния ошибок округления, требует минимального объема памяти и реализуется достаточно просто. Для системы  $(n+1)$  уравнения первого порядка, записанной в виде

$$y'_i = f_i(x, y_0(x), y_1(x), \dots, y_n(x)), \quad i = \overline{0, n}, \quad (35)$$

значения  $y$  хранятся в массиве  $y_{i,j}$ , в котором начальные условия записываются так:

$$y_{i0} = y_i(x_0), \quad i = \overline{0, n} \quad (36)$$

Процедура Гилла состоит в выполнении, в соответствии с алгоритмом, двух шагов Гилла размером  $\frac{h}{2}$ , необходимых для получения  $y_i(x_0 + h) = y_i^{(1)}$ , а затем одного шага размером  $h$  для вычисления  $y_i(x_0 + h) = y_i^{(2)}$ . ( $x_1 = x_0 + \frac{h}{2}, x_2 = x_1 + \frac{h}{2}, x_3 = x_0 + h$ .)

**Алгоритм метода Гилла**

1. Задать специальный набор коэффициентов.

$$a_1 = 1/2, \quad b_1 = 2, \quad c_1 = 1/2,$$

$$a_2 = 1 - \sqrt{1/2}, \quad b_2 = 1, \quad c_2 = 1 - \sqrt{1/2},$$

$$a_3 = 1 + \sqrt{1/2}, \quad b_3 = 2, \quad c_3 = 1 + \sqrt{1/2}.$$

2. Массив  $q_{i,j}$  используют со следующими начальнымиданными  $q_{i,j}(x_0) = 0, \quad i = \overline{0, n}$ .3. Полагают  $j = 1$ .4. Для  $i = \overline{0, n}$  вычисляют

$$k_{i,j} = f_i(x_{j-1}, y_{0,j-1}, \dots, y_{n,j-1}),$$

$$y_{i,j} = y_{i,j-1} + (a_j(h \cdot k_{i,j} - b_j \cdot q_{i,j-1})),$$

$$q_{i,j} = q_{i,j-1} + 3(a_j(h \cdot k_{i,j} - b_j \cdot q_{i,j-1})) - c_j \cdot h \cdot k_{i,j}.$$

5. Повторяют п.4 для  $i = 2, 3, 4$ .6. После четырех шагов значения  $y_i(x_0 + h) = y_{i4}$ .

Новые начальные значения  $y_{i0}$  и  $q_{i0}$  определяются следующим образом:

$$y_{i0} = y_{i4} \text{ и } q_{i0} = q_{i4}, \quad i = \overline{0, n},$$

затем повторяют п.п. (3) - (6) для вычисления следующих значений  $y_i$ .

Если разность двух значений (локальная ошибка округления) меньше некоторого наперед заданного  $E$ , результат считается достаточно точным. Если разность больше  $E$ , точность может быть повышена уменьшением шага вдвое и повторением вычислений.

Для этого метода ошибка округления

$$E_1 = \frac{1}{15} \sum_{i=0}^n \frac{|y_i^{(1)} - y_i^{(2)}|}{n+1}. \quad (37)$$

Если  $E_1 > E$ , то полагают  $h = \frac{h}{2}$  и вычисления повторяют до выполнения условия  $E_1 \leq E$ . Метод Гилла, если ошибка  $E_1$  мала, часто дополняют процедурой удвоения шага. Если  $E_1 < E/50$ , то  $h$  полагают равным  $2h$ . Такая процедура ускоряет процесс получения решения.

## 6.7. Уравнения высших порядков

Задача Коши для обыкновенного дифференциального уравнения  $n$ -го порядка

$$y^{(n)} = f(x, y, y', y'', \dots, y^{(n-1)}) \quad (38)$$

с начальными условиями

$$y(x_0) = y_0, \quad y'(x_0) = y'_0, \dots, y^{(n-1)}(x_0) = y_0^{(n-1)}$$

сводится к задаче Коши для системы  $n$  дифференциальных уравнений первого порядка.

$$\text{Обозначим } y' = y_1, \quad y'' = y_2, \dots, y^{(n-1)} = y_{n-1}. \quad (39)$$

Задача Коши (38) равносильна задаче Коши для системы

$$\begin{cases} y' = y_1, \\ y_1' = y_2, \\ \dots \dots \dots \\ y_{n-2}' = y_{n-1}, \\ y_{n-1}' = f(x, y, y_1, \dots, y_{n-1}) \end{cases} \quad (40)$$

с начальными условиями

$$y(x_0) = y_0, \quad y_1(x_0) = y'_0, \dots, y_{n-1}(x_0) = y_0^{(n-1)}.$$

## 6.8. Дифференциальные уравнения второго порядка

Ввиду особого значения и широкого применения дифференциальных уравнений второго порядка приведем схему решения, имеющую погрешность  $R \approx h^5$ , методом Рунге-Кутты [77] уравнения вида  $y'' = f(x, y, y')$ ,  $y(x_0) = y_0$ ,  $y'(x_0) = y_1$ .

Вычисляют

$$\begin{aligned} k_0 &= hf(x_i, y_i, y'_i), \\ k_1 &= hf(x_i + \frac{h}{2}, y_i + \frac{h}{2}y'_i + \frac{h}{8}k_0, y'_i + \frac{k_0}{2}), \\ k_2 &= hf(x_i + \frac{h}{2}, y_i + \frac{h}{2}y'_i + \frac{h}{8}k_0, y'_i + \frac{k_1}{2}), \\ k_3 &= hf(x_i + h, y_i + \frac{h}{2}y'_i + \frac{h}{8}k_2, y'_i + \frac{k_2}{2}), \end{aligned} \quad (41)$$

затем находят решение в точке  $x_{i+1}$

$$y_{i+1} = y_i + h(y'_i + \frac{1}{6}(k_0 + k_1 + k_2)),$$

$$y'_{i+1} = y'_i + \frac{1}{6}(k_0 + 2k_1 + 2k_2 + k_3).$$

Перед началом вычислений нужно задать шаг  $h$  и начальные значения  $x_0$ ,  $y(x_0)$ ,  $y'(x_0) = y_0$ .

## 6.9. Краевые (граничные) задачи

Пусть дано обыкновенное дифференциальное уравнение порядка  $N$  или эквивалентная ему система из  $N$  уравнений первого порядка. Необходимо найти решение уравнения  $y(x)$  на отрезке  $[x_0; b_0]$ , удовлетворяющее граничным условиям в виде функциональной связи между искомой функцией и ее производными

$$\varphi_k(y(x_k), y'(x_k), \dots, y^{(N-1)}(x_k)) = 0 \quad (42)$$

Для существования и единственности решения граничной задачи необходимо, чтобы  $k$  - число условий - совпадало с порядком дифференциального уравнения  $N$  ( $N \geq 2$ ). Условия (42) могут быть заданы в любой точке  $[x_0; b]$ , в том числе и на границах.

Рассмотрим для примера линейное дифференциальное уравнение второго порядка

$$y'' + p(x)y' + q(x)y = f(x) \quad (43)$$

с граничными условиями

$$y(a) = A, \quad y(b) = B. \quad (44)$$

Требуется найти решение уравнения на отрезке  $[a; b]$ , удовлетворяющее условиям (43) и (44).

Более общим видом условий (44) являются

$$\begin{cases} \alpha_0 y(a) + \alpha_1 y'(a) = A, \\ \beta_0 y(b) + \beta_1 y'(b) = B. \end{cases} \quad (45)$$

$$|\alpha_0| + |\alpha_1| \neq 0,$$

$$|\beta_0| + |\beta_1| \neq 0.$$

Разбивая отрезок  $[a; b]$  на  $n$  равных частей с шагом  $h = \frac{(b-a)}{n}$  получают точки  $x_i = a + ih$ , ( $i = \overline{0, n}$ ), в которых требуется найти значения решения.

**Структурная схема метода Рунге - Кутты решения дифференциального уравнения второго порядка**

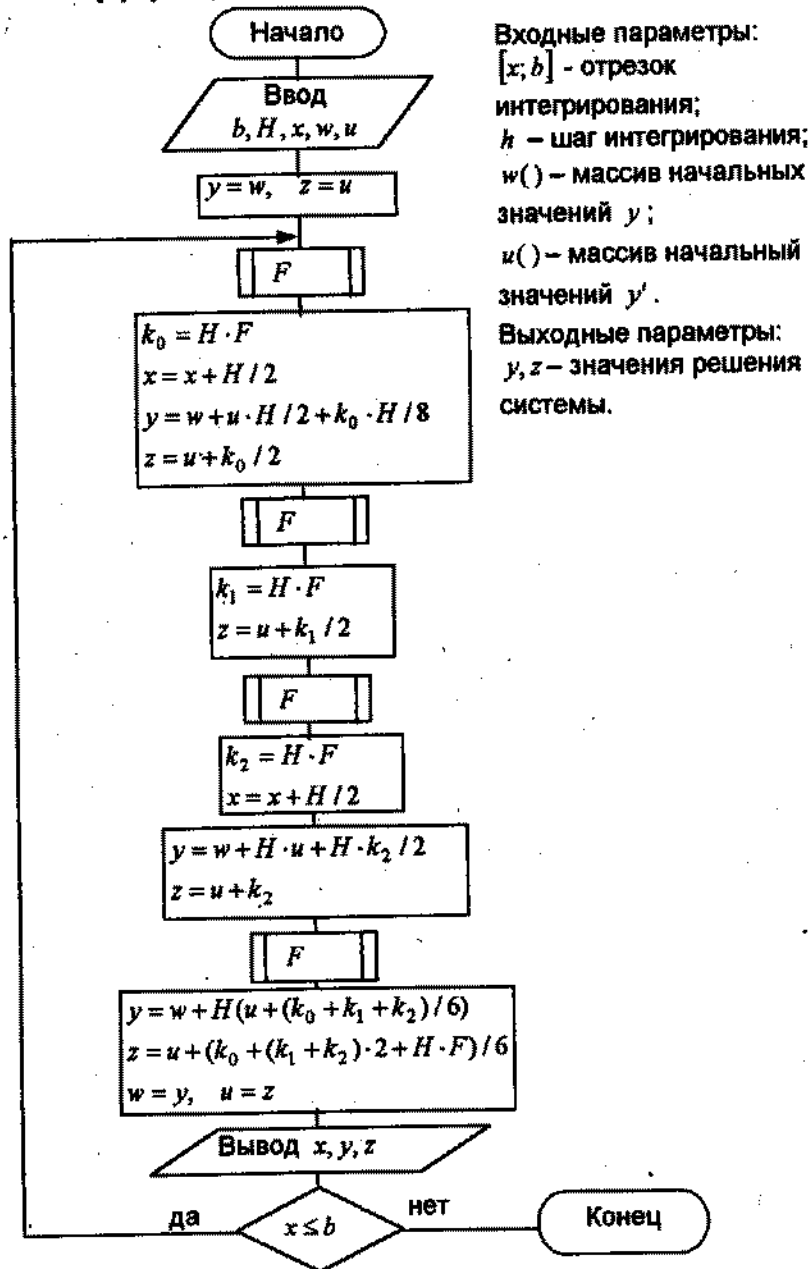


Рис. 6.9

Входные параметры:  
 $[x, b]$  - отрезок интегрирования;  
 $h$  - шаг интегрирования;  
 $w()$  - массив начальных значений  $y$ ;  
 $u()$  - массив начальных значений  $y'$ .  
 Выходные параметры:  
 $y, z$  - значения решения системы.

**6.10. Численные методы решения краевой (граничной) задачи**

В общем случае краевая задача может

- не иметь решений,
- иметь единственное решение,
- иметь несколько и даже бесконечно много решений.

Например, краевая задача  $y'' + y = 0, y(0) = y(\pi) = 0$

имеет бесчисленное множество решений:  $y = c \sin x$ ,  $c$  - произвольная постоянная.

Краевая задача  $y'' + y = 0, y(0) = 0, y(b) = 1$  имеет единственное решение

$$y = \frac{\sin x}{\sin b}, \text{ при } 0 < b < \pi,$$

при  $b = \pi$  не имеет решения.

Рассмотрим две группы методов решения граничной задачи:

- сведение решения краевой задачи к последовательности решений задачи Коши;
- конечно-разностные методы.

**6.10.1. Метод стрельбы**

Рассмотрим краевую задачу в виде

$$y'' = f(x, y, y'). \quad (46)$$

Будем искать решение этого уравнения на отрезке  $[0; 1]$ , удовлетворяющее граничным условиям

$$y(0) = y_0, \quad y(1) = y_1. \quad (47)$$

Отрезок интегрирования дифференциального уравнения  $[a, b]$

заменой  $t = \frac{x-a}{b-a}$  переводится в отрезок  $[0; 1]$ .

Суть метода стрельбы заключается в сведении краевой задачи (46) - (47) к решению задачи Коши для того же уравнения (46) с начальными условиями

$$y(0) = y_0, \quad y'(0) = k = \text{tg} \alpha. \quad (48)$$

Геометрически эта задача - поиск интегральной кривой, исходящей из точки  $(0, y_0)$  и попадающей в точку  $(1, y_1)$ .



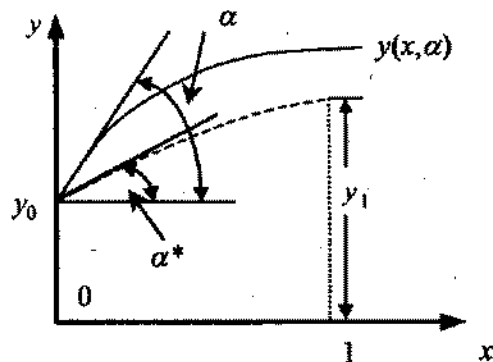


Рис. 6.10

Решение задачи Коши (46), (48)  $y = y(x, \alpha)$  зависит от параметра  $\alpha$ .

Если  $\alpha = \alpha^*$ , то решения задачи Коши  $y = y(x, \alpha^*)$  и краевой задачи с начальными условиями (47) совпадут. Возникает дополнительное условие

$$y(1, \alpha) - y_1 = 0. \quad (49)$$

Обозначим

$$F(\alpha) = y(1, \alpha) - y_1. \quad (50)$$

Уравнение  $F(\alpha) = 0$  можно решать, например, методом половинного деления, методом Ньютона и т.д.

В любом случае необходим отрезок  $[\alpha_0; \alpha_1]$ , содержащий  $\alpha^*$ . Отрезок  $[\alpha_0; \alpha_1]$  может быть найден из решения задачи Коши (46), (48) так, чтобы  $F(\alpha_0) F(\alpha_1) < 0$ . Далее решение ищется в соответствии с алгоритмом метода половинного деления для решения уравнения.

Решается задача Коши при  $\alpha = \alpha_2$ , где  $\alpha_2 = \frac{\alpha_0 + \alpha_1}{2}$ . Затем ищется очередное значение  $\alpha$  до тех пор, пока не выполнится условие  $|\alpha_i - \alpha_{i-1}| < \epsilon$ . Заметим, что в ряде случаев возможна неустойчивость решения.

Разложим решение по формуле Тейлора, сохраняя лишь линейные относительно  $\Delta\alpha$  члены разложения

$$y(1, \alpha_0 + \Delta\alpha) = y(1, \alpha_0) + \frac{\partial y(1, \alpha_0)}{\partial \alpha} \Delta\alpha, \quad (51)$$

откуда, при условии, что

$$\frac{\partial y}{\partial \alpha}(1, \alpha_0) = \frac{y(1, \alpha_0 + \delta\alpha) - y(1, \alpha_0)}{\delta\alpha},$$

$\delta\alpha$  - произвольное малое значение,

$$\Delta\alpha = \frac{y_1 - y(1, \alpha_0)}{\partial y / \partial \alpha}. \quad (52)$$

**Алгоритм решения краевой задачи методом стрельбы**

1. Задать  $y_0, y_1, \alpha_0, h, \delta\alpha, E$ .
2. Решить задачу Коши при  $\alpha = \alpha_0$ .
3. Решить задачу Коши при  $\alpha = \alpha_0 + \delta\alpha$ .
4. Вычислить  $\Delta\alpha$ .
5. Если  $|\Delta\alpha| < E$ , задача решена.

В противном случае, полагая  $\alpha = \alpha_0 + \Delta\alpha$ , возврат к п.2.

Метод стрельбы может применяться для решения системы дифференциальных уравнений и уравнений высших порядков.

Пусть дана система двух уравнений первого порядка

$$\begin{cases} y'_1 = f_1(x, y_1, y_2), \\ y'_2 = f_2(x, y_1, y_2) \end{cases} \quad (53)$$

и граничные условия

$$y_1(0) = y_{10}, \quad y_2(1) = y_{20}. \quad (54)$$

Выбирают некоторое  $\alpha$ , аппроксимирующее значение  $y_2(0) = \alpha$ . Получают решение при  $x = 1$  задачи Коши с начальными условиями  $y_1(0) = y_{10}, \quad y_2(0) = \alpha$ .

Если

$$|y_2(1, \alpha) - y_{20}| < E \quad (55)$$

то решение считается найденным, в противном случае находится уточненное значение  $\alpha$  и процесс повторяется.

## 6.10.2. Метод конечных разностей

Достоинство метода конечных разностей в том, что он сводит решение краевой задачи к решению системы алгебраических уравнений относительно значений искомой функции (решения) на заданном множестве точек. Это достигается заменой производных, входящих в дифференциальное уравнение, их конечными разностями.

**Структурная схема алгоритма метода стрельбы для дифференциального уравнения второго порядка**

Входные параметры:

$(x_0; y_0)$  - начальные условия,

$(l; y_l)$  - граничные условия,

$h_0, \alpha_0, \delta\alpha$  - параметры,

$\epsilon$  - точность.

Выходные параметры:

$(x_1; y_1)$  - решение.

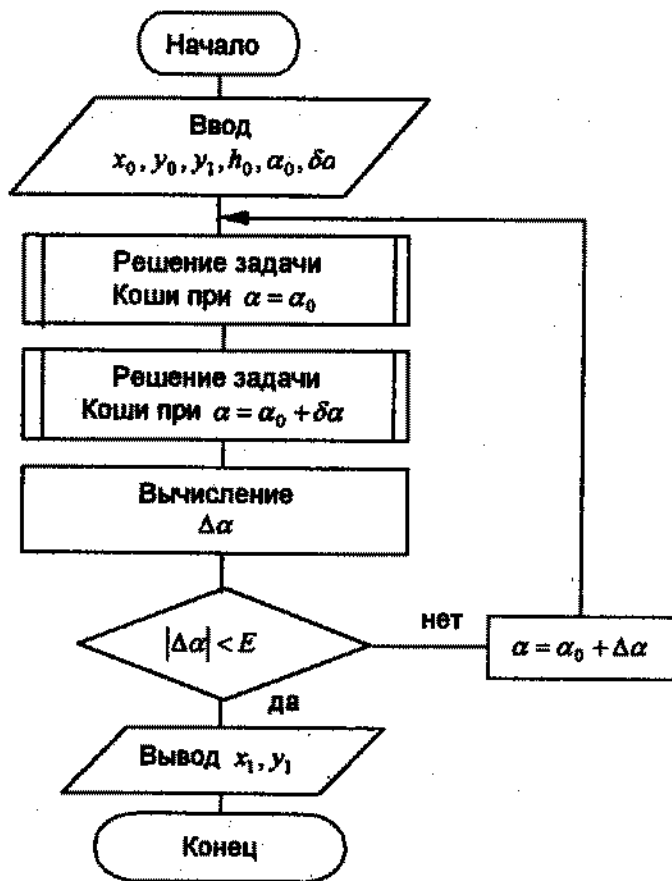


Рис. 6.11

Пусть дано уравнение

$y'' + p(x)y' + q(x)y = f(x)$  и граничные условия (45).

Заменяем  $y'', y', y'(a), y'(b)$  их значениями, получаем систему  $(n+1)$  уравнения с  $(n+1)$  неизвестными

$$\begin{cases} \frac{y_{i+2} - 2y_{i+1} + y_i}{h^2} + p_i \frac{y_{i+1} - y_i}{h} + q_i y_i = f_i, & i = \overline{0, n-2}, \\ \alpha_0 y_0 + \alpha_1 \frac{y_1 - y_0}{h} = A, \\ \beta_0 y_n + \beta_1 \frac{y_n - y_{n-1}}{h} = B, \end{cases} \quad (56)$$

где введены обозначения

$$p_i = p(x_i), \quad q_i = q(x_i), \quad f_i = f(x_i),$$

$$y_i'' = \frac{y_{i+2} - 2y_{i+1} + y_i}{h^2}, \quad y_i' = \frac{y_1 - y_0}{h}, \quad y_n' = \frac{y_n - y_{n-1}}{h},$$

$$x_i = a + ih, \quad (i = \overline{1, n-1}), \quad h = \frac{b-a}{n}, \quad n = \text{const.}$$

Решив систему (56) найдем таблицу приближенных значений искомого решения  $y(x)$  на сетке с шагом  $h$ .

Оценка погрешности метода конечных разностей имеет вид:

$$|y_i - y(x_i)| \leq \frac{h^2 M_4}{96} (b-a), \quad M_4 = \max |y^{IV}(x)|.$$

Условия разрешимости такой системы изложены в [8], систему (56) можно решать методом прогонки. В данном случае поступим так.

Первые  $(n-1)$  уравнения (56) перепишем в виде

$$y_{i+2} + m_i y_{i+1} + k_i y_i = h^2 f_i, \quad (57)$$

$$m_i = -2 + h p_i, \quad k_i = 1 - h p_i + h^2 q_i, \quad (i = \overline{0, n-2}).$$

Затем система приводится к виду

$$y_{i+1} = c_i (d_i - y_{i+2}), \quad (i = \overline{0, n-2}), \quad (58)$$

$$\begin{cases} c_0 = \frac{\alpha_1 - \alpha_0 h}{m_0 (\alpha_1 - \alpha_0 h) + k_0 \alpha_1}, \\ d_0 = \frac{k_0 A h}{\alpha_1 - \alpha_0 h} + f_0 h^2, \end{cases} \quad (59)$$

$$\begin{cases} c_i = \frac{1}{m_i - k_i c_{i-1}}, \\ d_i = f_i h^2 - k_i c_{i-1} d_{i-1}. \end{cases} \quad (i = \overline{1, n-2}) \quad (60)$$

**Алгоритм вычислений**

Прямой ход.

1. Вычислить  $m_i, k_i, (i=0, n-2)$  (формулы (57)).
2. Определить  $c_0, d_0$  (формулы (59)).
3. Вычислить последовательно  $c_i, d_i, i=1, 2, \dots, n-2$  (формулы (60)).

Обратный ход.

4. При  $i=n-2$ , вычислить

$$y_n = \frac{\beta_1 c_{n-2} d_{n-2} + B h}{\beta_1 (1 + c_{n-2}) + \beta_0 h} \quad ((58), (57))$$

5. При  $i=n-1, n-2, \dots, 1$  последовательно получить (58)

$$\begin{cases} y_{n-1} = c_{n-2} (d_{n-2} - y_n), \\ \dots \\ y_1 = c_0 (d_0 - y_2). \end{cases} \quad (61)$$

6. Из (57)

$$y_0 = \frac{\alpha_1 y_1 - A h}{\alpha_1 - \alpha_0 h} \quad (62)$$

**Пример 6.**

Решить краевую задачу

$$y'' - 2xy' - 2y = -4x,$$

$$y(0) - y'(0) = 0, \quad y(1) = 1 + e$$

методом конечных разностей.

**Решение.**

Пусть  $h=0.1$ . Заменяем дифференциальное уравнение разностным

$$\begin{cases} \frac{y_{i+2} - 2y_{i+1} + y_i}{h^2} - 2x_i \frac{y_{i+1} - y_i}{h} - 2y_i = -4x_i, & i = \overline{0, 8} \\ y_0 - \frac{y_1 - y_0}{0.1} = 0, \\ y_{10} = 3.718. \end{cases}$$

$$y_{i+2} + (-2 - 0.2x_i)y_{i+1} + (0.98 + 0.2x_i)y_i = -0.01x_i.$$

Согласно (56), (57)

$$\alpha_0 = 1, \quad \beta_0 = 1, \quad \alpha_1 = -1, \quad \beta_1 = 0,$$

$$A = 0, \quad B = 3.718, \quad m_i = -2 - 0.2x_i, \quad k_i = 0.98 + 0.2x_i, \quad f_i = -4x_i.$$

Прямой ход.

Вычисляем  $x_i, m_i, k_i, f_i$ .

Далее находим

$$c_0 = \frac{\alpha_1 - \alpha_0 h}{m_0 (\alpha_1 - \alpha_0 h) + k_0 \alpha_1} = -0.9016, \quad d_0 = \frac{k_0 A h}{\alpha_1 - \alpha_0 h} + f_0 h^2 = 0.$$

При  $i=1, 2, \dots, 8$  вычисляем  $c_i = -0.8941, d_i = -0.0040$  и т. д.

Обратный ход.

$$i = n - 2 = 10, \quad y_{10} = \frac{B}{\beta_0} = 3.718.$$

$i = 9, 8, \dots, 1$   $y_9 = c_8 (d_8 - y_{10}) = 3.214$ , затем вычисляем

$$y_8, y_7, \dots, y_1.$$

$$y_0 = 1.117.$$

Точное решение данной задачи  $y^T = x + e^{x^2}$ .

Результаты сведены в таблицу 6.

Таблица 6

x	m	k	f	Прямой ход		y	y <sup>T</sup>
				c	d		
0	2.00	.98		-0.9016	0.0000	1.117	1.000
.1	2.02	.00	0.4	-0.8941	-0.0040	1.229	1.110
.2	2.04	.02	0.8	-0.8865	-0.0117	1.363	1.241
.3	2.06	.04	1.2	-0.8787	-0.0228	1.521	1.394
.4	2.08	.06	1.6	-0.8706	-0.0372	1.704	1.574
.5	2.10	.08	2.0	-0.8623	-0.0550	1.916	1.784
.6	2.12	.10	2.4	-0.8536	-0.0781	2.164	2.033
.7	2.14	.12	2.8	-0.8446	-0.1007	2.455	2.332
.8	2.16	.14	3.2	-0.8354	-0.1290	2.800	2.696
.9						3.214	3.148
1						3.718	3.718

## 7. ОПТИМИЗАЦИЯ

Оптимизация - это процесс выбора наилучшего варианта из всех возможных.

Задача оптимизации - составная часть многих инженерных задач. К примеру, при анализе электронных схем ставится задача оптимизации схем с целью реализации заданных частотно-временных характеристик, минимизации чувствительности к изменению параметров.

При решении задач оптимизации рассматривается некоторая скалярная или векторная функция, называемая целевой функцией, критерием оптимальности. Аргументы (переменные целевой функции) называются параметрами оптимизации. Результат решения задачи - определение таких значений параметров, при которых целевая функция достигает экстремума.

По количеству независимых переменных различают задачи одномерные и многомерные ( $n \geq 2$ ).

По способу задания параметров задачи делятся на дискретные и непрерывные.

Задачи параметрической оптимизации предполагают поиск группы параметров, доставляющих функции экстремум. Задачи поиска функций, на которых функционалы достигают оптимальных значений, называются задачами оптимального управления.

Методы решения задачи оптимизации различаются в зависимости от того, имеет ли функция один или нескольких экстремумов.

Если имеются некоторые ограничения на искомые параметры, то задача называется задачей с ограничениями, задачей условной оптимизации. В противном случае - задачей безусловной оптимизации. При этом существенен и тот факт, являются ли эти ограничения линейными или нелинейными.

Задача нахождения максимума функции сводится к задаче поиска минимума, поэтому в дальнейшем будем говорить только о поиске минимума функции.

Целью одномерного поиска является определение минимума функции одной переменной, заданной на отрезке  $[a, b]$ , т.е. определение  $x^* \in [a, b]$  такого, что

$$f(x^*) = \min_{x \in [a, b]} f(x); \quad (1)$$

Задача (1) имеет единственное решение, если  $f(x)$  является унимодальной функцией.

Функция  $f(x)$  называется унимодальной на отрезке  $[a, b]$ , (отрезок неопределенности), если она непрерывна на  $[a, b]$  и существуют числа  $\alpha$  и  $\beta$ ,  $a \leq \alpha \leq \beta \leq b$ , такие, что:

- если  $a < \alpha$ , то на отрезке  $[a, \alpha]$   $f(x)$  монотонно убывает,
- если  $\beta < b$ , то на отрезке  $[\beta, b]$   $f(x)$  монотонно возрастает,
- при  $x \in [\alpha, \beta]$   $f^* = \min f(x)$ .

Для проверки унимодальности функции практически используются критерии:

- если  $f(x)$  дифференцируема на  $[a, b]$  и производная  $f'(x)$  не убывает на этом отрезке, то  $f(x)$  - унимодальная на  $[a, b]$ ,
- если  $f(x)$  дважды дифференцируема на отрезке  $[a, b]$  и  $f''(x) \geq 0$  при  $x \in [a, b]$ , то  $f(x)$  - унимодальная на  $[a, b]$ .

Если  $f(x)$  мультимодальная, то нахождение глобального (абсолютного) минимума обычно сопряжено с некоторыми сложностями: нужно найти все локальные минимумы, и затем определить глобальный минимум, т.к. всякая точка глобального минимума функции  $f(x)$  является и точкой локального минимума (обратное, вообще говоря, неверно).

Число  $x^* \in G$  называется точкой глобального минимума в области  $G$ , если для любого  $x \in G$   $f(x^*) \leq f(x)$ .

Точка  $\bar{x}$  называется точкой локального минимума, если существует такое  $E > 0$ , что для всех  $x \in G$ , удовлетворяющих условию

$$|x - \bar{x}| < E, \quad f(\bar{x}) \leq f(x).$$

### Пример 1.

Убедиться, что  $f(x) = \frac{1}{4}x^4 + x^2 - 8x + 12$  унимодальная на отрезке  $[0, 2]$ .

Решение.

Многочлен - непрерывная, дифференцируемая функция.

Вторая производная  $f''(x) = 3x^2 + 2 \geq 0$  при любом  $x$ , в том числе и при  $x \in [0, 2]$ , т.е.  $f(x)$  унимодальная на отрезке  $[0, 2]$ .

Существует класс задач одномерной оптимизации, имеющих следующие особенности:

- значения исследуемой функции определяются в процессе эксперимента,

- функция сложна и желательно уменьшить количество ее вычислений,
  - невозможно (или трудно) найти решение классическими методами,
  - функция не имеет непрерывных производных или разрывная.
- В этом случае прибегают к методам одномерного поиска, не требующим вычисления производных.

### 7.1. Метод прямого перебора

Если  $f(x)$  задана в  $n$  точках, то решение очевидно: находят наименьшее из  $n$  значений и полагают его равным искомому.

Пусть  $f(x)$  - унимодальная на  $[a, b]$  и требуется найти минимум  $f(x)$  с абсолютной погрешностью  $E$ .

#### Геометрическая иллюстрация метода перебора

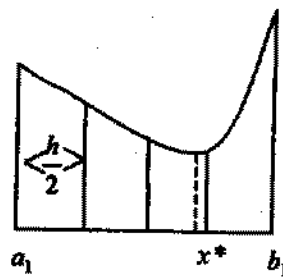
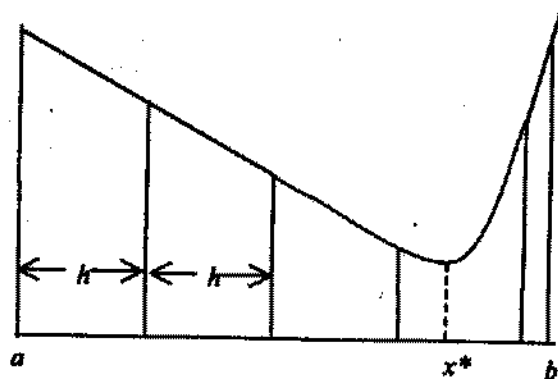
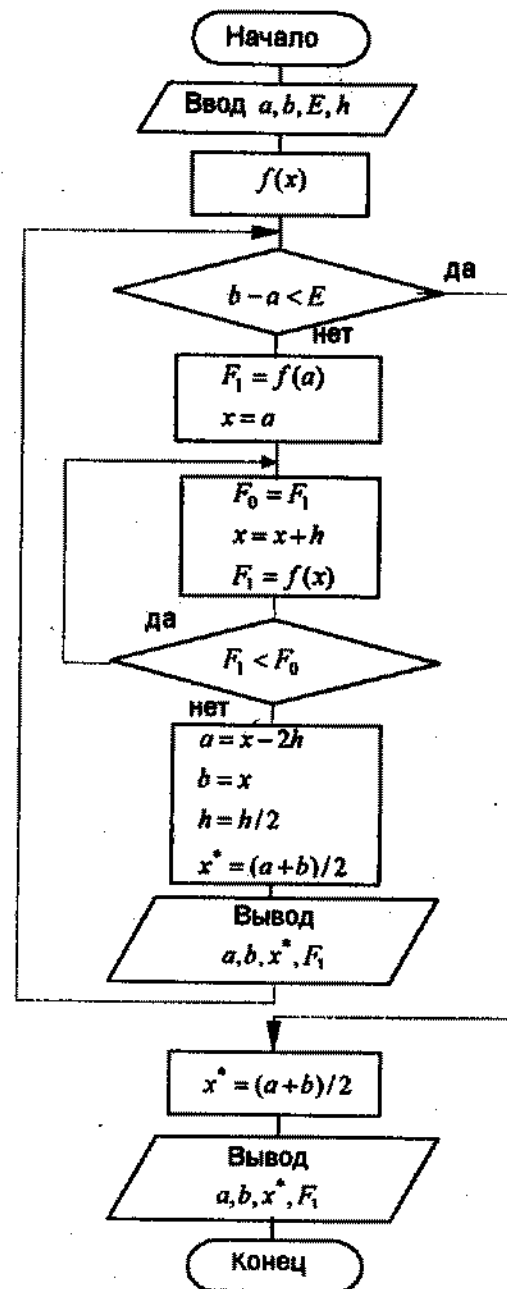


Рис. 7.1

#### Структурная схема алгоритма метода перебора



Входные параметры:  
 $[a, b]$  – отрезок неопределенности;  
 $E$  – точность;  
 $h$  – параметр.  
 Выходные параметры:  
 $(x^*, f^*(x^*))$  – точка минимума.

Рис. 7.2

Определяют значения функции:  $F_0 = f(a)$  в точке  $x = a$  и значение  $F_1 = f(x)$ , где  $x = a + h$ ,  $h$  - некоторый шаг. Увеличение значения  $x$  происходит до тех пор, пока функция  $f(x)$  убывает.

В противном случае определяют новый отрезок неопределенности  $[a; b]$ :  $a = x - 2h$ ,  $b = x$ . Полагают  $h = \frac{h}{2}$ .

Процесс повторяют до момента достижения заданной точности поиска минимума.

Схема алгоритма метода перебора - на рис. 2.

## 7.2. Методы последовательного перебора

Методы последовательного перебора заключаются в построении последовательности отрезков  $[a_1; b_1]$ ,  $[a_2; b_2]$ , ... , содержащих точку минимума.

Методы дихотомии, Фибоначчи, золотого сечения отличаются лишь способами получения последовательности вложенных отрезков  $[a_k; b_k]$ ,  $k = 1, 2, 3, \dots$

Эти методы существенно опираются на свойство унимодальности  $f(x)$  и используют предыдущую информацию о величине функции в некоторых точках для выбора следующих точек отсчета оптимизируемой функции  $f(x)$ .

### 7.2.1. Метод дихотомии

Выберем  $h \in [0; 2E]$ ,  $E$  - точность определения точки минимума функции  $f(x)$ , унимодальной на отрезке на  $[a; b]$ .

На каждой итерации сравнивают значения исследуемой функции в двух точках  $x_1^{(N-1)}$  и  $x_2^{(N-1)}$ , определяемых рекуррентными формулами:

$$\begin{aligned} x_1^{(N-1)} &= (a_{N-1} + b_{N-1} - h) / 2, \\ x_2^{(N-1)} &= (a_{N-1} + b_{N-1} + h) / 2. \end{aligned} \quad (2)$$

Если  $f(x_1^{(N-1)}) \leq f(x_2^{(N-1)})$ , то  $a_N = a_{N-1}$ ;  $b_N = x_2^{(N-1)}$ ,  
иначе  $a_N = x_1^{(N-1)}$ ;  $b_N = b_{N-1}$ .

$N = 1, 2, \dots$  - номер итерации,

$a_0 = a$ ,  $b_0 = b$ .

Вычисления проводятся до тех пор, пока

$$(b - a - h) / 2^{N+1} + h / 2 > E.$$

Тогда с абсолютной погрешностью, не превосходящей  $E$ ,

$$x^* = (a_N + b_N) / 2.$$

### Геометрическая иллюстрация метода дихотомии

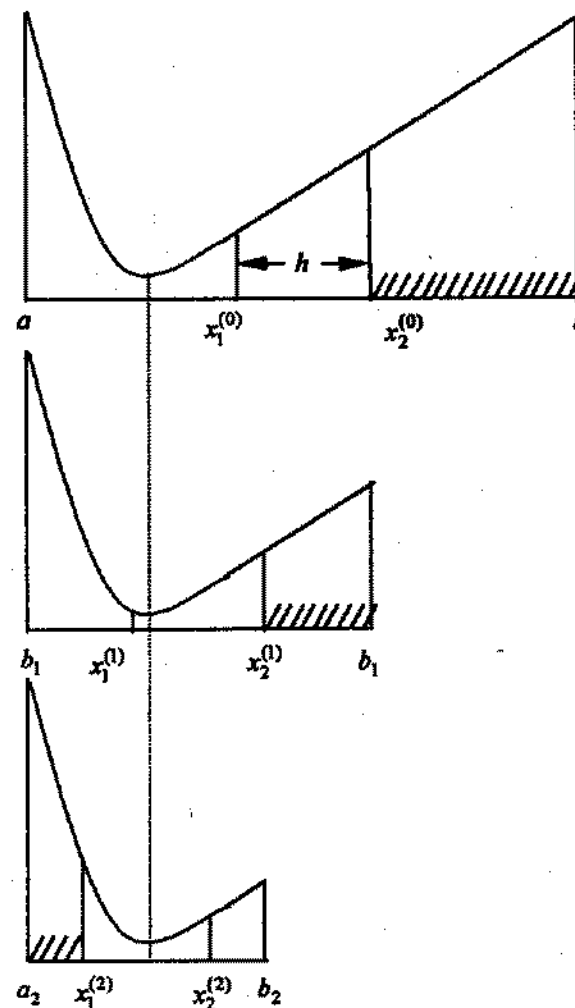


Рис. 7.3

### Структурная схема алгоритма метода дихотомии

Входные параметры:  
 $[a, b]$  – отрезок  
 неопределенности;  
 $E$  – точность;  
 $d$  – параметр.  
 Выходные параметры:  
 $(x^*, f^*(x^*))$  – точка  
 минимума.

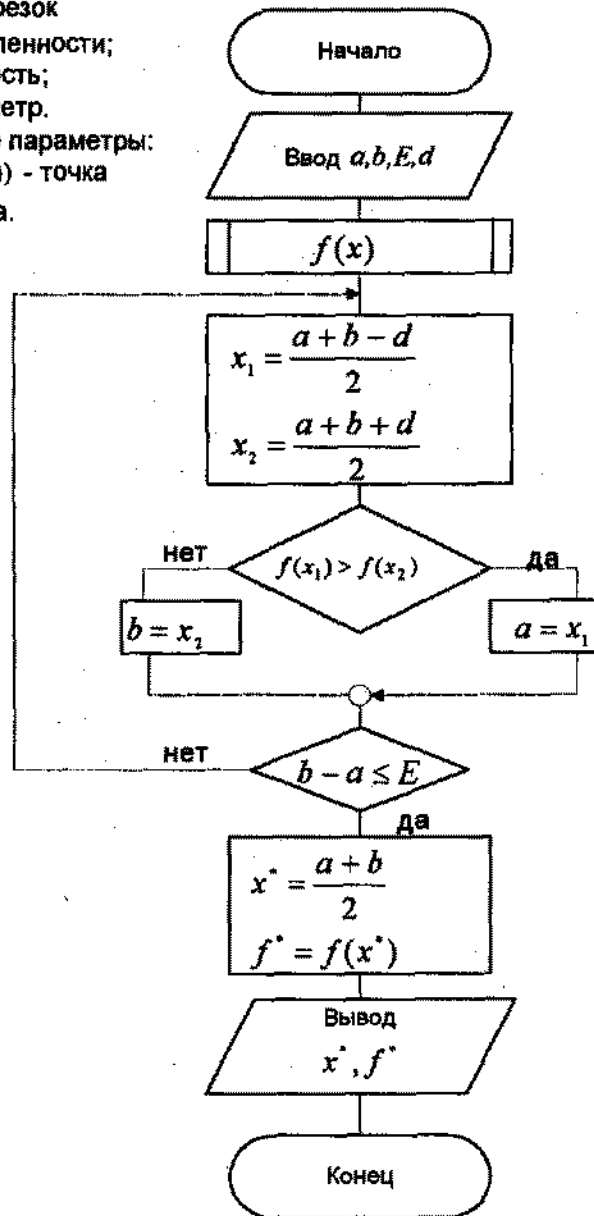


Рис. 7.4

Число отсчетов функции  $n$  (количество вычислений значений функции) и число итераций  $N$  связаны соотношением  $N = n/2$ . На каждом шаге (итерации) отрезок неопределенности  $[a_N; b_N]$  уменьшается примерно в 2 раза. Длина конечного отрезка неопределенности

$$L_N = \frac{L_0 + (2^N - 1)E}{2^N}, \quad (3)$$

$L_0 = b - a$  – начальный отрезок.

#### Пример 2.

Найти минимальное значение функции  $f(x) = 2x^2 - e^x$  на отрезке  $[0; 0.5]$  с погрешностью  $E = 0.05$  методом дихотомии.

Решение.

Функция  $f(x) = 2x^2 - e^x$  унимодальная, поскольку на отрезке  $[0; 0.5]$   $f'(x) = 4 - e^x > 0$

Положим  $\delta = 0.02 < 2E = 0.1$ . По формулам (2) вычисляем последовательно границы интервалов неопределенности. Результаты приведены в таблице 1.

Таблица 1

$N$	$a$	$x_1$	$x_2$	$b$	$b - a$	$f_1(x_1)$	$f_2(x_2)$	
0	0	0.24	0.26	0.5	0.5	-1.1580	-1.1617	$f_1 > f_2$ $a_1 = x_1^{(0)}$
1	0.24	0.36	0.38	0.5	0.26	-1.1741	-1.1735	$f_1 < f_2$ $b_2 = x_2^{(1)}$
2	0.24	0.3	0.32	0.38	0.14	-1.1699	-1.1723	$f_1 > f_2$ $a_3 = x_1^{(2)}$
3	0.3	0.33	0.35	0.38	0.08	-1.1732	-1.1741	$f_1 > f_2$ $a_4 = x_1^{(3)}$
4	0.33	0.345	0.365	0.38	0.05			

Точность достигнута, так как  $b - a = 0.05 = E$ .

Полагаем  $x^* = 0.355$ ,  $f^* = f(0.355) = -1.174$ .

## 7.2.2. Метод золотого сечения

Золотым сечением отрезка  $[a, b]$  называется деление его в отношении  $\frac{l_1}{l_2} = \frac{l_2}{l} = \tau$ , где  $l_1 + l_2 = l$ ,  $l_2 > l_1$ . (4)

Положим в выражении (4)  $l = 1$ , тогда  $l_1 = 1 - l_2$ .

Из условия (4) при  $l = 1$  следует  $l_2^2 + l_2 - 1 = 0$ , откуда  $l_2 = \frac{-1 + \sqrt{5}}{2}$ ,  $l_1 = \frac{3 - \sqrt{5}}{2}$ ,  $\tau = \frac{\sqrt{5} - 1}{2} \approx 0.618$ .

Золотое сечение осуществляется двумя точками  $x_1$  и  $x_2$ ,

$$x_1 = a + \frac{3 - \sqrt{5}}{2}(b - a) = a + k_1(b - a), \quad (5)$$

$$x_2 = a + \frac{\sqrt{5} - 1}{2}(b - a) = a + k_2(b - a),$$

причем  $x_1$  есть вторая точка золотого сечения отрезка  $[a, x_2]$ , а  $x_2$  - первая точка золотого сечения отрезка  $[x_1, b]$ ,  $\tau = 0.618$ . Таким образом, на каждой итерации ( $N = 2, 3, \dots$ ) требуется вычисление одного значения  $f(x)$  (в методе дихотомии вычисляются два значения функции).

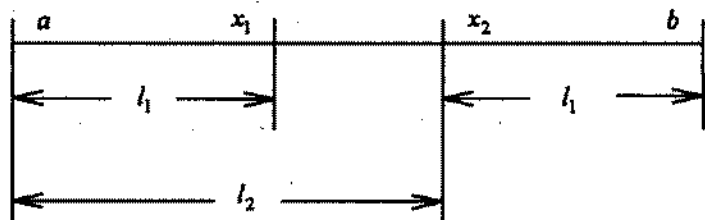


Рис. 7.5

Схема алгоритма приведена на рис. 7. Заметим, что после каждой итерации длина интервала неопределенности сокращается примерно в 1.618 раза.

Точность определения точки минимума

$$L_N = \left( \frac{\sqrt{5} - 1}{2} \right)^N L_0. \quad (6)$$

Число отсчетов  $n$  связано с числом итераций  $N$  зависимостью  $n = N + 1$ . (7)

Условие окончания поиска минимума с заданной точностью  $E$   $L_N \leq E$ . (8)

Из условий (6) и (8) можно по заданной точности  $E$  вычислить необходимое число итераций  $N$ :

$$0.618^N = \frac{L_N}{L_0},$$

$L_0, L_N$  - длины начального и конечного (после  $N$  итераций) отрезков неопределенности.

### Геометрическая иллюстрация метода золотого сечения

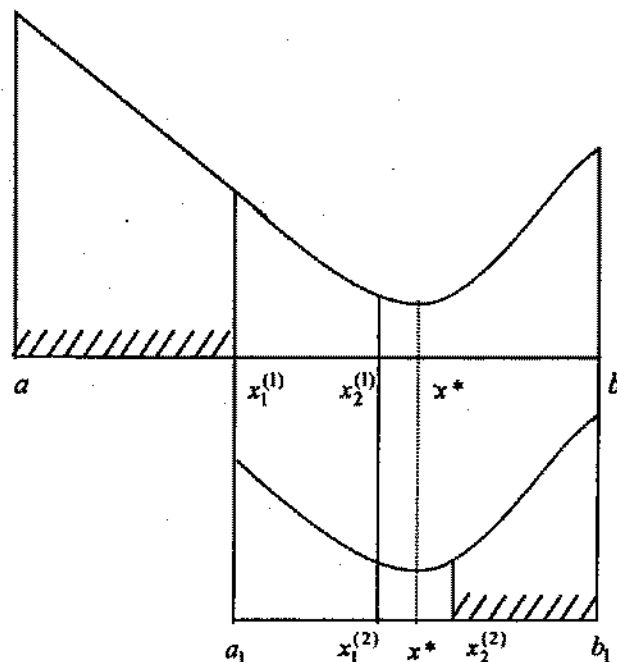


Рис. 7.6



**Структурная схема метода золотого сечения**

Входные параметры:  
 $[a, b]$  – отрезок  
 неопределенности;  
 $E$  – точность.  
 Выходные параметры:  
 $(x^*, f^*(x^*))$  – точка  
 минимума.

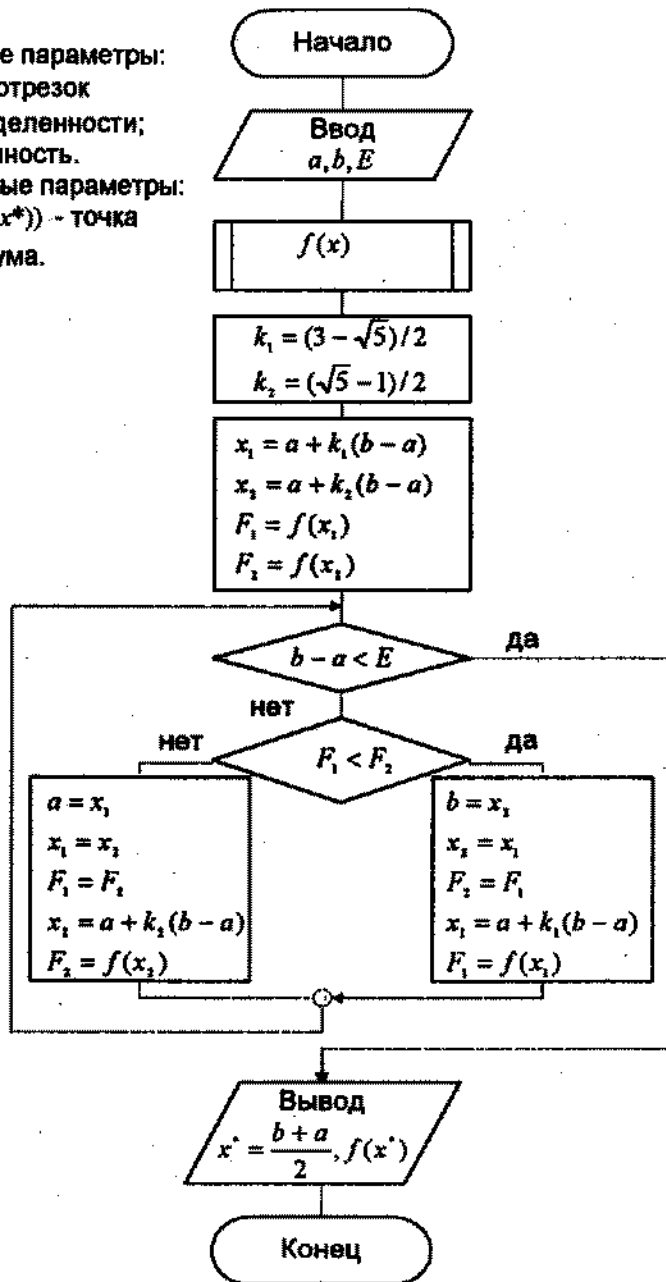


Рис. 7.7

**Пример 3.**

Требуется определить количество итераций, необходимых, чтобы локализовать точку минимума с точностью  $E = 0.01$ , если длина начального отрезка  $L_0 = 2$ .

*Решение.*

Полагаем  $L_N = 0.01$ ,  $L_N = 0.618^N L_0$  (формула (6)).

$$0.618^N \cdot 2 = 0.01, \quad N = 11.$$

**Пример 4.**

Методом золотого сечения на отрезке  $[5; 20]$  вычислить точку минимума функции  $f(x) = 24 - \frac{2}{3}x + \frac{1}{30}x^2$  с точностью  $E = 1$  и значение функции в точке минимума.

*Решение.*

Функция  $f(x)$  на отрезке  $[5; 20]$  унимодальная, т.к.

$$f''(x) = \frac{1}{15} > 0. \text{ Результаты вычислений занесены в таблицу 2.}$$

Таблица 2

N	$x_1$	a	$x_2$	b	$b-a$	$f(x_1)$	$f(x_2)$	
0	10.7	5	14.3	20	15	20.7	21.3	$f_2 > f_1$ $b = x_2$
1	8.6	5	10.7	14.3	9.3	20.73	20.68	$f_1 > f_2$ $a = x_1$
2	10.7	8.6	12.1	14.3	5.7	20.68	20.81	$f_2 > f_1$ $b = x_2$
3	9.9	8.6	10.7	12.1	3.5	20.66	20.68	$f_2 > f_1$ $b = x_2$
4	9.4	8.6	9.9	10.7	2.1	20.68	20.68	$f_1 > f_2$ $a = x_1$
5	9.9	9.4	10.2	10.7	1.3	20.67	20.67	$f_2 > f_1$ $b = x_2$
6		9.4		10.2	0.8			

При  $N = 6$   $b - a = 0.8 < E$ .  $x^* = \frac{a+b}{2} = 9.8$ ,  $f(x^*) = 20.7$ .

### 7.2.3. Оптимизация функций методом квадратичной интерполяции

Метод квадратичной интерполяции использует аппроксимацию функции  $f(x)$  в окрестности точки минимума квадратичной функцией  $\varphi(x) = Ax^2 + Bx + C$ .

Известно, что точка минимума квадратичной функции  $\varphi(x)$  имеет координаты

$$\bar{x} = -\frac{B}{(2A)}, \quad (9)$$

$$\bar{y} = f(\bar{x}).$$

Далее полагают:  $x^* \approx \bar{x}$ ,  $f(x^*) = \bar{y}$ .

Если известны значения функции  $f(x)$  в трех точках  $x_1, x_2, x_3$ , то из предположения  $f(x) \approx \varphi(x)$  на достаточно малом отрезке, содержащем точки  $x_1, x_2, x_3$ , имеем

$$\begin{cases} Ax_1^2 + Bx_1 + C = f_1, \\ Ax_2^2 + Bx_2 + C = f_2, \\ Ax_3^2 + Bx_3 + C = f_3. \end{cases}$$

$$f_1 = f(x_1), \quad f_2 = f(x_2), \quad f_3 = f(x_3).$$

Решая систему, находим  $A, B, C$ , затем из условия (9) [4]

$$XM = \frac{1}{2}(x_1 + x_2) + \frac{\frac{1}{2}(f_1 - f_2) \cdot (x_2 - x_3) \cdot (x_3 - x_1)}{(x_2 - x_3)f_1 + (x_3 - x_1)f_2 + (x_1 - x_2)f_3}, \quad (10)$$

где  $XM$  — приближенное значение  $x^*$ .

Пусть  $f(x)$  — унимодальная функция,  $t$  — приближенное значение  $x^*$  — точки минимума. Выберем  $H = \text{const}$ ,  $H \approx x^* - t$ .

#### Алгоритм метода квадратичной интерполяции

1. Вычислить  $f(t)$  и  $f(t+H)$ .
2. Если  $f(t) < f(t+H)$ , то положить  $x_1 = t - H$  и вычислить  $f(x_1)$  (рис. 8),  $x_2 = t + H$ .
3. В противном случае положить  $x_2 = t + 2H$  и вычислить  $f(x_2)$ ,  $x_1 = t$  (рис. 9).

4. Используя точки  $t, x_1, x_2$  вычислить по формуле (10)  $XM$ , затем  $f(XM)$ . Таким образом, имеем значения функции в четырех точках.

5. Если разница между наименьшим значением функции и следующим наименьшим значением функции меньше заданной точности, то процедура поиска минимума завершена. Положить  $x^* = XM$ .

6. Если процедура не завершена на шаге 4, определить границы нового интервала неопределенности так, чтобы точка с наименьшим значением функции содержалась в этом новом интервале (рис. 10). Перейти к п.3.

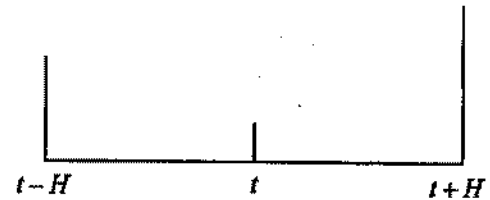


Рис. 7.8

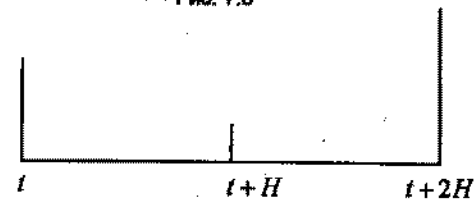


Рис. 7.9

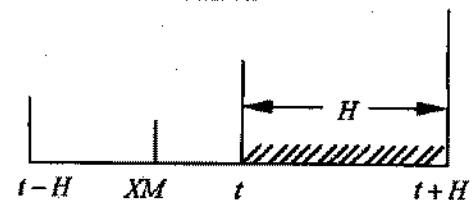


Рис. 7.10

Структурная схема алгоритма метода квадратичной интерполяции на рис. 11.

#### Пример 5. [4]

Используя квадратичную интерполяцию найти минимум функции  $f(x) = 2x^2 - e^x$  с точностью 0.001. Начальные значения  $t = 1$ ,  $H = 0.5$ .

### Структурная схема метода квадратичной интерполяции

Входные параметры:

$h$  – шаг;

$E$  – точность;

$t$  – начальное значение  $x^*$ .

Выходные параметры:

$(x^*, f^*(x^*))$  – точка

минимума.

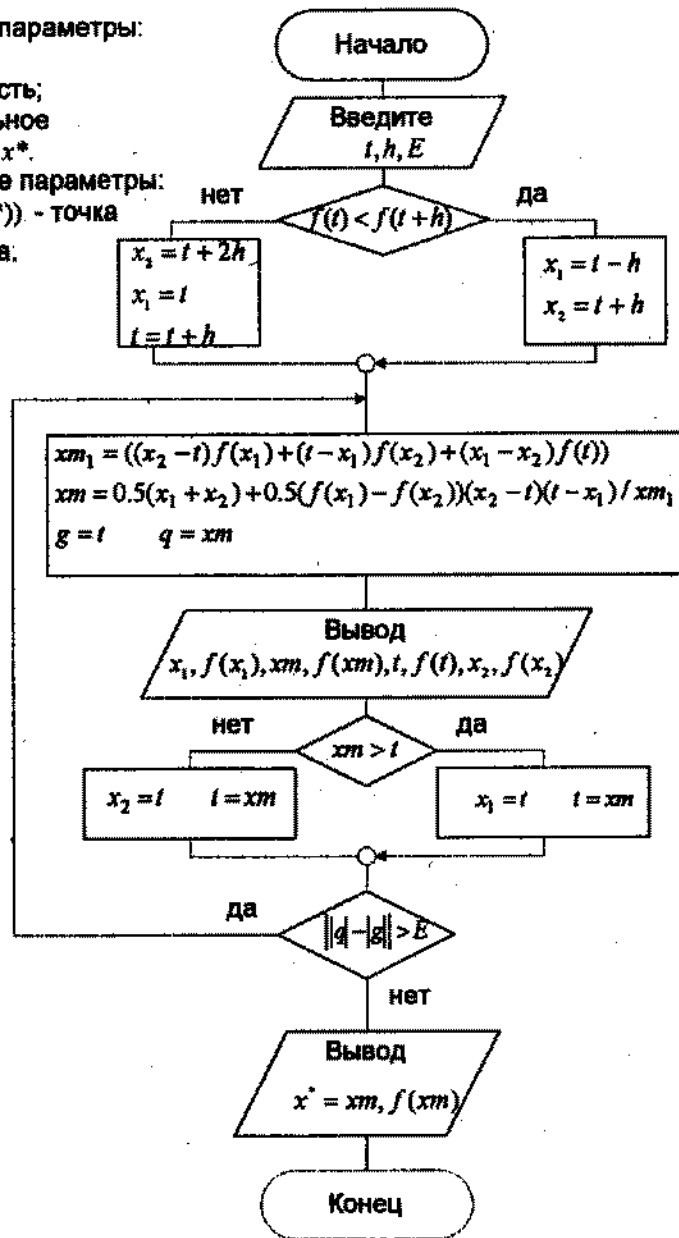


Рис. 7.11

Решение.

Результаты вычислений по приведенному алгоритму с использованием формул (9) и (10) сведены в таблицу 3.

Таблица 3

$n$	$x_i$	$f(x_i)$	$n$	$x_i$	$f(x_i)$
1	0.5	-1.148721	4	0.3541262	-1.174124
	0.04	-1.043721		0.3789044	-1.173546
	1.00	-0.7182818		0.3745916	-1.17376
	1.50	0.0183110		1.00	-0.7182818
2	0.3745916	-1.17376	5	0.3745916	-1.17376
	0.5	-1.148721		0.35738	-1.174138
	0.0470198	-1.043721		0.3541262	-1.174124
	1.00	-0.718282		0.3789044	-1.173546
3	0.3789044	-1.173546	6	0.3541262	-1.174124
	0.3745916	-1.17376		0.3574118	-1.174138
	1.00	-0.7182818		0.35738	-1.174138
	0.0470198	-1.043721		0.3789044	-1.173546

Точность достигнута, полагаем  $x^* = 0.3575$ ,  $f^* = -1.1741$ .

### 7.2.4. Метод Фибоначчи

Процедура предназначена для поиска точки минимума  $x^*$  при заданном числе отсчетов  $n$  (вычисления  $n$  значений минимизируемой функции  $f(x)$ ). Предполагается, что функция  $f(x)$  унимодальная на  $[a, b]$ . Оптимальный последовательный метод определения минимума связан с числами Фибоначчи [14].

Числа Фибоначчи определяются соотношениями

$$\begin{cases} F_1 = 1, & F_2 = 1, \\ F_{n+2} = F_{n+1} + F_n, & n = 1, 2, 3, \dots \end{cases} \quad (11)$$

Последовательный метод Фибоначчи использует на каждом шаге тройку чисел  $(a_k, b_k, \tilde{x}_k)$ , локализирующую минимум функции на отрезке

$$\Delta_k = b_k - a_k = (b - a)F_{n-k+3} / F_{n+2}, \quad 1 \leq k \leq n, \quad a_1 = a, \quad b_1 = b. \quad (12)$$

Точка  $\tilde{x}_k$ ,  $a_k < \tilde{x}_k < b_k$ , с вычисленным значением  $f(\tilde{x}_k) = \min f(x_i)$ , где  $1 \leq i \leq k$ , совпадает с одной из точек

$$\begin{cases} x'_k = a_k + (b - a)F_{n-k+1} / F_{n+2}, \\ x''_k = a_k + (b - a)F_{n-k+2} / F_{n+2}, \end{cases} \quad (13)$$

расположенных на  $[a_k, b_k]$  симметрично относительно его середины.

На каждом  $[a_k; b_k]$ , содержащим точку  $\bar{x}_k$  с предыдущего шага, при выборе следующей точки за  $\bar{x}_{k+1}$  следует принимать ту из точек (13), которая наиболее удалена от  $\bar{x}_k$ .

При заданном количестве отсчетов  $n$  точность вычисления минимизирующей точки  $x^*$ , удовлетворяющей неравенству

$|x^* - x_{\min}| \leq E$ , ( $x_{\min}$  - истинное значение аргумента точки минимума функции) не превышает [11]  $\frac{b-a}{F_{n+2}}$ .

$0 < E < (b-a)/F_{n+1}$ ,  $n \geq 1$ , откуда следует  $\frac{b-a}{F_{n+2}} \leq E < \frac{b-a}{F_{n+1}}$ .

Отметим, что длины последовательных отрезков  $[a_k; b_k]$  удовлетворяет ряду соотношений [4]:

$$L_{n-1} = 2L_n - E, \quad L_{m-1} = L_m + L_{m+1}, \quad 1 < m < n, \quad (14)$$

$$L_{n-j} = F_{j+1}L_n - F_{j-1}E, \quad j=1, 2, \dots, (n-1).$$

Число итераций  $N$  и число отсчетов  $n$  связаны равенством  $n = N + 1$ . (15)

Длина конечного интервала минимизации по методу Фибоначчи определяется по формуле

$$L^n \leq \frac{2L^0}{F_{n+2}}. \quad (16)$$

Если значения функции можно получить только в некоторых заранее известных точках (не обязательно равноотстоящих), то поиск экстремума может быть реализован по алгоритму модифицированного метода Фибоначчи [11].

**Пример 6.**

Выполнить три итерации методом Фибоначчи и определить минимальное значение функции  $f(x) = x^3 - 6x - 1$ , заданной на отрезке  $[1; 2]$ . Указать конечную длину интервала неопределенности.

**Решение.**

Поскольку число отсчетов  $n$  на единицу больше числа итераций  $N$ , то на первом шаге решения вычисляются (13) значения в точках:

$$x_1' = 1 + \frac{F_4}{F_6}(2-1) = 1 + \frac{3}{8}(2-1) = 1.375;$$

$$x_1'' = 1 + \frac{F_5}{F_6}(2-1) = 1 + \frac{5}{8}(2-1) = 1.625.$$

**Структурная схема алгоритма метода Фибоначчи**

Входные параметры:  
 $[a; b]$  – отрезок  
 неопределенности;  
 $n$  – количество отсчетов.  
 Выходные параметры:  
 $(x^*, g^*(x^*))$  – точка минимума.

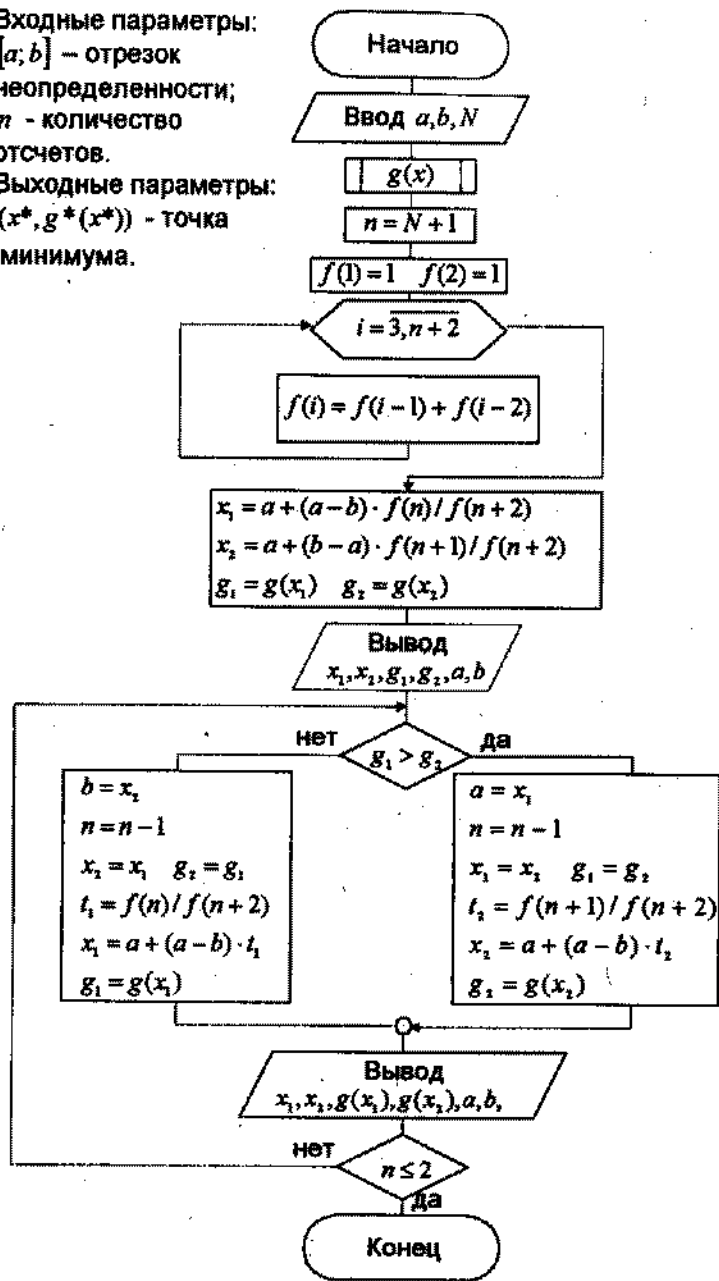


Рис. 7.12

### 7.3. Метод ломаных

Дальнейшие вычисления в соответствии с приведенным алгоритмом приведены в таблице 4.

Таблица 4

$N$	$x'_N$	$x''_N$	$f(x'_N)$	$f(x''_N)$	$[a_N; b_N]$
1	1.375	1.625	-6.65	-6.46	[1; 1.625]
2	1.25	1.375	-6.547	-6.65	[1.25; 1.625]
3	1.375	1.5	-6.65	-6.625	[1.25; 1.5]

Длина конечного интервала  $\Delta = 1.5 - 1.25 = 0.25$  совпадает с

длиной, вычисленной по формуле (16)  $L^n = \frac{2L^0}{F_{n+2}} = \frac{2}{8} = 0.25$ .

$$x^* = 1.4375 \quad f(x^*) = -6.65451.$$

Приведем сравнительную таблицу величин конечных интервалов неопределенности, получаемых после  $n$  отсчетов функции методами золотого сечения, Фибоначчи и дихотомии. Длина начального интервала неопределенности равна  $L^0$ ,  $E$ -точность.

Величины конечных интервалов вычисляются по формулам:

$$L_C^n = 0.618^{n-1} L^0 \text{ - метод золотого сечения;}$$

$$L_\Phi^n = \frac{2L^0}{F_{n+2}} \text{ - метод Фибоначчи;} \quad (17)$$

$$L_D^n = \frac{L^0 + (2^{\frac{n}{2}} - 1)E}{2^{\frac{n}{2}}} \text{ - метод дихотомии.}$$

Таблица 5

$n$	$L_C$	$L_D$	$L_\Phi$
2	0.618	$0.5 + 0.5 * E$	0.666
4	0.236	$0.25 + 0.75 * E$	0.250
6	0.090	$0.125 + 0.875 * E$	0.095
8	0.034	$0.062 + 0.93 * E$	0.036
10	0.0131	$0.031 + 0.96 * E$	0.0138
20	0.0001	$0.0009 + 0.99 * E$	0.0001

Из сравнения результатов следует, что наименее эффективен метод дихотомии.

Метод применяют для минимизации произвольных, в том числе не унимодальных функций [66]. Функция  $f(x)$  должна удовлетворять условию Липшица

$$|f(x') - f(x'')| \leq L |x' - x''|; \quad (18)$$

для  $x', x'' \in [a; b]$ ,  $L > 0$  - константа Липшица.

$$\text{Полагаем } L = \max |f'(x)|.$$

#### Пример 7.

Проверить, удовлетворяет ли функция  $f(x) = \frac{1}{3}x^3 + 2x^2 - 5x + 6$

на отрезке  $[0; 1]$  условию Липшица.

*Решение.*

Вычислим производные функции  $f(x)$ .

$$f'(x) = x^2 + 4x - 5, \quad f''(x) = 2x + 4.$$

На отрезке  $[0; 1]$   $f''(x) > 0$ , следовательно, функция  $f'(x)$  возрастает ( $f''(x) = 0$ , если  $x = -2$ , но  $x = -2 \notin [0; 1]$ ).

Наибольшее значение  $|f'(x)|$  на отрезке  $[0; 1]$  равно 5.

Таким образом, функция  $f(x)$  удовлетворяет условию Липшица, константа Липшица  $L = 5$ .

Пусть функция  $f(x)$  удовлетворяет на отрезке  $[a; b]$  условию Липшица с константой  $L$ .

#### Алгоритм метода ломаных

1. Положим

$$x_1^* = \frac{1}{2L}(f(a) - f(b) + L \cdot (a + b)), \quad (19)$$

$$p_1^* = \frac{1}{2}(f(a) + f(b) + L \cdot (a - b)).$$

2. Вместо пары чисел  $(x_1^*, p_1^*)$  образуем две новые пары  $(x', p_1)$  и  $(x'', p_1)$ :

$$x'_1 = x_1^* - \Delta_1, \quad x''_1 = x_1^* + \Delta_1, \quad p_1 = \frac{1}{2}(f(x_1^*) + p_1^*),$$

$$\Delta_1 = \frac{1}{2L}(f(x_1^*) - p_1^*).$$

3. Из полученных двух пар  $(x'_1, p_1)$  и  $(x''_1, p_1)$  выбираем ту, у которой вторая компонента минимальна (в данном случае - любую па-

ру). Обозначим ее  $(x_2^*, p_2^*)$  и исключим из рассматриваемого множества точек.

4. Вместо пары  $(x_2^*, p_2^*)$  добавляем две новые пары  $(x_2', p_2)$  и  $(x_2'', p_2)$ , компоненты которых находятся по формулам

$$x_2' = x_2^* - \Delta_2, \quad x_2'' = x_2^* + \Delta_2, \quad p_2 = \frac{1}{2}(f(x_2') + p_2^*), \quad \Delta_2 = \frac{1}{2L}(f(x_2') - p_2^*).$$

В результате получим множество из трех пар точек.

5. Из  $n$  полученных на предыдущих шагах пар  $(x, p)$  выбираем ту, у которой вторая компонента минимальна. Обозначим ее  $(x_n^*, p_n^*)$ . Исключаем эту пару из рассматриваемого множества и добавляем вместо нее две новые пары чисел  $(x_n', p_n)$  и  $(x_n'', p_n)$  по формулам

$$x_n' = x_n^* - \Delta_n, \quad x_n'' = x_n^* + \Delta_n, \quad p_n = \frac{1}{2}(f(x_n') + p_n^*), \quad (20)$$

$$\text{где } \Delta_n = \frac{1}{2L}(f(x_n') - p_n^*).$$

6. Полагая  $x^* = \min x_n^*$ ,  $f^*(x) = f(x_n^*)$ , получим приближенное решение задачи минимизации.

Точность определения  $f^*$

$$|f(x_n^*) - f^*| \leq 2L\Delta_n. \quad (21)$$

Геометрически метод ломаных состоит в построении последовательности ломаных, приближающихся к графику  $f(x)$  снизу и имеющих угловые коэффициенты всех звеньев, равные  $\pm L$ .

### Геометрическая иллюстрация метода ломаных

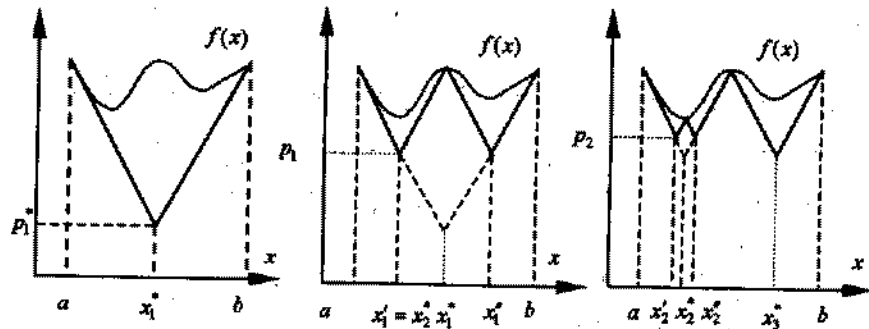


Рис. 7.13

### Пример 8. [66]

Методом ломаных найти точку минимума  $x^*$  и значение  $f^*(x^*)$  функции  $f(x) = \frac{\sin x}{x}$  в точке минимума на отрезке  $[10; 15]$  с точностью  $E = 0.01$ .

Решение.

Функция не унимодальная на отрезке  $[10; 15]$ . Дифференцируемая функция  $f(x)$  имеет на отрезке  $[10; 15]$  ограниченную производную, поскольку

$$|f'(x)| = \left| \frac{x \cos x - \sin x}{x^2} \right| < \frac{|x \cos x| + |\sin x|}{x^2} < \frac{x+1}{x^2} \leq 0.11.$$

Константа Липшица равна 0.11.

Находим по формуле (19)

$$x_1^* = \frac{1}{2 \cdot 0.11} \left( \frac{\sin 10}{10} - \frac{\sin 15}{15} + 0.11 \cdot 25 \right) = 12.056,$$

$$p_1^* = \frac{1}{2} \left( \frac{\sin 10}{10} + \frac{\sin 15}{15} + 0.11 \cdot (-5) \right) = -0.281.$$

Вместо пары точек  $(x_1^*, p_1^*)$  образуем две новые пары  $(x', p_1)$  и  $(x'', p_1)$ :

$$\Delta_1 = \frac{1}{2L}(f(x_1^*) - p_1^*) = \frac{-0.0405 + 0.281}{0.22} = 1.0932,$$

$$x_1' = x_1^* - \Delta_1 = 10.963, \quad x_1'' = x_1^* + \Delta_1 = 13.149,$$

$$p_1 = \frac{1}{2}(f(x_1') + p_1^*) = -0.16075 \approx -0.161.$$

$$2L\Delta_1 = 0.240$$

У двух пар точек  $(10.963; -0.161)$ ,  $(13.149; -0.161)$  одинаковая вторая компонента.

Вычислим две новые пары  $(x_2', p_2)$  и  $(x_2'', p_2)$

$$x_2' = x_2^* - \Delta_2, \quad x_2'' = x_2^* + \Delta_2, \quad p_2 = \frac{1}{2}(f(x_2') + p_2^*),$$

$$\Delta_2 = \frac{1}{2L}(f(x_2') - p_2^*).$$

$$\Delta_2 = \frac{-0.0912 + 0.161}{0.22} = 0.3173, \quad p_2 = -0.126.$$

$$x_2' = x_2^* - \Delta_2 = 10.646, \quad x_2'' = x_2^* + \Delta_2 = 11.28.$$

$$2L\Delta_2 = 0.0698,$$

### Структурная схема алгоритма метода ломаных

Входные параметры:  
 $[a, b]$  – отрезок  
 неопределенности;  
 $E$  – точность;  
 $l$  – параметр.

Выходные параметры:  
 $(x^*, f^*(x^*))$  – точка  
 минимума.

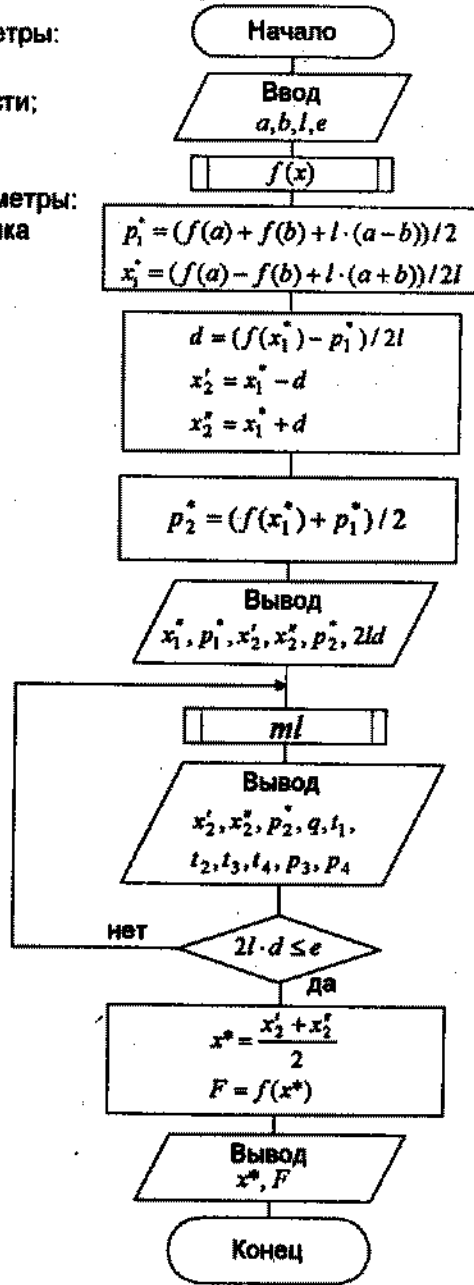


Рис. 7.14

### Процедура метода ломаных

Входные параметры:

$x_2^*, x_3^*, p_2^*$  – координаты точек;  
 $l$  – константа Липшица.

Выходные параметры:

$x_2^*, x_3^*, p_2^*, q, l_1, l_2, l_3, l_4, p_3, p_4$ .

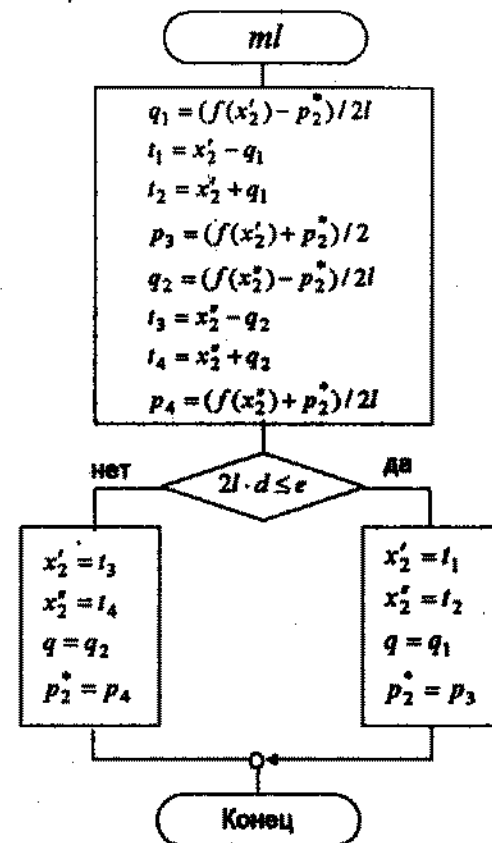


Рис. 7.15

Образует две новые пары:

$$\Delta_3 = \frac{0.0418 + 0.161}{0.22} = 0.9218, \quad p_3 = -0.0596.$$

$$x_3' = 12.227, \quad x_3'' = 14.071.$$

$$2L\Delta_3 = 0.203.$$

Из четырех точек

$$(x_2', p_2) = (10.646; -0.126), \quad (x_2'', p_2) = (11.28; -0.126),$$

$$(x_3', p_3) = (12.227; -0.0596), \quad (x_3'', p_3) = (14.071; -0.0596)$$

выбираем пару  $(x_2^*, p_2)$ ,  $(x_2^*, p_2)$ , т.к.  $p_2 < p_3$ .

Дальнейшие вычисления приведены в таблице 6.

Итерации продолжаются до тех пор, пока  $2L\Delta n > E$ .

Если неравенство  $2L\Delta n \leq E$  истинно, полагают решение найденным,  $x^* = \min x_n^*$ ,  $f^*(x) = f(x^*)$ .

Таблица 6

n	Исключаемые пары (x, p)		2LΔ <sub>n</sub>	Включаемые пары		
	x <sub>n</sub> <sup>*</sup>	p <sub>n</sub> <sup>*</sup>		x <sub>n</sub> <sup>*</sup>	x <sub>n</sub> <sup>*</sup>	p <sub>n</sub>
1	12.056	-0.281	0.240	10.965	13.146	-0.161
2	10.965	-0.161	0.069	10.649	11.280	-0.126
3	13.146	-0.161	0.202	12.227	14.066	-0.059
4	10.649	-0.126	0.038	10.479	10.820	-0.107
5	11.280	-0.126	0.041	11.095	11.465	-0.105
6	10.479	-0.107	0.024	10.369	10.589	-0.095
7	10.820	-0.107	0.016	10.747	10.893	-0.099
8	10.747	-0.099	0.009	10.707	10.787	-0.095
9	10.893	-0.099	0.008	10.858	10.928	-0.095

Из таблицы находим

$$x^* = 10.893, \quad f^*(10.893) = -0.091.$$

#### 7.4. Метод касательных

При решении задачи оптимизации применяются также методы, использующие производные различных порядков, так называемые не прямые методы. Во многих случаях эти методы обеспечивают большую скорость сходимости, чем прямые методы. Иногда применяют смешанную стратегию: на начальном этапе применяют менее трудоемкие методы, а на завершающем этапе минимизации – более точные и быстрые.

Пусть  $f(x)$  – дифференцируемая выпуклая на отрезке  $[a; b]$  функция, причем  $f'(a) \cdot f'(b) < 0$ . Геометрически метод касательных (рис.15) заключается в построении последовательных точек  $x_i \in [a_i; b_i]$ ,  $i = 0, 1, 2, \dots, n$ , являющихся абсциссами точек пересечения касательных к графику функции  $y = f(x)$ , проведенных в граничных точках отрезка  $[a_i; b_i]$ , сходящихся к точке минимума  $x^*$ .

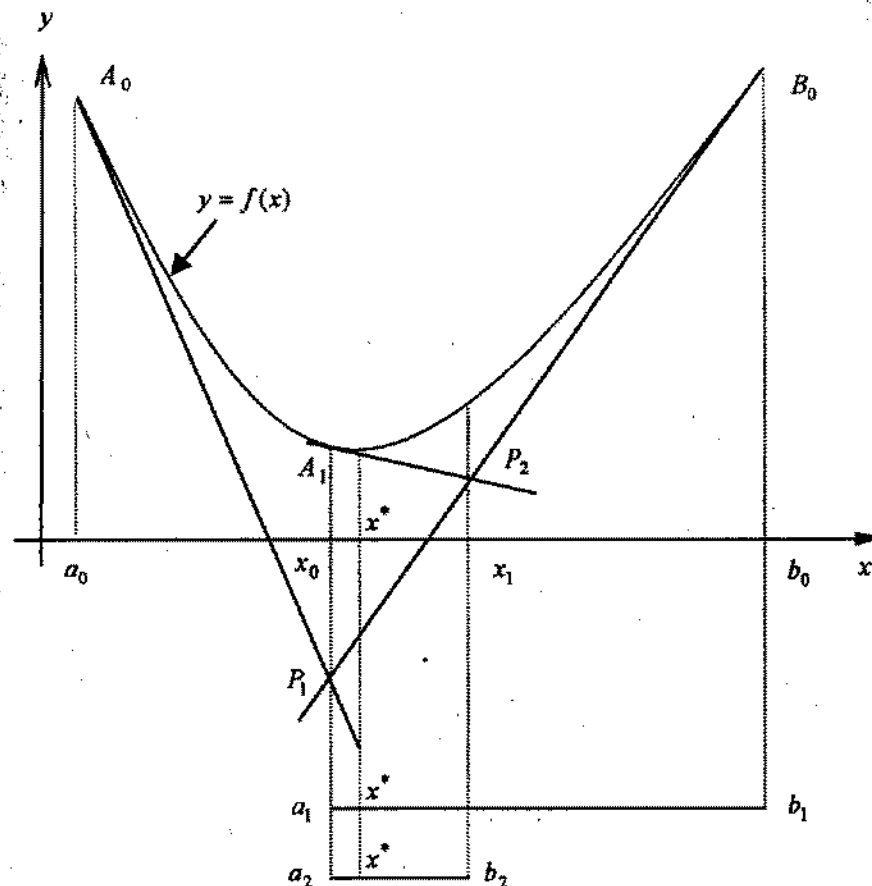


Рис. 7.15

- 1)  $A_0P_1$  и  $B_0P_1$  – касательные к графику функции  $y = f(x)$ , проведенные в точках  $A(a_0, f(a_0))$  и  $B(b_0, f(b_0))$ .  $P_1 \in (A_0P_1; B_0P_1)$ .
- 2)  $P_2$  – точка пересечения касательных  $A_1P_2$  и  $B_0P_2$ .  $A_1(x_0, f(x_0))$ .

#### Алгоритм метода касательных

7. Положить  $a_0 = a$ ,  $b_0 = b$ ,  $i = 0$ .
8. Вычислить  $x_i = \frac{b_i f'(b_i) - a_i f'(a_i) + f(a_i) - f(b_i)}{f'(b_i) - f'(a_i)}$ .
9. Если  $f'(x_i) \geq 0$ , положить  $a_{i+1} = a_i$ ,  $b_{i+1} = x_i$ , перейти к п. 5.
10. в противном случае перейти к п. 4.



11. Положить  $a_{i+1} = x_i$ ,  $b_{i+1} = b_i$ , перейти к п. 5.

12. Если  $|f'(x_i)| < E$  ( $E > 0$  – точность), положить  $x^* = x_i$ ,

$f^* = f(x^*)$ , конец алгоритма.

В противном случае положить  $i = i + 1$  и перейти к п. 2.

Если условие  $f'(a) \cdot f'(b) < 0$  не выполняется, то полагают

- $x^* = a$ , если  $f'(a) > 0$ ,  $f'(b) > 0$ ,
- $x^* = b$ , если  $f'(a) < 0$ ,  $f'(b) < 0$ ,
- $x^* = a$ , если  $f'(a) = 0$ ,
- $x^* = b$ , если  $f'(b) = 0$ .

Таким образом рекуррентная формула метода

$$x_i = \frac{b_i f'(b_i) - a_i f'(a_i) + f(a_i) - f(b_i)}{f'(b_i) - f'(a_i)}, \quad i = 0, 1, 2, \dots \quad (22)$$

Условие окончания поиска минимума

$$|f'(x_n)| < E. \quad (23)$$

#### Пример 9.

Минимизировать функцию  $f(x) = x^2 - \sin x$  на отрезке  $[0; \frac{\pi}{2}]$  методом касательных с точностью  $E = 0.05$ .

*Решение.*

Убедимся, что на отрезке  $[0; \frac{\pi}{2}]$  функция  $f(x)$  выпуклая.

$$f'(x) = 2x - \cos x, \quad f''(x) = 2 + \sin x, \quad f''(x) > 0, \text{ следовательно } f(x)$$

- выпуклая функция,  $f'(0) \cdot f'(\frac{\pi}{2}) < 0$ .

Результаты вычислений сведены в таблицу 7.

Таблица 7

$N$	$a_N$	$b_N$	$x_N$	$f(x_N)$	$f'(x_N)$
0	0	1.507963	0.80374	-0.073958	0.913461
1	0	0.80374	0.42235	-0.2315	-0.06744
2	0.42235	0.80374	0.61688	-0.19795	0.418084
3	0.42235	0.61688	0.52071	-0.226357	0.173957
4	0.42235	0.52071	0.47182	-0.231894	0.052901
5	0.42235	0.47182	0.471598	-0.232454	-0.00736

Точность достигнута.  $x^* = 0.459$ ,  $f^* = -0.232$ .

## 7.5. Метод Ньютона

Пусть  $f(x)$  – выпуклая, дважды дифференцируемая функция на отрезке  $[a; b]$ , причем  $f(x)$  удовлетворяет условию Липшица с константой  $L$  и  $f''(x) \neq 0$ ,  $\forall x \in [a; b]$ .

Метод Ньютона использует производные до второго порядка включительно.

Применяя метод Ньютона–Рафсона к решению уравнения  $f'(x) = 0$  (необходимое условие экстремума), получим рекуррентную формулу метода Ньютона

$$x_{n+1} = x_n - \frac{f'(x_n)}{f''(x_n)}, \quad n = 0, 1, \dots \quad (24)$$

Начальное приближение, обеспечивающее сходимость последовательности (24) к точке минимума, определяется из условия

$$q = \frac{L}{2m^2} |f'(x_0)| < 1, \quad (25)$$

$$0 < m \leq f''(x), \quad \forall x \in [a; b].$$

Практически метод Ньютона применяют на завершающем этапе минимизации функции, когда  $x_0$  – найденное приближенное значение абсциссы точки минимума. Поиск минимума считают завершенным, если выполнено условие (23).

Скорость сходимости

$$|x^* - x_{\min}| \leq \frac{2m}{L} q^{2n}. \quad (26)$$

#### Пример 10.

Вычислить минимум функции  $y = 2x^2 - e^x$  с точностью  $E = 0.001$  методом Ньютона.

*Решение.*

В примере 2 методом дихотомии найдена точка минимума (0.355; -1.174131).

Положим  $x_0 = 0.355$ .

$$f'(x) = 4x - e^x, \quad f''(x) = 4 - e^x, \quad f'(x_0) = 6.18 \cdot 10^{-3}.$$

По формуле (24)

$$x_1 = x_0 - \frac{4x_0 - e^{x_0}}{4 - e^{x_0}} = 0.355 - \frac{4 \cdot 0.355 - e^{0.355}}{4 - e^{0.355}} = 0.35740.$$

$$f'(x_1) = -7.598 \cdot 10^{-3} - \text{условие окончания выполнено.}$$

Уточненные координаты точки минимума  
 $x^* = 0.3574$ ,  $f(x^*) = -1.1741$ .

## 8. ОПТИМИЗАЦИЯ ФУНКЦИИ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

### 8.1. Основные понятия и определения

Пусть на некотором множестве  $U \in \mathbb{R}$  определена функция  $f(\bar{x}) = f(x_1, x_2, \dots, x_n)$ . Требуется найти хотя бы одну точку минимума  $\bar{x}^*$  и (если она существует) вычислить  $f^* = f(\bar{x}^*)$ .

Точка  $(x_1^*, x_2^*, \dots, x_n^*)$  называется точкой глобального (абсолютного) минимума функции  $f(\bar{x})$  на множестве  $U$ , если для всех  $x \in U$   $f(\bar{x}^*) \leq f(\bar{x})$ .

Точка  $(x_1^*, x_2^*, \dots, x_n^*)$  называется точкой локального минимума функции  $f(\bar{x})$ , если существует такое число  $\delta > 0$ , что для всех

$$x \in U_\delta = \{\bar{x} \mid \bar{x} \in U, |\bar{x} - \bar{x}^*| < \delta\},$$

$$f(\bar{x}^*) \leq f(\bar{x}).$$

Всякая точка глобального минимума функции  $f(\bar{x})$  является точкой локального минимума.

Некоторое множество  $U$  называется выпуклым, если вместе с двумя точками  $\bar{x}^{(1)}$  и  $\bar{x}^{(2)} \in U$  оно содержит и отрезок, соединяющий эти точки, т.е. для всех  $\alpha \in [0; 1]$

$$\alpha \bar{x}^{(1)} + (1 - \alpha) \bar{x}^{(2)} \in U. \quad (1)$$

Проверка условия (1), как правило, требует громоздких выкладок. На практике при исследовании выпуклости часто используют геометрические иллюстрации.

#### Пример 1

Является ли множество  $U = \{(x_1, x_2) \mid x_1 x_2 > 1, x_1 > 0, x_2 > 0\}$  выпуклым.

*Решение.*

Покажем, что множество точек  $\bar{x} = (x_1, x_2)$  плоскости

$$U = \{(x_1, x_2) \mid x_1 x_2 > 1, x_1 > 0, x_2 > 0\}$$
 выпукло.

Пусть  $\bar{x}^{(1)} = (x_1^{(1)}, x_2^{(1)})$  и  $\bar{x}^{(2)} = (x_1^{(2)}, x_2^{(2)}) \in U$ .

На отрезке, соединяющем точки  $\bar{x}^{(1)}$  и  $\bar{x}^{(2)}$ , возьмем произвольную точку  $\bar{x} = (x_1, x_2) = \alpha \cdot \bar{x}^{(1)} + (1 - \alpha) \bar{x}^{(2)}$ . Убедимся, что  $\bar{x} \in U$ .

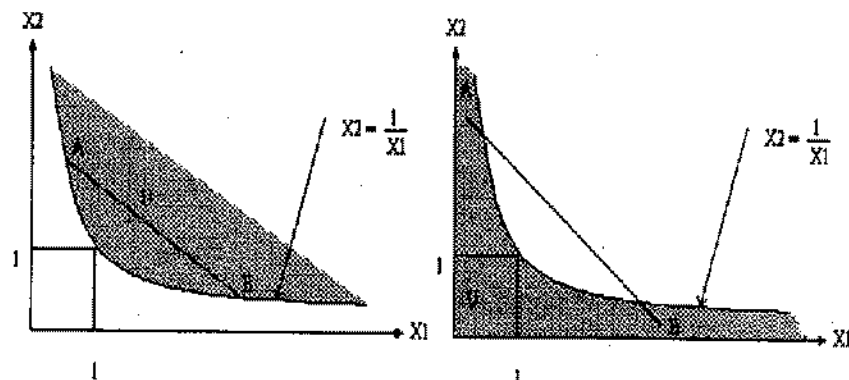


Рис. 8.1

Рис. 8.2

$$\begin{aligned} x_1 \cdot x_2 &= (\alpha x_1^{(1)} + (1 - \alpha) x_1^{(2)}) (\alpha x_2^{(1)} + (1 - \alpha) x_2^{(2)}) = \\ &= \alpha^2 x_1^{(1)} x_2^{(1)} + \alpha(1 - \alpha)(x_1^{(2)} x_2^{(1)} + x_1^{(1)} x_2^{(2)}) + (1 - \alpha)^2 x_1^{(2)} x_2^{(2)} > \\ &> \alpha^2 + 2\alpha(1 - \alpha) + (1 - \alpha)^2 = 1, \text{ т.е. } \bar{x} \in U. \end{aligned}$$

Прямая  $AB$ , соединяющая произвольные точки области  $U$ , принадлежит области  $U$  (рис. 1), следовательно область  $U$  – выпуклая.

#### Пример 2

Является ли множество  $U = \{(x_1, x_2) \mid x_1 x_2 < 1, x_1 > 0, x_2 > 0\}$  выпуклым?

*Решение.*

Выбираем точки  $A$  и  $B$ , принадлежащие области  $U$  (рис. 2). Отрезок  $AB \notin U$ , следовательно, область – не выпуклая.

Функция  $f(\bar{x})$  называется выпуклой на выпуклом множестве  $U$ , если для любых точек  $\bar{x}^{(1)}, \bar{x}^{(2)} \in U$  и произвольного числа  $\alpha \in [0; 1]$  справедливо неравенство:

$$f[\alpha \bar{x}^{(1)} + (1 - \alpha) \bar{x}^{(2)}] \leq \alpha f(\bar{x}^{(1)}) + (1 - \alpha) f(\bar{x}^{(2)}). \quad (2)$$

Возможна геометрическая иллюстрация.

Обозначим

$$\begin{cases} \tilde{x} = \alpha \bar{x}^{(1)} + (1 - \alpha) \bar{x}^{(2)}, \\ \tilde{y} = \alpha f(\bar{x}^{(1)}) + (1 - \alpha) f(\bar{x}^{(2)}). \end{cases}$$

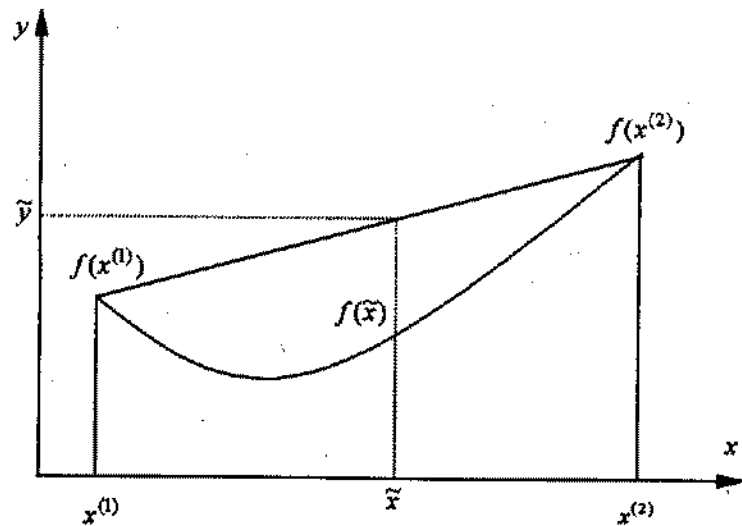


Рис.8.3

**Теорема 1.** (Критерий выпуклости функции.)

Если  $f(\bar{x})$  - дважды дифференцируемая на выпуклом множестве  $U$  функция и матрица ее вторых производных (матрица Гессе)

$$G = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix}, \quad (3)$$

положительно определена при всех  $\bar{x} \in U$ , то функция  $f(\bar{x})$  является выпуклой на  $U$ .

Для проверки положительной определенности матрицы можно воспользоваться критерием Сильвестра:

если все угловые миноры матрицы положительны при  $\bar{x} \in U$ , т.е.

$$\frac{\partial^2 f}{\partial x_1^2} > 0, \quad \frac{\partial^2 f}{\partial x_1^2} \frac{\partial^2 f}{\partial x_2^2} - \left( \frac{\partial^2 f}{\partial x_1 \partial x_2} \right)^2 > 0,$$

матрица положительно определена и функция  $f(\bar{x})$  выпукла на множестве  $U$ .

**Пример 3**

Проверить, является ли функция  $e^{x_1^2+x_2^2}$  выпуклой.

**Решение.**

Составим матрицу Гессе.

$$G = \begin{pmatrix} 2e^{x_1^2+x_2^2}(2x_1^2+1) & 4x_1x_2e^{x_1^2+x_2^2} \\ 4x_1x_2e^{x_1^2+x_2^2} & 2e^{x_1^2+x_2^2}(2x_2^2+1) \end{pmatrix}$$

Угловые миноры:

$$\Delta_1 = 2e^{x_1^2+x_2^2}(2x_1^2+1) > 0, \text{ для } \forall x_1, x_2.$$

$$\Delta_2 = \begin{vmatrix} 2e^{x_1^2+x_2^2}(2x_1^2+1) & 4x_1x_2e^{x_1^2+x_2^2} \\ 4x_1x_2e^{x_1^2+x_2^2} & 2e^{x_1^2+x_2^2}(2x_2^2+1) \end{vmatrix} =$$

$$= 4(e^{x_1^2+x_2^2})^2(2x_1^2+2x_2^2+1) > 0,$$

следовательно функция  $e^{x_1^2+x_2^2}$  - выпуклая.

Справедливо утверждение: всякий локальный минимум выпуклой функции является одновременно и глобальным.

Любую квадратичную функцию

$$f(\bar{x}) = \sum_{i,j=1}^n c_{ij}x_ix_j + \sum_{j=1}^n r_jx_j \quad (4)$$

можно представить в виде.

$$f(\bar{x}) = \frac{1}{2}(Q\bar{x}, \bar{x}) + (\bar{r}, \bar{x}), \quad (5)$$

где

$$\bar{x}^T = (x_1, x_2, \dots, x_n)^T, \quad \bar{r}^T = (r_1, r_2, \dots, r_n)^T - \text{векторы - столбцы,}$$

$Q$  - матрица Гессе для квадратичной функции,  $(\bar{x}, \bar{y})$  - скалярное произведение векторов  $\bar{x}$  и  $\bar{y}$ .

$$Q = (q_{ij}) = (c_{ij} + c_{ji}). \quad (6)$$

**Пример 4.**

$$f(\bar{x}) = x_1^2 - 3x_1x_2 + 3x_2^2 + x_1 - x_2.$$

Найти матрицу  $Q$  и вектор  $\bar{r}$  в представлении (5) функции  $f(\bar{x})$ , выяснить, является ли  $f(\bar{x})$  выпуклой.

**Решение.**

В данном случае  $c_{11}=1$ ;  $c_{12}=-3$ ;  $c_{22}=3$ ;  $r_1=1$ ;  $r_2=-1$ ,

$$Q = \begin{pmatrix} 2 & -3 \\ -3 & 6 \end{pmatrix}, \quad \bar{r} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \bar{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Угловые миноры матрицы  $Q$  положительны ( $\Delta_1 = 2 > 0$ ,  $\Delta_2 = 12 - 9 > 0$ ), следовательно,  $f(x)$  - выпуклая функция.

## 8.2. Классический метод решения задачи оптимизации

Если  $f(\bar{x}) = f(x_1, x_2, \dots, x_n)$  - дважды дифференцируемая функция, то возможно решение задачи оптимизации аналитическим (классическим) методом.

### Алгоритм решения задачи оптимизации

1. Составляется и решается система уравнений

$$\begin{cases} \frac{\partial f(\bar{x})}{\partial x_1} = 0, \\ \dots \\ \frac{\partial f(\bar{x})}{\partial x_n} = 0 \end{cases} \quad \text{— необходимые условия существования минимума}$$

мама  $f(\bar{x})$ , находят  $(\bar{x}_k^1, \bar{x}_k^2, \dots, \bar{x}_k^n)$ ,  $k \geq 1$ .

2. Отбираются точки  $(\bar{x}_m^1, \bar{x}_m^2, \dots, \bar{x}_m^n)$ ,  $m \leq k$ , в которых выполняются и достаточные условия минимума  $f(\bar{x})$ .

3. Отбирается одна точка  $(\bar{x}^*)$ , в которой значение  $f(\bar{x})$  минимально.

### Пример 5.

Решить задачу оптимизации функции

$$f(x, y) = x^2 + 2y^2 - xy + 6y + 15$$

аналитическим методом.

*Решение.*

Проверим, является ли  $f(x, y)$  выпуклой.

Матрица Гессе:

$$Q = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}, \quad Q = \begin{pmatrix} 2 & -1 \\ -1 & 4 \end{pmatrix}.$$

$\Delta_1 = 2 > 0$ ,  $\Delta_2 = 7$  - функция  $f(x, y)$  - выпуклая.

Необходимые условия

$$\begin{cases} \frac{\partial f}{\partial x} = 2x - y, \\ \frac{\partial f}{\partial y} = 4y - x + 6, \end{cases} \quad \begin{cases} 2x - y = 0, \\ 4y - x + 6 = 0, \end{cases} \quad \begin{cases} \bar{x} = -\frac{6}{7}, \\ \bar{y} = -\frac{12}{7}. \end{cases}$$

Достаточные условия минимума

$$\begin{cases} \Delta = \frac{\partial^2 f}{\partial x^2} \cdot \frac{\partial^2 f}{\partial y^2} - \left( \frac{\partial^2 f}{\partial x \partial y} \right)^2 > 0, \\ \frac{\partial^2 f}{\partial x^2} > 0. \end{cases}$$

$$\frac{\partial^2 f}{\partial x^2} = 2 > 0, \quad \frac{\partial^2 f}{\partial y^2} = 4, \quad \frac{\partial^2 f}{\partial x \partial y} = -1.$$

$$\Delta = 2 \cdot 4 - 1 = 7 > 0, \quad \begin{cases} x^* = -\frac{6}{7}, \\ y^* = -\frac{12}{7} \end{cases} \quad \text{— точка минимума, } f(\bar{x}^*) = \frac{69}{7}.$$

Классический метод имеет ограниченную область применения, поскольку на практике:

- функция является сложной или разрывной,
- нельзя или трудно вычислить частные производные,
- проверка достаточных условий - сложная.

## 8.3. Поисковые методы оптимизации

Поисковые методы являются универсальным средством решения задач оптимизации, в том числе и таких, которые не решаются классическими методами.

Поисковыми называются методы нахождения экстремума произвольной функции  $F(\bar{x})$ , относительно которой мы не располагаем полными данными, например:

- аналитическое выражение функции  $F(\bar{x})$  не известно или известно не полностью и имеется возможность только вычислять или измерять ее значения в отдельных точках,
- аналитическое выражение  $F(\bar{x})$  известно, но сложно для вычислений.

Различают следующие виды поисковых методов:

- полный перебор (слепой поиск),
  - случайный поиск,
  - целенаправленный поиск.
- Остановимся на методах целенаправленного поиска.

## 8.4. Метод конфигураций

Методы поиска минимума функции, основанные на вычислении только самой функции, называются методами нулевого порядка. В других методах используется первые, вторые или более высокие степени производных.

В основе метода нулевого порядка - метода конфигураций - лежит идея покоординатного поиска с постоянным приращением  $\Delta x$ . На каждой итерации координаты поочередно меняются на  $\pm \Delta x_i$ ,  $i = 1, 2, \dots, n$  и фиксируется точка с наилучшим (наименьшим при поиске точки минимума) значением целевой функции  $f(x)$ . Цикл изменения всех координат заканчивается определением исходной точки для новой итерации.

Начальными данными поиска являются начальная точка  $\bar{X}^0 = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$ , величина приращения  $\Delta x$  и точность поиска  $\varepsilon < \Delta x$ . При выполнении условия

$$|\bar{x}^{(k-1)} - \bar{x}^{(k)}| < \varepsilon, \quad (7)$$

где  $k$  - номер итерации, процесс определения минимума завершается.

Траектория поиска - ломаная линия с ортогональными равными звеньями. В результате поиска будет получена последовательность вершин траектории, ведущей к стационарной точке  $\bar{x}^*$ , либо к точке, лежащей на "хребте". В обоих случаях процедура поиска заканчивается тем, что для конечной точки  $k$ -й итерации не удастся найти приращения, улучшающее целевую функцию.

Уменьшив приращение  $\Delta x$  можно получить новое решение. Новая остановка процедуры при  $\Delta x \geq \varepsilon$  произойдет в точке  $\bar{X}_{\min}$ .

## 8.5. Градиентные методы оптимизации

Пусть задана некоторая функция  $f(x_1, x_2, \dots, x_n) = f(\bar{x})$ ,  $i = \overline{1, n}$ .

Требуется определить такие значения  $x_i \in C_n$ ,  $i = \overline{1, n}$ , для которых функция  $f(\bar{x})$  принимает экстремальные значения, т.е.  $f^* = \text{extr } f(\bar{x})$ .

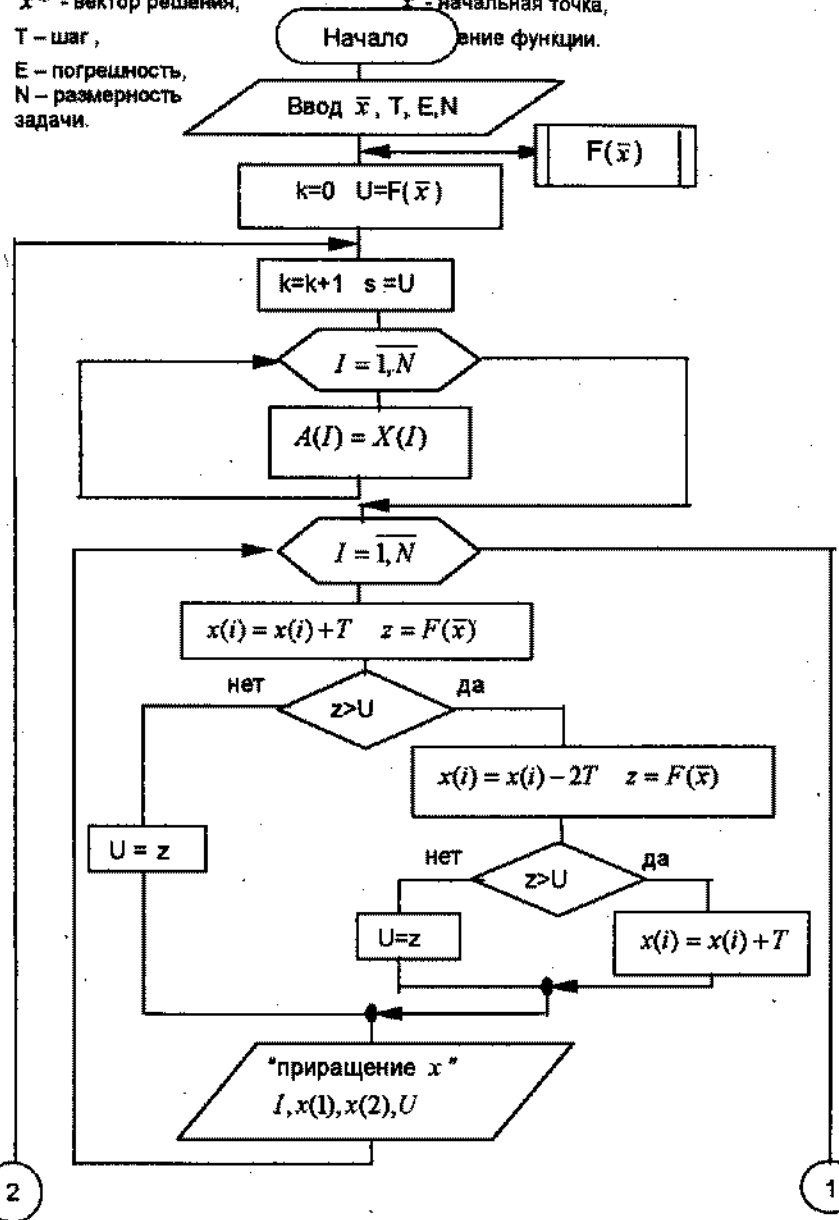
Переменные  $x_i$  - параметры оптимизации (корни уравнений, параметры технической системы).

Выпуклая функция  $f(\bar{x})$  - функция критерия оптимальности, или целевая функция (показатель качества и т.п.).

## Структурная схема алгоритма метода конфигураций

Входные параметры:  
 $\bar{x}^*$  - вектор решения,  
 $T$  - шаг,  
 $E$  - погрешность,  
 $N$  - размерность задачи.

Выходные параметры:  
 $\bar{x}$  - начальная точка,  
 $F(\bar{x})$  - значение функции.



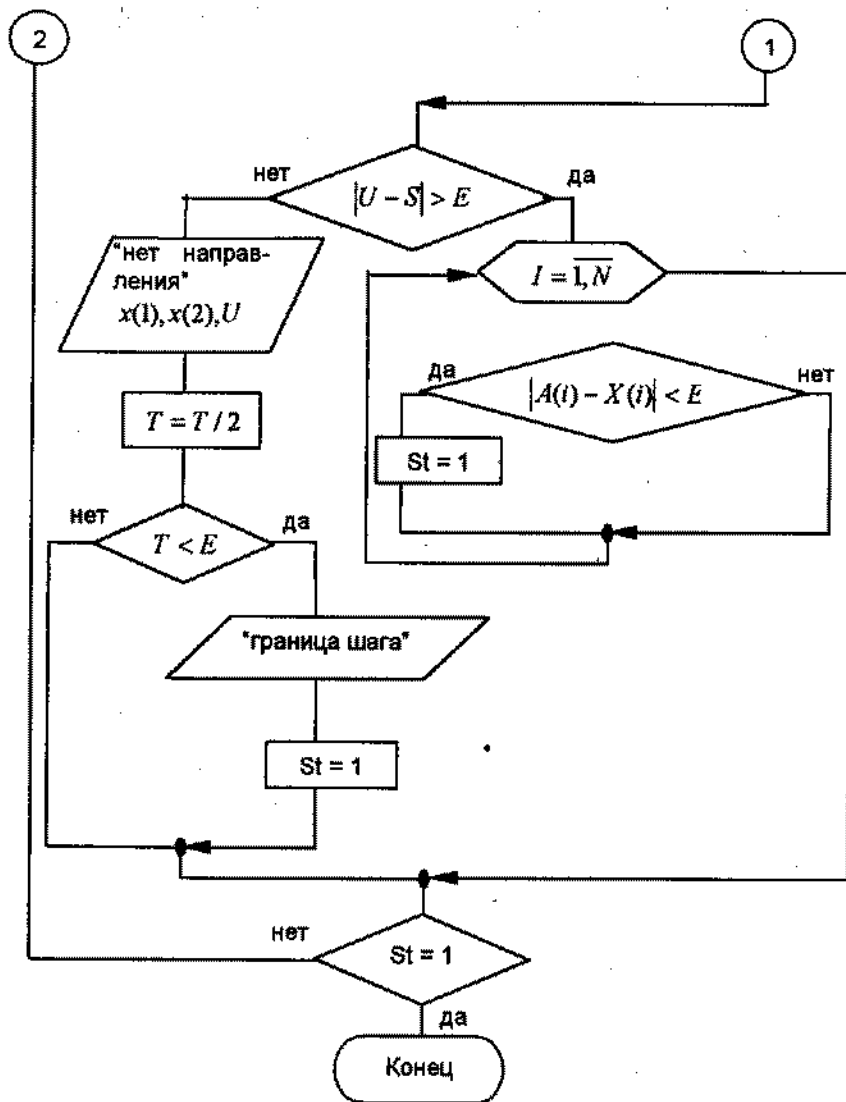


Рис.В. 4.

Будем говорить о минимуме  $f(\bar{x})$ . Не ограничивая общности остановимся на случае функции двух переменных.  
 Линией уровня функции  $z = f(x, y)$  называется множество тех и только тех точек области определения функции  $f(x, y)$ , для которых значение функции равно постоянной (const).  
 Градиент функции – вектор

$$\nabla(x, y) = \overline{\text{grad } f(x, y)} = \frac{\partial f(x, y)}{\partial x} \bar{i} + \frac{\partial f(x, y)}{\partial y} \bar{j}.$$

Свойства градиента:

- указывает направление наибольшего возрастания функции  $f(x, y)$ ,
- длина вектора равна наибольшей скорости возрастания,
- $\overline{\text{grad } f(x, y)}$  направлен по нормали к линии уровня, проходящей через  $(x, y)$ ,
- $\overline{\text{grad } f(x, y)}$  направлен по касательной к поверхности  $z = f(x, y)$  в точке  $(x, y)$ .

Градиентные методы оптимизации – итерационные методы, в которых направление движения к точке минимума  $f(\bar{x})$  задаётся (с точностью до знака) направлением градиента этой функции.

Методы называются  $p$ -шаговыми, если при построении следующей итерации используются результаты  $p$  предыдущих итераций.

Рассмотрим одношаговые методы, использующие первые производные  $f'(\bar{x})$ .

Итерационный процесс градиентных методов описывается рекуррентной формулой

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} - \lambda^{(k)} \overline{f'(\bar{x}^{(k)})}, \quad k = 0, 1, \dots, \quad (8)$$

причем  $f(\bar{x}^{(0)}) > f(\bar{x}^{(1)}) > \dots > f(\bar{x}^{(k)}) > \dots \rightarrow \min f(\bar{x}^*)$ .

Методы поиска с использованием градиента функции различаются выбором параметра  $\lambda$ .

Процесс поиска заканчивается, если выполняется какое-либо из условий окончания поиска ( $E$ -достаточно малое число, точность):

1. Если  $f^*$  известно, то требуется выполнение условия  $f(\bar{x}) - f^*(\bar{x}^*) \leq E$ . (9)

2.  $|\overline{\text{grad } f(\bar{x})}| \cdot \lambda^{(k)} \leq E$ . (10)

3.  $\begin{cases} \left| \frac{\partial f}{\partial x_1} \right| < E, \\ \left| \frac{\partial f}{\partial x_2} \right| < E. \end{cases}$  (11)

4.  $\|\overline{f'(\bar{x}^{(k)})}\| = \sqrt{\sum_{i=1}^2 \left( \frac{\partial f(\bar{x}^{(k)})}{\partial x_i} \right)^2} \leq E$ . (12)

## 8.6. Метод градиентного поиска с дроблением шага (GR)

Рекуррентная формула (координатная форма записи формулы (8)).

$$\begin{cases} x^{(k+1)} = x^{(k)} - \lambda^{(k)} \frac{\partial f}{\partial x} \Big|_{(x_k, y_k)} \\ y^{(k+1)} = y^{(k)} - \lambda^{(k)} \frac{\partial f}{\partial y} \Big|_{(x_k, y_k)} \end{cases} \quad k=0, 1, \dots$$

$\lambda^{(k)}$  - шаг (параметр).

Величина  $\lambda_k$  на каждой итерации выбирается так, чтобы выполнялось условие

$$f(x^{(k+1)}, y^{(k+1)}) < f(x^{(k)}, y^{(k)}). \quad (14)$$

Если при выбранном значении  $\lambda^{(k)}$  условие (14) не выполняется, следует уменьшить шаг ( $\lambda = \lambda/2$ ) и вновь вычислить координаты точки  $(x^{(k+1)}, y^{(k+1)})$  (рис. 7).

Процедура продолжается до выполнения условия (14). Затем переходят к вычислению координат следующей точки траектории поиска минимума. Итерации заканчиваются, если выполнено какое-либо из условий (9)-(12).

*Геометрическая иллюстрация методов градиентного спуска*

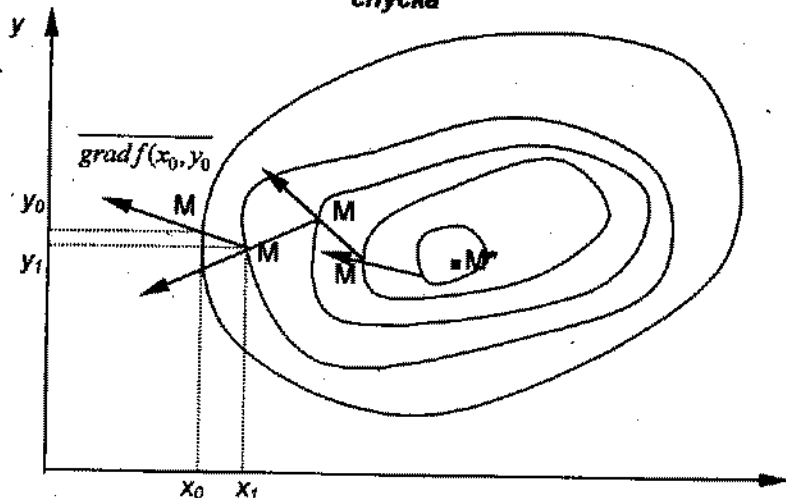


Рис.8.5

## Схема алгоритма методов градиентного спуска

Входные параметры:

$(x, y)$  - начальная точка,

$\lambda_0$  - начальный шаг,

$E$  - точность.

Выходные параметры:

$x^*, y^*, F^*$  - координаты точки минимума и значение целевой функции в точке минимума,

$k$  - число итераций.

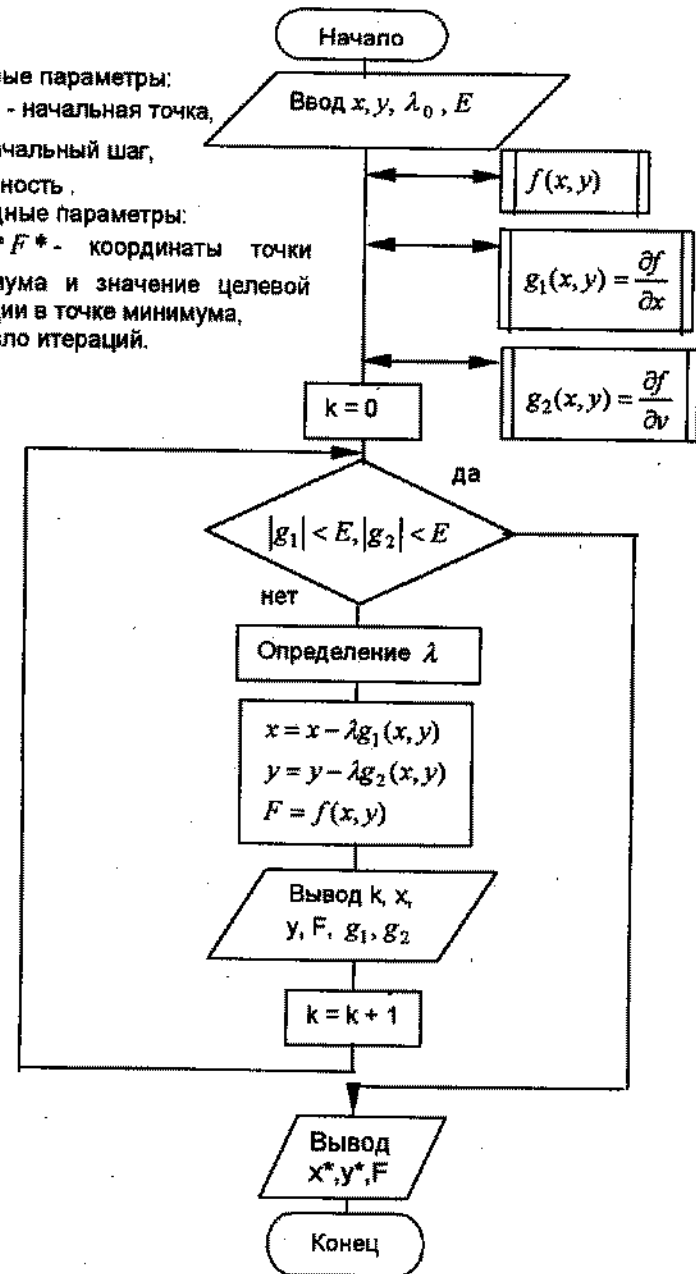


Рис.8.6

Пример 6

Определить точку минимума функции

$$f(x, y) = x^2 + 2y^2 + xy - 7x - 7y,$$

приняв начальное значение  $\lambda_0 = 0.5$ ,  $x^{(0)} = 1$ ,  $y^{(0)} = 0$  методом GR.

Решение.

Вычислим  $f_0(x^{(0)}, y^{(0)})$ ,  $\frac{\partial f}{\partial x} \Big|_{(x^{(0)}, y^{(0)})}$ ,  $\frac{\partial f}{\partial y} \Big|_{(x^{(0)}, y^{(0)})}$

$$f_0(x^{(0)}, y^{(0)}) = f_0(1; 0) = -6,$$

$$\frac{\partial f}{\partial x} = 2x + y - 7, \quad \frac{\partial f}{\partial y} = 4y + x - 7, \quad \frac{\partial f}{\partial x}(1, 0) = -5, \quad \frac{\partial f}{\partial y}(1, 0) = -6.$$

По формуле (13)

$$\begin{cases} x^{(1)} = 1 - 0.5(-5) = 3.5, \\ y^{(1)} = 0 - 0.5(-6) = 3, \end{cases} \quad f_1(x^{(1)}, y^{(1)}) = f(3.5; 3) = -4.75.$$

Так как  $f_1 > f_0$ , следует уменьшить шаг  $\lambda^{(0)}$ .

Пусть  $\lambda^{(0)} = 0.25$ . Пересчитаем  $x^{(1)}$  и  $y^{(1)}$ .

$$\begin{cases} x^{(1)} = 1 - 0.25(-5) = 2.25, \\ y^{(1)} = 0 - 0.25(-6) = 1.5. \end{cases} \quad f_1(2.25; 1.5) = -13.3125 < f_0.$$

Дальнейшие вычисления сведены в таблицу 1.

Таблица 1

k	$x_k$	$y_k$	$\lambda_k$	$\frac{\partial f}{\partial x}$	$\frac{\partial f}{\partial y}$	$f(x_k, y_k)$	
0	1	0		-5	-6	-6	
1	3.5000	3	0.5			-4.7500	$f_1 > f_0$ $\lambda = \lambda/2$
	2.2500	1.5000	0.25	-1	1.250	-13.3125	$f_1 < f_0$
2	2.5000	1.1800	0.25	-1	1.250	-13.7752	$f_2 < f_1$
3	2.9100	1.0700	0.5	-0.820	0.22	-13.9880	$f_3 < f_2$
4	2.9370	1.0225	0.25	0.110	0.190	-13.9965	$f_4 < f_3$
5	2.9870	1.0090	0.5	-0.100	0.027	-13.9980	$f_5 < f_4$
6	2.9955	0.9975	0.5	-0.017	0.023	-14	$f_6 < f_5$

Принимаем  $\begin{cases} x^* = 2.9955, \\ y^* = 0.9975, \end{cases} \quad f^* = -14.$

Необходимость вычисления шага  $\lambda$  на каждой итерации связана с тем, что при достаточно большом значении  $\lambda$  возможно нарушение условия (14). Малые значения шага существенно уменьшают скорость сходимости процесса поиска точки минимума. Как правило, начинают с шага  $\lambda_0$ , удовлетворяющего условию  $0 < \lambda_0 \leq 1$ .

Схема алгоритма поиска шага в методе GR

Входные параметры:

$\lambda_0$  - начальный шаг,

$(x, y)$  - текущая точка.

Выходные параметры:

$\lambda$  - шаг.

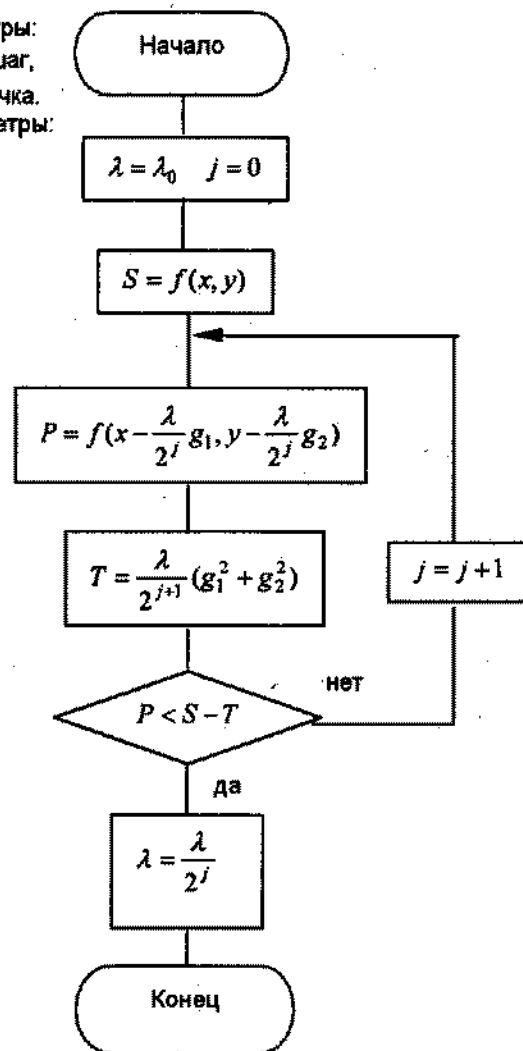


Рис. 8.7



## 8.7. Методы наискорейшего спуска (NS)

Основные пункты алгоритма поиска минимума  $f(\bar{x})$ ,  $\bar{x} = (x_1, x_2, \dots, x_n)$  методом наискорейшего спуска.

1. Определить  $\overline{\text{grad}} f(\bar{x})$  в начальной точке  $M_0$ .
2. Осуществить перемещение по направлению антиградиента из точки  $M_0(\bar{x}^{(0)})$  в точку  $M_1(\bar{x}^{(1)})$ , в которой функция  $f(\bar{x})$  достигает условного минимума на направлении  $M_0M_1$  - касательной к линии уровня в точке  $M_1$ .

В точке  $M_1$  функция  $f(x, y)$  принимает минимальное значение на направлении  $M_0M_1$ . В точке  $M_2$ , функция имеет минимальное значение на направлении  $M_1M_2$  и т.д.

Итерации повторяются для точек  $(\bar{x}^{(1)})$ ,  $(\bar{x}^{(2)})$ , ... до тех пор, пока не выполнится условие окончания поиска.

**Геометрическая иллюстрация метода наискорейшего спуска**

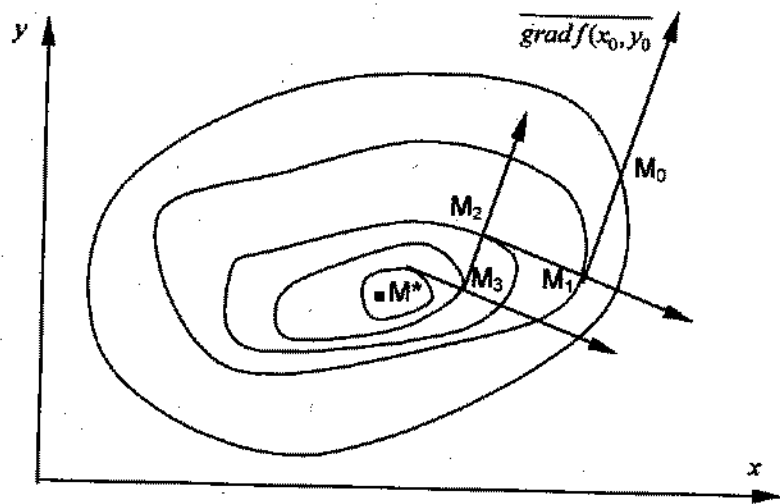


Рис.8.8

Свойства траектории поиска  $M_0, M_1, M_2, \dots, M^*$ .

- уровня  $f(x, y)$  в точке  $M_i$  и является касательной к линии уровня в точке  $M_j$ ,
- смежные отрезки траектории  $M_iM_j$  и  $M_jM_k$  взаимно перпендикулярны,
- для функций с эллиптическими линиями уровня точки  $M_0, M_2, \dots, M^*$  лежат на одной прямой  $M_0M^*$ , а точки  $M_1, M_3, \dots, M^*$  на прямой  $M_1M^*$ .

Рассмотрим методы определения длины отрезка  $M_iM_j$  тако- го, чтобы в точке  $M_j$  функция  $f(\bar{x})$  достигала условного минимума.

1) Аналитический метод наискорейшего спуска (NSA).

Пусть функция  $f(\bar{x})$  непрерывна и имеет непрерывные част- ные производные  $\frac{\partial f}{\partial x}$ ,  $\frac{\partial f}{\partial y}$ . Составляем функцию

$$\Phi(\lambda) = f(x^{(k)} - \lambda^{(k)} \frac{\partial f(x^{(k)}, y^{(k)})}{\partial x}, y^{(k)} - \lambda^{(k)} \frac{\partial f(x^{(k)}, y^{(k)})}{\partial y}). \quad (15)$$

Из условия минимума функции  $\Phi(\lambda)$

$$\Phi'(\lambda) = 0 \quad (16)$$

получим искомое значение  $\lambda^{(k)} = \lambda^*(x^{(k)}, y^{(k)})$ .

Итак, в методе NSA (наискорейшего спуска аналитический) следует решать задачу одномерной оптимизации, исходя из аналитического условия существования минимума функции одной пере- менной.

Практически NSA имеет смысл применять лишь в том случае, когда уравнение (16), например, является алгебраическим уравне- нием и его корни легко находятся. В качестве  $\lambda^{(k)}$  берется наи- больший положительный корень. В случае, если  $f(x)$  - квадратич- ная функция, величина  $\lambda^{(k)}$  может быть найдена в явном виде

$$\lambda^{(k)} = \frac{(\overline{f'}^{(k)}, \overline{f'}^{(k)})}{(Q\overline{f'}^{(k)}, Q\overline{f'}^{(k)})}. \quad (17)$$

где

$$\overline{f'}^{(k)} = Q\bar{x} + \bar{r}, \quad f(\bar{x}) = \frac{1}{2}(Q\bar{x}, \bar{x}) + (\bar{r}, \bar{x}). \quad (18)$$

### Пример 7.

Найти аналитическое выражение  $\lambda^{(k)}$  для минимизации методом NSA функции  $f(x, y) = x^2 + 4y^2$ .

Решение.

Составим функцию

$$\Phi(\lambda) = (x^{(k)} - \lambda^{(k)} \cdot 2x^{(k)})^2 + 4(y^{(k)} - \lambda^{(k)} \cdot 8y^{(k)})^2.$$

Найдем производную  $\Phi'(\lambda)$  и приравняем ее к нулю.

$$\Phi'(\lambda) = -4(x^{(k)})^2(1 - 2\lambda^{(k)}) - 64(y^{(k)})^2(1 - 8\lambda^{(k)}) = 0,$$

откуда

$$\lambda^{(k)} = \frac{(x^{(k)})^2 + 16(y^{(k)})^2}{2(x^{(k)})^2 + 128(y^{(k)})^2}.$$

Решим эту же задачу с помощью формул (17) - (18).

Согласно (4) - (6)

$$c_{11} = 1, c_{12} = 0, c_{22} = 4, c_{21} = 0, r_1 = 0, r_2 = 0.$$

$$Q = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}, \quad \bar{r} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \bar{x} = \begin{pmatrix} x_k \\ y_k \end{pmatrix}.$$

$$\text{Вычислим } \overline{f'(k)} = \left( \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}, \begin{pmatrix} x^{(k)} \\ y^{(k)} \end{pmatrix} \right) = \begin{pmatrix} 2x^{(k)} \\ 8y^{(k)} \end{pmatrix}.$$

$$Q \cdot \overline{f'(k)} = \left( \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}, \begin{pmatrix} 2x^{(k)} \\ 8y^{(k)} \end{pmatrix} \right) = \begin{pmatrix} 4x^{(k)} \\ 64y^{(k)} \end{pmatrix}.$$

Скалярное произведение

$$\overline{f'(k)}, \overline{f'(k)} = \left( \begin{pmatrix} 2x^{(k)} \\ 8y^{(k)} \end{pmatrix}, \begin{pmatrix} 2x^{(k)} \\ 8y^{(k)} \end{pmatrix} \right) = 4(x^{(k)})^2 + 64(y^{(k)})^2,$$

$$\begin{aligned} (Q \cdot \overline{f'(k)}, \overline{f'(k)}) &= \left( \begin{pmatrix} 4x^{(k)} \\ 64y^{(k)} \end{pmatrix}, \begin{pmatrix} 2x^{(k)} \\ 8y^{(k)} \end{pmatrix} \right) = \\ &= 8(x^{(k)})^2 + 512(y^{(k)})^2. \end{aligned}$$

Подставляя в (17), получим

$$\lambda^{(k)} = \frac{(x^{(k)})^2 + 16(y^{(k)})^2}{2(x^{(k)})^2 + 128(y^{(k)})^2}.$$

2) Численный метод наискорейшего спуска (NSD) предполагает определение значения  $\lambda^{(k)}$  производить методами одномерной оптимизации; начальный отрезок оптимизации [0;1].

Структурная схема определения  $\lambda^{(k)}$  приведена на рис. 10.

Градиентные методы достаточно просты в реализации и могут быть использованы для минимизации различных функций. Градиентные методы обладают свойством самоисправления.

Сходимость методов зависит от свойств функции и определена теоремой 1 [11].

### Теорема 2.

Если функция  $f(\bar{x})$  ограничена снизу, ее градиент удовлетворяет условию Липшица, то, какова бы ни была начальная точка  $\bar{x}_0$ :

$$\|\bar{r}'(\bar{x}^{(k)})\| \rightarrow 0 \text{ при } k \rightarrow \infty$$

Скорость сходимости методов можно оценить в случае, если функция  $f(\bar{x})$  удовлетворяет достаточно жестким условиям (например, является сильно выпуклой [10]). В этом случае скорость сходимости можно оценить [48]

$$\|x^{(k+1)} - x^*\| \leq \frac{M-m}{M+m} \|x^{(k)} - x^*\|, \quad (19)$$

$m, M$  - оценки минимального и максимального собственных чисел матрицы  $f''(x^*)$ .

Если  $m \approx M$ , то сходимость метода достаточно высока, если

$\frac{m}{M} \ll 1$ , то метод сходится медленно.

Для улучшения сходимости существуют различные приемы. Некоторые из них изложены в [48].

Проиллюстрируем так называемый "овражный эффект" - случай медленной сходимости.

$f(x, y) = x^2 + 16y^2$ . Линии уровня функции сильно вытянуты вдоль оси  $ox$ .

$$\text{Матрица вторых производных } G = \begin{pmatrix} 2 & 0 \\ 0 & 32 \end{pmatrix}.$$

Схематически траектория поиска минимума приведена на рис 9.

Траектория характеризуется довольно быстрым спуском на "дно" оврага и затем медленным зигзагообразным движением в точку минимума. Возможный метод решения задачи оптимизации предложен в [48].

**Структурная схема нахождения шага  $\lambda$  в методе NSD**

Входные параметры:  
 $(x, y)$  – текущая точка  
 $[a, b]$  – отрезок для поиска  $\lambda$   
 $\epsilon$  – точность поиска минимума.  
 Выходные параметры:  
 $\lambda$  – искомый шаг.

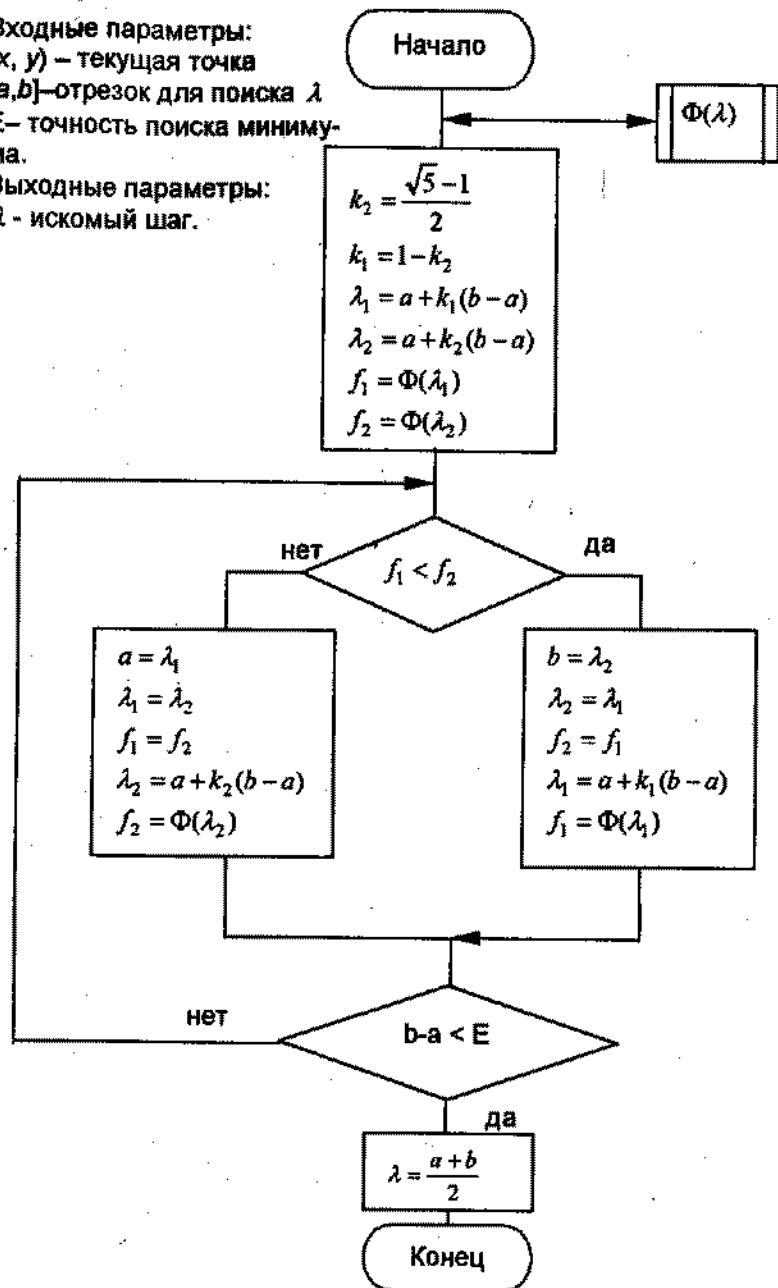


Рис. 8.10

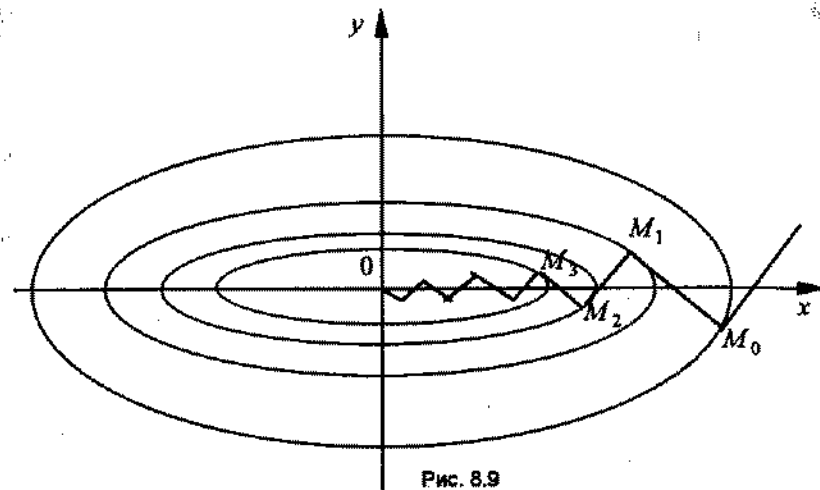


Рис. 8.9

**8.8. Метод сопряженных направлений**

Сформулируем ряд определений и утверждений.

**Определение 1.**

Два вектора  $\bar{x}$  и  $\bar{y}$  называются  $Q$ -сопряженными в пространстве  $E_n$  (или сопряженными по отношению к матрице  $Q$ ), если  $(\bar{x}, Q\bar{y}) = 0$ . (20)

Если  $Q = E$  - единичная матрица, то вектора  $\bar{x}$  и  $\bar{y}$  ортогональны.

**Определение 2.**

Векторы  $\bar{p}_0, \bar{p}_1, \dots, \bar{p}_k$  называются сопряженными относительно матрицы  $Q$ , если скалярное произведение  $(Q\bar{p}_i, \bar{p}_j) = 0$  при всех  $i \neq j, 0 \leq i, j \leq k$ . (21)

**Утверждение 1.**

В  $n$ -мерном пространстве не может быть более  $n$  ненулевых взаимно ортогональных векторов.

**Утверждение 2.**

$Q$ -сопряженные вектора линейно-независимы.

Доказательства этих утверждений приведены в [36].

Метод сопряженных направлений заключается в построении последовательности

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} + \lambda^{(k)} \bar{S}^{(k)}, \quad (22)$$

где  $\lambda^{(k)}$  выбирают из условия минимума функции  $f(\bar{x})$  в направлении  $\bar{S}_k$ ;  $\bar{S}_k$  - направление спуска на шаге  $k$ .

### Алгоритм метода сопряженных направлений

1. Задать произвольные начальное приближение  $\bar{x}^{(0)}$  и начальное направление спуска  $\bar{S}^{(0)}$ .
2. На  $k$ -ом шаге ( $k=0, 1, \dots, n$ ) решить задачу одномерной оптимизации по  $\lambda > 0$  функции  $f(\bar{x}^{(k)} + \lambda^{(k)} \bar{S}^{(k)})$ .

Определить  $\lambda^{(k)}$  и  $\bar{x}^{(k+1)} = \bar{x}^{(k)} + \lambda^{(k)} \bar{S}^{(k)}$ .

3. Вычислить  $\bar{f}'(\bar{x}^{(k+1)})$ .
4. Если  $\bar{f}'(\bar{x}^{(k+1)}) = 0$  - минимум найден, переход на конец алгоритма. В противном случае переход к пункту 5.
5. Вычислить  $f''(\bar{x}^{(k)})$ , произведение  $(f''(\bar{x}^{(k)}) \cdot \bar{S}^{(k)})$ .
6. Определить направление  $\bar{S}^{(k+1)}$  из условия равенства нулю скалярного произведения  $(\bar{S}^{(k+1)}, f''(\bar{x}^{(k)}) \cdot \bar{S}^{(k)}) = 0$ .
7. Переход к п 2.

### Пример 8.

Методом сопряженных направлений вычислить минимум функции  $f(x, y) = x^2 + 16y^2$ .

Решение.

Выберем  $\bar{x}^{(0)} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ ,  $\bar{S}^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ , тогда  $f(x^{(0)}, y^{(0)}) = 4$ .

Обозначим

$$\bar{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \bar{S}^{(k)} = \begin{pmatrix} S_1^{(k)} \\ S_2^{(k)} \end{pmatrix}$$

Первая итерация (формула(22)):

$$\bar{x}^{(1)} = \bar{x}^{(0)} + \lambda^{(0)} \cdot \bar{S}^{(0)}$$

Координатная форма записи формулы(22)):

$$\begin{cases} x^{(1)} = 2 + \lambda^{(0)}, \\ y^{(1)} = 0 + \lambda^{(0)}. \end{cases}$$

$$\Phi(\lambda^{(0)}) = f(x^{(1)}, y^{(1)}, \lambda^{(0)}) = (2 + \lambda^{(0)})^2 + 16(\lambda^{(0)})^2$$

Из условия  $\frac{df}{d\lambda^{(0)}} = 0$  определим  $\lambda^{(0)}$ .

$$2 \cdot (2 + \lambda^{(0)}) + 32\lambda^{(0)} = 0, \quad \lambda^{(0)} = -\frac{2}{17}$$

$$\begin{cases} x^{(1)} = 2 - \frac{2}{17} = \frac{32}{17}, \\ y^{(1)} = -\frac{2}{17}. \end{cases}$$

$$f(x^{(1)}, y^{(1)}, \lambda^{(0)}) = \left(\frac{32}{17}\right)^2 + 16 \cdot \left(-\frac{2}{17}\right)^2 = 3.765$$

$$\bar{f}' = \begin{pmatrix} 2x \\ 32y \end{pmatrix}, \quad \bar{f}'(x^{(1)}, y^{(1)}) = \begin{pmatrix} \frac{64}{17} \\ -64 \\ \frac{64}{17} \end{pmatrix}$$

$\bar{f}'(x^{(1)}, y^{(1)}) \neq 0$  - условие окончания не выполнено.

Найдем направление спуска  $\bar{S}^{(1)}$ .

$$f''(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}, \quad f''(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 32 \end{pmatrix}$$

$$f''(x, y) \cdot \bar{S}^{(0)} = \begin{pmatrix} 2 & 0 \\ 0 & 32 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 32 \end{pmatrix}$$

Скалярное произведение

$$\left( \begin{pmatrix} S_1^{(1)} \\ S_2^{(1)} \end{pmatrix}, \begin{pmatrix} 2 \\ 32 \end{pmatrix} \right) = 2S_1^{(1)} + 32S_2^{(1)}$$

Из условия  $2S_1^{(1)} + 32S_2^{(1)} = 0$  находим  $S_2^{(1)} = -\frac{1}{16}$ , полагая

$S_1^{(1)} = 1$  (произвольно, так как для определения двух координат вектора есть лишь одно условие).

Вторая итерация.

$$\begin{cases} x^{(2)} = \frac{32}{17} + \lambda^{(1)}, \\ y^{(2)} = -\frac{2}{17} + \lambda^{(1)} \left(-\frac{1}{16}\right). \end{cases}$$

$$f(x^{(2)}, y^{(2)}, \lambda^{(1)}) = \left(\frac{32}{17} + \lambda^{(1)}\right)^2 + 16\left(-\frac{2}{17} - \frac{1}{16}\lambda^{(1)}\right)^2$$

$$\frac{df}{d\lambda^{(1)}} = 0, \quad \frac{32}{17} + \lambda^{(1)} + \frac{2}{17} + \frac{\lambda^{(1)}}{16} = 0, \quad \lambda^{(1)} = -\frac{32}{17}$$

$$\begin{cases} x^{(2)} = \frac{32}{17} - \frac{32}{17} = 0, \\ y^{(2)} = -\frac{2}{17} - \frac{1}{16}\left(-\frac{32}{17}\right) = 0. \end{cases}$$

$\bar{f}'(\bar{x}^{(2)}) = \begin{pmatrix} 2x \\ 32y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$  - условие окончания выполнено, следовательно за две итерации найдены координаты точки минимума  $x^* = 0$ ,  $y^* = 0$ ; значение функции  $f(x, y)$  в точке минимума  $f(x^*, y^*) = 0$ .

### 8.9. Метод Флетчера - Ривза

Сначала рассмотрим метод для квадратичной функции с положительно определенной ( $n \times n$ ) матрицей  $Q$ .

Известно, что любую квадратичную функцию можно записать в виде  $f(\bar{x}) = \frac{1}{2}(\bar{x}, Q\bar{x}) + (\bar{x}, b) + a$ .

Метод Флетчера - Ривза - метод сопряженных градиентов.

Метод использует последовательность направлений поиска минимума, каждое из которых является линейной комбинацией антиградиента в текущей точке и предыдущего направления спуска.

При минимизации  $f(\bar{x})$  методом Флетчера - Ривза векторы  $\bar{S}^{(k)}$  вычисляются [48] по формулам

$$\bar{S}^{(k)} = -\bar{f}'(\bar{x}^{(k)}) + \beta_{k-1}\bar{S}^{(k-1)}, \quad k \geq 1, \quad (24)$$

$$\bar{S}^{(0)} = -\bar{f}'(\bar{x}^{(0)}).$$

Величины  $\beta_{k-1}$  выбирают так, чтобы направления  $\bar{S}^{(k)}$  были Q-сопряженными, т.е.  $(\bar{S}^{(k)}, Q\bar{S}^{(k-1)}) = 0$ .

$$\beta_{k-1} = \frac{(\bar{f}'(\bar{x}^{(k)}), Q\bar{S}^{(k-1)})}{(\bar{S}^{(k-1)}, Q\bar{S}^{(k-1)})}. \quad (25)$$

При  $\lambda = \lambda^{(k)}$  функция  $f(\bar{x}^{(k)} + \lambda\bar{S}^{(k)})$  достигает минимума, который может быть вычислен по формуле:

$$\lambda^{(k)} = -\frac{(\bar{f}'(\bar{x}^{(k)}), \bar{S}^{(k)})}{(\bar{S}^{(k)}, Q\bar{S}^{(k)})}. \quad (26)$$

Доказана [10], [48] справедливость утверждений.

**Утверждение 3.**

$$\bar{f}'(\bar{x}^{(k+1)}) = \bar{f}'(\bar{x}^{(k)}) + \lambda^{(k)}Q\bar{S}^{(k)}. \quad (27)$$

**Утверждение 4.**

Векторы  $\bar{f}'(\bar{x}^{(k)})$  и  $\bar{f}'(\bar{x}^{(k+1)})$  ортогональны между собой.

**Утверждение 5.**

Векторы  $\bar{S}^{(0)}, \bar{S}^{(1)}, \dots, \bar{S}^{(k)}$ ,  $k \leq n$ , в методе Флетчера - Ривза являются Q-сопряженными, а градиенты  $\bar{f}'(\bar{x}^{(0)}), \bar{f}'(\bar{x}^{(1)}), \dots, \bar{f}'(\bar{x}^{(k)})$  взаимно ортогональными.

**Утверждение 6.**

Справедливо соотношение

$$\beta_{k-1} = \frac{(\bar{f}'(\bar{x}^{(k)}), \bar{f}'(\bar{x}^{(k)}))}{(\bar{f}'(\bar{x}^{(k-1)}), \bar{f}'(\bar{x}^{(k-1)}))}. \quad (28)$$

### Алгоритм метода Флетчера - Ривза

1. Задать произвольную точку  $\bar{x}^{(0)}$ , положить  $k = 0$ .
2. Вычислить в точке  $\bar{x}^{(0)}$  координаты вектора  $\bar{S}^{(0)} = -\bar{f}'(\bar{x}^{(0)})$ .
3. На  $k$ -ом шаге решить задачу минимизации по  $\lambda \geq 0$  функции  $f(\bar{x}^{(k)} + \lambda\bar{S}^{(k)})$ , для чего определить шаг  $\lambda^{(k)}$  (по формуле (26)) и вычислить координаты  $\bar{x}^{(k+1)} = \bar{x}^{(k)} + \lambda^{(k)}\bar{S}^{(k)}$ .
4. Вычислить значения  $\bar{f}'(\bar{x}^{(k+1)})$  и  $\bar{f}'(\bar{x}^{(k)})$ .
5. Если  $\bar{f}'(\bar{x}^{(k+1)}) = 0$ , то  $\bar{x}^{(k+1)}$  - решение задачи, в противном случае - определить  $\bar{S}^{(k+1)}$

$$\bar{S}^{(k+1)} = -\bar{f}'(\bar{x}^{(k+1)}) + \frac{(\bar{f}'(\bar{x}^{(k+1)}), \bar{f}'(\bar{x}^{(k+1)}))}{(\bar{f}'(\bar{x}^{(k)}), \bar{f}'(\bar{x}^{(k)}))} \bar{S}^{(k)}, \quad (29)$$

положить  $k = k + 1$  и перейти к пункту 3.

Выражение (29) можно записать в виде

$$\bar{S}^{(k+1)} = -\bar{f}'(\bar{x}^{(k+1)}) + \frac{\sum_{i=1}^n \left(\frac{\partial f(\bar{x}^{(k+1)})}{\partial x_i}\right)^2}{\sum_{i=1}^n \left(\frac{\partial f(\bar{x}^{(k)})}{\partial x_i}\right)^2} \bar{S}^{(k)}. \quad (30)$$

Для квадратичной функции  $n$  переменных методом Флетчера - Ривза точка минимума может быть найдена не более, чем за  $n$  шагов при любом начальном приближении.

**Пример 9.**

Методом Флетчера - Ривза вычислить минимум функции

$$f(x, y) = x^2 + 16y^2.$$

Решение.

$$\text{Пусть } \bar{x}^{(0)} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad f(\bar{x}^{(0)}) = 20, \quad Q(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 32 \end{pmatrix}.$$

Вычислим  $\bar{S}^{(0)} = -\bar{f}'(\bar{x}^{(0)})$ .

$$\bar{f}'(\bar{x}) = \begin{pmatrix} 2x \\ 32y \end{pmatrix}, \quad f'(\bar{x}^{(0)}) = \begin{pmatrix} 4 \\ 32 \end{pmatrix}, \quad \bar{S}^{(0)} = \begin{pmatrix} -4 \\ -32 \end{pmatrix}.$$

По формуле (26)

$$\lambda^{(0)} = - \frac{\begin{pmatrix} 4 \\ 32 \end{pmatrix} \begin{pmatrix} -4 \\ -32 \end{pmatrix}}{\begin{pmatrix} -4 \\ -32 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 32 \end{pmatrix} \begin{pmatrix} -4 \\ -32 \end{pmatrix}} = - \frac{-16 - 1024}{\begin{pmatrix} -4 \\ -32 \end{pmatrix} \begin{pmatrix} -8 \\ -1024 \end{pmatrix}} = \frac{1040}{32 + 32768} = 0.03171.$$

$$\bar{x}^{(1)} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} + 0.03171 \begin{pmatrix} -4 \\ -32 \end{pmatrix} = \begin{pmatrix} 1.87317 \\ -0.01463 \end{pmatrix}, \quad f(\bar{x}^{(1)}) = 3.51219,$$

$$\bar{f}'(\bar{x}^{(1)}) = \begin{pmatrix} 3.74634 \\ -0.46829 \end{pmatrix} \neq 0.$$

$$\bar{S}^{(1)} = \begin{pmatrix} -3.74634 \\ 0.46829 \end{pmatrix} + \frac{14.25435}{1040} \begin{pmatrix} -4 \\ -32 \end{pmatrix} = \begin{pmatrix} -3.80116 \\ 0.02969 \end{pmatrix}.$$

Переходим ко второй итерации.

$$\lambda^{(1)} = - \frac{(\bar{f}'(\bar{x}^{(1)}), \bar{S}^{(1)})}{(\bar{S}^{(1)}, Q\bar{S}^{(1)})} = 0.49279, \quad f(\bar{x}^{(2)}) = 0.$$

$$\bar{f}'(\bar{x}^{(2)}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \text{следовательно } \bar{x}^* = \bar{x}^{(2)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ - точка минимума,}$$

$$f(\bar{x}^*) = 0.$$

Заметим, что при использовании градиентных методов для решения задачи оптимизации этой функции наблюдается "овражный эффект"!

Минимизация неквадратичных функции  $n$  - переменных может быть проведена модифицированным методом Флетчера - Ривза [48].

**Алгоритм метода Флетчера - Ривза минимизации неквадратичной функции  $n$  переменных**

1. Задать начальную точку  $\bar{x}^{(0)}$ , вычислить  $\bar{S}^{(0)} = -\bar{f}'(\bar{x}^{(0)})$ .
2. Вычислить  $\bar{x}^{(k+1)} = \bar{x}^{(k)} + \lambda^{(k)} \bar{S}^{(k)}$ ,

$$(31)$$

$$\bar{S}^{(k)} = -\bar{f}'(\bar{x}^{(k)}) + \beta_{k-1} \bar{S}^{(k-1)}, \quad k \geq 1.$$

$$f(\bar{x}^{(k)} + \lambda^{(k)} \bar{S}^{(k)}) = \min f(\bar{x}^{(k)} + \lambda^{(k)} \bar{S}^{(k)}).$$

$$\beta_{k-1} = \begin{cases} \frac{(\bar{f}'(\bar{x}^{(k)}), \bar{f}'(\bar{x}^{(k)}) - \bar{f}'(\bar{x}^{(k-1)}))}{(\bar{f}'(\bar{x}^{(k-1)}), \bar{f}'(\bar{x}^{(k-1)}))}, & k \notin I, \\ 0, & k \in I, \end{cases}$$

$I = 0, n, 2n, \dots$  - множество индексов.

После  $n$  шагов итерационная процедура 1-2 повторяется с заменой  $\bar{x}^{(k)}$  на  $\bar{x}^{(n+k)}$ .

Условие окончания вычисления минимума функции

$$\|\bar{f}'(\bar{x}^{(k)})\| < E. \quad (32)$$

Скорость сходимости метода Флетчера - Ривза - квадратичная и может быть оценена при достаточно жестких предположениях относительно свойств функции  $f(\bar{x})$  [78].

$$\|x_{k+n} - x^*\| \leq q \|x_k - x^*\|^2, \quad k \in I, k \geq N. \quad (33)$$

**8.10. Метод Ньютона**

Метод Ньютона - метод второго порядка; при построении последовательности приближений к точке  $\bar{x}^*$  - минимуму функции  $f(x_1, x_2, \dots, x_n)$ , используют информацию не только о первых, но и о вторых производных функции  $f(\bar{x})$ . При этом  $f(\bar{x})$  должна быть выпуклой, дважды дифференцируемой функцией.

**Структурная схема нахождения точки минимума выпуклой квадратичной функции методом Ньютона**

Входные параметры:

$\bar{x}^0$  - вектор начальных приближений.

$E$  - точность

Выходные параметры:

$\bar{x}^*$  - вектор оптимального решения.

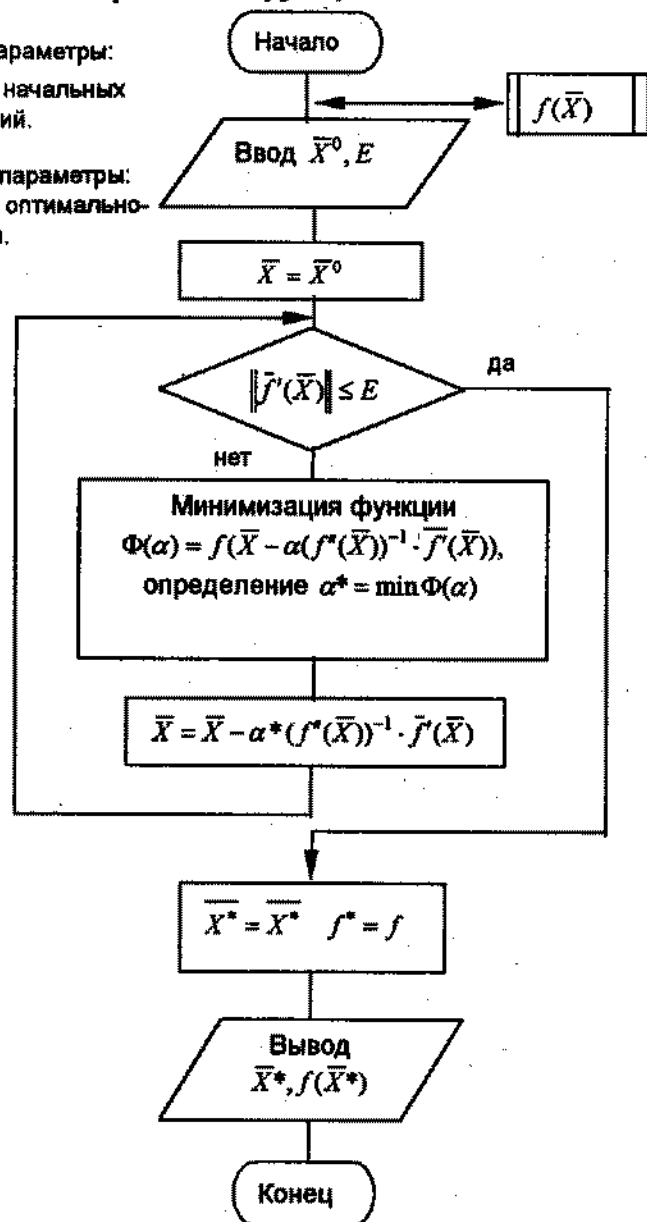


Рис. 8.11

Метод Ньютона используют, как правило, для уточнения решения, найденного более простым способом. Рекуррентная формула метода:

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} - [f''(\bar{x}^{(k)})]^{-1} \cdot \bar{f}'(\bar{x}^{(k)}), \quad k=0,1,\dots \quad (34)$$

$\bar{x}^{(0)}$  - начальное приближение,  $[f''(\bar{x}^{(k)})]^{-1}$  - матрица, обратная матрице вторых производных.

Метод сходится к  $\bar{x}^*$ , если начальное приближение  $\bar{x}^{(0)}$  достаточно близко к  $\bar{x}^*$ . Существуют различные модификации метода Ньютона [66], в частности метод Ньютона - Рафсона с выбором шага  $\alpha_k$ .

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} - \alpha_k [f''(\bar{x}^{(k)})]^{-1} \bar{f}'(\bar{x}^{(k)}), \quad k=0,1,\dots \quad (35)$$

Из условия  $\Phi_k(\alpha_k) = \min \Phi_k(\alpha)$  определяется  $\alpha_k$ .

$$\Phi_k(\alpha) = f(\bar{x}^{(k)} - \alpha_k [f''(\bar{x}^{(k)})]^{-1} \cdot \bar{f}'(\bar{x}^{(k)})), \quad (36)$$

$$0 \leq \alpha_k \leq 1.$$

Схема (35) - (36) дает сходящейся к  $\bar{x}^*$  процесс (за счет выбора  $\alpha_k$ ), тогда как метод (34) может не сходить (или сходить очень медленно) даже для выпуклых, дважды дифференцируемых функций. Заметим также, что для (35) на некоторых итерациях может нарушаться условие  $f(\bar{x}^{(k+1)}) \leq f(\bar{x}^{(k)})$ .

Схема алгоритма представлена на рис. 11.

Рассмотрим несколько другой подход.

Необходимое условие экстремума приводит к решению системы уравнений (37).

$$\frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_m} = 0, \quad m=1,2,\dots,n. \quad (37)$$

При тех же требованиях к  $f(\bar{x})$  предлагается искать приближение к  $\bar{x}^*$ :

$$\bar{x}_j^{(k+1)} = \bar{x}_j^{(k)} + \bar{y}_j, \quad j=1, n, \quad (38)$$

$\bar{y} = (y_1, y_2, \dots, y_n)$  образуют решение системы  $n$  линейных алгебраических уравнений





$H_i, \text{grad } f(\bar{x}), H_i$  - положительно определенная симметричная матрица.

Матрица  $H_i$  обновляется на каждом шаге и в пределе становится равной обратному гессиану.

На шаге  $i = 0$   $H_0$  - единичная матрица,  $\bar{x}_0$  - начальная точка.

Пусть  $\bar{g}_i = \bar{g}(\bar{x}_i) = \text{grad}f(\bar{x}_i)$ .

При нахождении  $\lambda_i$  для минимизации  $f(\bar{x}_i - \lambda_i H_i \bar{g}_i)$  производится одномерный поиск методом кубической интерполяции.

Теоретическое обоснование метода ДФП приведено в [4].

### Алгоритм метода ДФП

1. Задать начальную точку  $\bar{x}_0$  и матрицу  $H_0$ . Положить  $i = 0$ , задать точность.

2. Определить направление поиска

$$\bar{d}_i = -H_i \bar{g}_i. \quad (44)$$

3. Произвести одномерный поиск минимума функции  $f(\bar{x}_i + \lambda_i \bar{d}_i)$  вдоль прямой  $\bar{x}_i + \lambda \bar{d}_i$ .

$$4. \text{ Положить } \bar{v}_i = \lambda_i \bar{d}_i. \quad (45)$$

5. Положить  $\bar{x}_{i+1} = \bar{x}_i + \bar{v}_i$ .

6. Найти  $f(\bar{x}_{i+1})$  и  $\bar{g}_{i+1}$ .

$$7. \text{ Если } |\bar{g}_{i+1}| < E \text{ или } |\bar{v}_i| < E, \quad (46)$$

процедуру завершить, иначе - переход к пункту 8.

$$8. \text{ Положить } \bar{U}_i = \bar{g}_{i+1} - \bar{g}_i. \quad (47)$$

9. Обновить матрицу  $H$  по правилу:

$$H_{i+1} = H_i + A_i + B_i, \quad (48)$$

$$A_i = \frac{\bar{v}_i \cdot \bar{v}_i^T}{\bar{v}_i^T \cdot \bar{U}_i}, \quad (49)$$

$$B_i = \frac{-H_i \bar{U}_i \bar{U}_i^T H_i}{\bar{U}_i^T H_i \bar{U}_i}. \quad (50)$$

10. Увеличить  $i$  на единицу и перейти к п. 2.

### Замечания.

- Одномерный поиск методом кубической интерполяции рекомендуется завершить после того, как будет получено улучшенное значение функции.

- Алгоритм поиска предполагает наличие процедур вычисления рассматриваемой функции и ее производных.

### Структурная схема алгоритма метода Ньютона

Входные параметры:

$\bar{X}^0$  - вектор начальных приближений.

Выходные параметры:

$\bar{X}^*$  - вектор оптимального решения.



Рис. 8.12

## 9. ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

### 9.1. Постановка задачи

Дана система  $m$  линейных уравнений с  $n$  неизвестными

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots \quad \dots \quad \dots \quad \dots \quad \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m. \end{cases} \quad (1)$$

неизвестные,

$$x_i \geq 0 \quad (2)$$

и линейная функция

$$F = c_1x_1 + c_2x_2 + \dots + c_nx_n, \quad i = \overline{1, n}. \quad (3)$$

Требуется среди всех решений системы уравнений (1) найти такое неотрицательное решение, при котором линейная форма (3) принимает наименьшее значение. Система (1) называется системой ограничений, причем условия (2) также можно считать ограничениями.

Постановку задачи в форме (1-3) называют основной формой ЗЛП (задачи линейного программирования). Функция  $F$  - линейная функция - называется критерием оптимальности, целевой функцией.

Система ограничений (1), (2) - линейные функции своих аргументов. Матрица системы (1) называется матрицей условий, ее столбцы - векторами условий.

Всякое неотрицательное решение (1) называется допустимым. Допустимое решение, оптимизирующее линейную форму  $F$  называется оптимальным. Заметим, что поставленная задача может

- не иметь решений,
- иметь единственное решение,
- иметь бесчисленное множество решений.

Задача может иметь решение, если система (1) совместна. Если при этом система имеет единственное решение, то и задача линейного программирования имеет единственное решение, если это решение неотрицательно (допустимо) и не имеет решения, если единственное решение не является допустимым.

Ограничимся рассмотрением задач, для которых все уравнения (1) являются линейно - независимыми, т.е. число уравнений ( $m$ ) системы ограничений равно рангу  $r$  матрицы системы ограничений ( $r = m$ ). В этом случае система (1) имеет более одного решения.

Помимо указанной постановки задача линейного программирования в виде (1-3) возможны и другие формулировки. Например, в случае, если требуется найти решение (1-2), доставляющее ее линейной форме  $F$  наибольшее значение, достаточно решить задачу минимизации функции  $F_1 = -F$ .

Допустим теперь, что среди ограничений задачи помимо уравнений имеются неравенства вида

$$\alpha_1x_1 + \alpha_2x_2 + \dots + \alpha_nx_n \geq 0. \quad (4)$$

Введем новую (дополнительную) переменную

$$x_{n+1} = \alpha_1x_1 + \alpha_2x_2 + \dots + \alpha_nx_n.$$

Тогда требование  $x_{n+1} \geq 0$  эквивалентно условию (4). Таким образом, все неравенства заменяются ограничениями - равенствами. При этом увеличивается число неизвестных.

В математических моделях некоторых задач какие-то переменные по своему физическому смыслу могут принимать и отрицательные значения. В этом случае для такой переменной  $x_j$  следует сделать замену

$$x_j = x'_j - x''_j, \quad (5)$$

$x'_j, x''_j$  - неотрицательные переменные.

В некоторых случаях вместо системы линейных уравнений (1) целесообразнее иметь систему линейных неравенств. Чтобы перейти от системы линейных уравнений к системе линейных неравенств, достаточно заметить, что уравнение

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i$$

равносильно системе двух неравенств

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \leq b_i, \\ a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n > b_i, \end{cases} \quad (6)$$

или

$$\begin{cases} a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \leq b_i, \\ -(a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n) < -b_i. \end{cases} \quad (6')$$

В общем случае задачу с ограниченными — неравенствами запишем так:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b_2, \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m, \\ x_i \geq 0, \\ F = c_1x_1 + c_2x_2 + \dots + c_nx_n, \quad i = \overline{1, n}. \end{cases} \quad (7)$$

## 9.2. Геометрическая иллюстрация задачи линейного программирования с ограничениями — неравенствами

Для решения задачи линейного программирования разработаны различные методы. Остановимся на геометрической иллюстрации основной задачи с системой ограничений вида (7), которая возможна, если  $n-m=2$  или  $n-m=3$ .

Известно, что геометрическое место точек, координаты которых удовлетворяют системе линейных неравенств, образует при  $n-m=2$  выпуклый, возможно и не замкнутый многоугольник (симплекс), называемый многоугольником решений. Оптимальное решение, если оно существует и единственно, достигается в вершине симплекса; если таких решений бесчисленное множество — во всех точках некоторой стороны (в том числе, в ограничивающих эту сторону вершинах).

Графическое решение состоит из следующих шагов:

### I. Построение области допустимых решений (О.Д.Р.)

1) Строят прямые, уравнения которых получаются в результате замены в системе ограничений (7) знаков неравенств на знаки точных равенств.

2) Находят полуплоскости, определяемые каждым из ограничений задачи.

3) Находят многоугольник решений.

### II. Определение оптимального решения

1) Строят вектор  $\overline{\text{grad}} F = (c_1, c_2)$ .

2) Строят прямую  $c_1x_1 + c_2x_2 = c = \text{const}$ , проходящую через многоугольник решений.

3) Перемещают прямую  $c_1x_1 + c_2x_2 = c$  в направлении, противоположном направлению градиента (антиградиент). В результате чего либо находят точку (точки), в которой целевая функция принимает минимальное значение, либо устанавливают неограниченность снизу целевой функции.

4) Определяют координаты точки минимума функции  $F$  и вычисляют значение целевой функции в этой точке.

### Пример 1.

Найти графическое решение ЗЛП

$$\begin{cases} 3x_2 \leq 15, \\ 2x_1 + 3x_2 \leq 19, \\ 2x_1 + x_2 \leq 13, \\ 3x_1 \leq 18, \end{cases} \quad (*)$$

$$x_1 \geq 0,$$

$$x_2 \geq 0.$$

$$F = -7x_1 - 5x_2 \rightarrow \min.$$

Решение.

Построим ОДР — область допустимых решений. Строим прямые

$$\begin{cases} 3x_2 = 15, \\ 2x_1 + 3x_2 = 19, \\ 2x_1 + 3x_2 = 13, \\ x_1 = 6, \end{cases} \quad (6)$$

каждая из которых делит координатную плоскость на две полуплоскости. Выбираем ту полуплоскость, где выполнены условия (\*). В результате получаем выпуклый многоугольник OPQRST решений (рис. 1).

Ищем точку (точки) ОДР, где функция  $F$  достигает наименьшего значения. Уравнение  $-7x_1 - 5x_2 = c$  определяет при каждом  $c$  некоторую линию уровня. Вектор  $\overline{\text{grad}} F = -7i - 5j$  указывает направление возрастания функции  $F$ .

Точки А и В принадлежат ОДР. Прямая АВ — одна из линий уровня. Перемещаем прямую АВ параллельно самой себе в направлении антиградиента (что обеспечивает убывание функции  $F$ ) до тех пор, пока прямая, параллельная АВ будет иметь хотя бы одну общую точку с ОДР. Координаты точки можно найти, решая совместно уравнения прямых, пересечением которых она является.

$$\begin{cases} 2x_1 + 3x_2 = 19, \\ 2x_1 + x_2 = 13, \end{cases} \quad \text{откуда } x_1 = 5, \quad x_2 = 3.$$

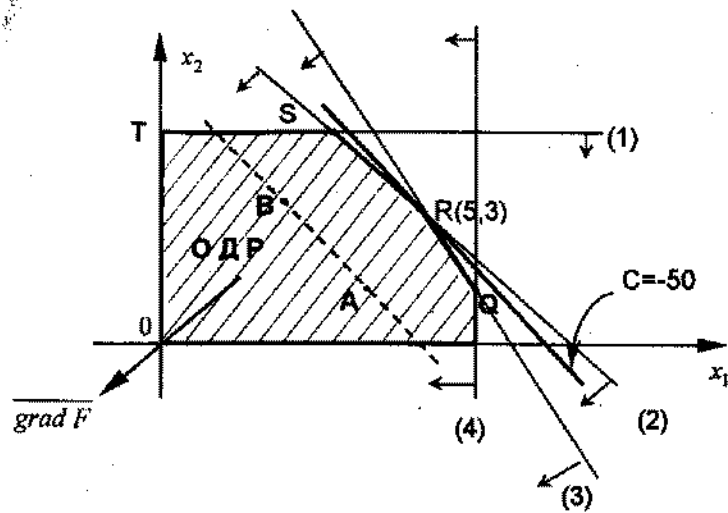


Рис. 9.1

Подставляя полученные координаты в выражение  $F$ , получим наименьшее значение линейной формы.

Итак, оптимальное решение  $x_1 = 5$ ,  $x_2 = 3$ ,  $F_{\min} = -50$ .

**Замечания.**

Если решение не единственное, то предельное положение прямой АВ совпадет с одной из сторон симплекса для замкнутой области ОДР.

В случае, если ОДР - незамкнутая область, ЗЛП может и не иметь решения.

Если при тех же ограничениях линейная форма  $F = 8 - x_1 - x_2$ , то оптимальных решений - бесчисленное множество - любая точка стороны SR и, в том числе, в ограничивающие эту сторону вершины многоугольника допустимых решений. Во всех точках минимальное значение равно 0, т.к. при  $x_1 = 5$ ,  $x_2 = 3$ ,  $F = 8 - x_1 - x_2 = 0$ .

**Пример 2**

Минимизировать функцию  $F = 4 - x_1 - x_2$  при ограничениях

$$x_1 \geq 0, \quad x_2 \geq 0,$$

$$\begin{cases} x_1 - 2x_2 \leq 2 & (1), \\ x_1 + x_2 \geq 2 & (2). \end{cases}$$

Решение.

Строим прямые  $x_1 = 0$ ,  $x_2 = 0$ ,  $x_1 - 2x_2 = 2$ ,  $x_1 + x_2 = 2$ .

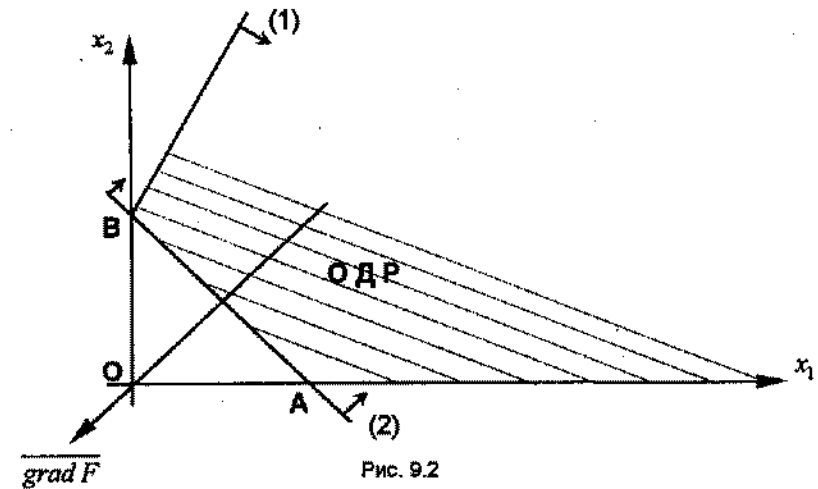


Рис. 9.2

Допустимая область не ограничена в направлении убывания линейной формы  $F$ , т.е. не существует точки, в которой  $F$  достигла бы минимального значения. Если бы требовалась найти максимальное значение функции  $F$ , то в этом случае задача имела бы множество (отрезок АВ) решений в неограниченной области допустимых решений.

Сформулируем ряд утверждений линейного программирования [36]

- Допустимая область является выпуклым множеством.
- Допустимое решение называется базисным, если его координаты можно разбить на две группы так, чтобы в первую группу попали ровно  $m$  координат и соответствующие им вектора условий были бы линейно - независимы, а вторая группа состояла из нулей. Список координат первой группы принято называть допустимым базисом.
- Если ограничения имеют допустимое решение, то они имеют и базисное решение.
- Базисные допустимые решения - суть вершины выпуклого множества.
- Если целевая функция имеет конечный минимум, то по крайней мере одно оптимальное решение является базисным.
- У задачи линейного программирования может быть не более  $C_n^m = \frac{n!}{m!(n-m)!}$  допустимых базисных решений, среди них есть оптимальные. При поиске оптимального решения в допустимой области можно ограничиться допустимыми базисными решениями.

### 9.3. Симплексный метод решения ЗЛП

Решение ЗЛП - определение координат той вершины симплекса, в которой функция  $F$  достигает минимума. Это решение может быть найдено перебором всех вершин многоугольника О.Д.Р. С ростом числа переменных ( $n$ ) число вершин резко возрастает. Кроме того, требуется в системе ограничений выделять линейно-независимые уравнения.

Сущность симплексного метода состоит в том, что отправляясь из некоторой вершины переходят к рассмотрению последовательности только тех вершин, в которых значения целевой функции меньше, чем в предыдущей вершине.

Таким образом производится упорядоченный частичный перебор вершин, сопровождающийся уменьшением значения целевой функции до тех пор, пока не останется вершин, где функция  $F$  убывает.

На каждой итерации устанавливается один из фактов:

- базисное решение оптимально,
- линейная функция  $F$  не ограничена снизу,
- возможно уменьшение  $F$ , т.е. переход к новому базису.

Предполагается, что найден начальный допустимый базис.

Запишем систему (1) в виде

$$\begin{cases} x_1 + \dots + a'_{1,m+1}x_{m+1} + a'_{1,m+2}x_{m+2} + \dots + a'_{1n}x_n = b'_1, \\ x_2 + \dots + a'_{2,m+1}x_{m+1} + \dots + a'_{2n}x_n = b'_2, \\ \dots \\ x_m + a'_{m,m+1}x_{m+1} + \dots + a'_{mn}x_n = b'_m. \end{cases} \quad (6)$$

Для исключения переменных  $x_1, x_2, \dots, x_m$  из (3) следует умножить ограничения системы (6) на  $c_1, c_2, \dots, c_m$  и вычесть из (3), тогда

$$\begin{aligned} c'_{m+1}x_{m+1} + c'_{m+2}x_{m+2} + \dots + c'_n x_n &= \alpha - \alpha_0, \\ F_0 &= \sum_{i=1}^m c_i b'_i. \end{aligned} \quad (7)$$

Соотношения (6), (7) являются канонической формой для базиса  $x_1, x_2, \dots, x_m$ .

Базисное решение

$$x_1 = b'_1, \quad x_2 = b'_2, \dots, \quad x_m = b'_m, \quad x_{m+1} = 0, \quad x_{m+2} = 0, \dots, \quad x_n = 0$$

при  $b'_i \geq 0$  ( $i = \overline{1, m}$ ) является допустимым.

Если в выражении (7) все коэффициенты  $c'_s \geq 0$  ( $s = \overline{m+1, n}$ ), то значение функции  $F$  не может быть уменьшено, т.е. минимум найден.

Итерационный процесс симплекс-метода состоит из трех шагов.

1. Поиск переменной для включения в базис.

Среди отрицательных коэффициентов линейной формы (7) выбирается наибольший по модулю  $c'_s$ . Переменная  $x_s$  должна быть включена в базис.

2. Поиск переменной для исключения из базиса.

В системе ограничений (6) среди переменных с положительными коэффициентами  $a'_{is}$  выбирают базисную переменную принадлежащую строке с номером  $r$ . Если  $r$  - номер строки, в которой достигается

$$\max x_s = \min \left( \frac{b'_i}{a'_{is}} \right), \quad i = \overline{1, m}, \quad a'_{is} > 0, \quad (8)$$

то переменная  $x_r$  исключается из базиса.

Элемент  $a'_{rs}$  называется ведущим, строка  $r$  - ведущая строка, столбец  $s$  - ведущий столбец.

3. Построение новой канонической формы.

Переменную  $x_s$  следует исключить из ограничений (6) и из целевой функции (7).

а) Образует новую ведущую строку, поделив коэффициенты строки на  $a'_{rs}$ .

б) Удалим  $x_s$  из ограничений. Для этого из  $i$ -ой строки ( $i \neq r$ ) с коэффициентом  $a'_{is}$  при  $x_s$  вычтем произведение  $a'_{is}$  и новой ведущей строки.

в) Для преобразования целевой функции с коэффициентом  $c'_s < 0$  при  $x_s$  вычтем из строки целевой функции произведение  $c'_s$  и новой ведущей строки.

Исключение переменной также может быть произведено по правилам (28) раздела "Системы линейных уравнений":

$$b_r = \frac{b'_r}{a'_{rs}}, \quad a_{ij} = \frac{a'_{ij}}{a'_{rs}}, \quad i=r, \quad b_i = b'_i - a'_{is} \cdot b_r, \quad i \neq r, \quad (9)$$

$$F_{ij} = a'_{ij} - a'_{is} \cdot a_{ij}, \quad c_j = c'_j - c'_s \cdot a_{ij}, \quad F = F_0 + c'_s \cdot b_r.$$

**Замечания.**

- Поиск переменной  $x_s$  становится невозможным, если среди коэффициентов линейной формы нет отрицательных. Это соответствует достижению минимума линейной формы  $F$ .
- В случае, если все коэффициенты  $a_{is} \leq 0$ , нельзя найти переменную  $x_r$ . В симплекс - методе условие  $a'_{is} \leq 0$  – условие неограниченности функции  $F$ .
- Наличие нулевых коэффициентов при базисных переменных в выражении целевой функции свидетельствует о неединственности оптимального решения.

Вычислительный процесс удобнее вести, записывая условия задачи и первоначальные данные, полученные после определения начального базиса, в таблицу.

Столбец  $\bar{C}$  содержит положительные коэффициенты при неизвестных целевой функции, имеющие те же индексы, что и базисные переменные.

В столбце  $b$  записывают положительные компоненты начального базиса, на следующих итерациях – полученные положительные компоненты оптимального решения.

В последней строке – значение целевой функции, которое она получает при данном базисном решении и коэффициенты при неизвестных целевой функции.

Последние  $n$  столбцов содержат коэффициенты системы ограничений, получившейся после определения начального базиса.

Таблица 1

$\bar{C}$	Б	$b$	$c_1$ $x_1$	$c_2$ $x_2$	...	$c_m$ $x_m$	$c_{m+1}$ $x_{m+1}$	...	$c_n$ $x_n$
$c_1$	$x_1$	$b_1$	1	0	...	0	$a_{1,m+1}$	...	$a_{1,n}$
$c_2$	$x_2$	$b_2$	0	1	...	0	$a_{2,m+1}$	...	$a_{2,n}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$c_m$	$x_m$	$b_m$	0	0	...	1	$a_{m,m+1}$	...	$a_{m,n}$
	$F$								

**Пример 3.**

Минимизировать функцию

$F = -6x_1 - 2x_2$  при ограничениях

$$\begin{cases} x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, \\ 2x_1 + 4x_2 + x_3 = 9, \\ 3x_1 + x_2 + x_4 = 6. \end{cases}$$

**Решение.**

Пусть  $x_3, x_4$  – базисные переменные.

Наибольший по модулю отрицательный коэффициент в целевой функции  $F$  стоит перед переменной  $x_1$ , поэтому полагаем

$x_5 = x_1$ .  $\max x_1 = \min_{a_n > 0} \left( \frac{9}{2}, \frac{6}{3} \right)$  находится в строке с базисной переменной

$x_4$ , следовательно  $x_r = x_4$ . Итак, включаем в базис переменную  $x_1$ , исключаем из базиса  $x_4$ , новые базисные переменные  $x_1$  и  $x_3$ .

Получим новую каноническую форму. Ведущая строка ( $a'_{rs} = a'_{41} = 3$ )  $x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_4 = 2$ , исключаем  $x_1$  из ограничений.

$$\begin{cases} 2x_1 + 4x_2 + x_3 - 2\left(x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_4\right) = 9 - 2 \cdot 2, \\ -6x_1 - 2x_2 - (-6)\left(x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_4\right) = 0 - (-6) \cdot 2. \end{cases}$$

или

$$\begin{cases} \frac{10}{3}x_2 + x_3 - \frac{2}{3}x_4 = 5, \\ x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_4 = 2, \end{cases} \quad F = 0 \cdot x_2 + 2x_4 + 12.$$

Переменная  $x_2$ , не являясь базисной, входит в целевую функцию с коэффициентом ноль, что свидетельствует о том, что базисное решение не единственное.

Включим  $x_2$  в число базисных. Новая ведущая строка (делим коэффициенты 1-го уравнения на  $\frac{10}{3}$ )  $x_2 + \frac{3}{10}x_3 - \frac{2}{10}x_4 = \frac{3}{2}$ .

Так как  $\min\left(\frac{5 \cdot 3}{10}, \frac{2 \cdot 3}{1}\right) = \frac{3}{2}$  соответствует  $r = 3$ , исключим  $x_3$  из числа базисных переменных. Удалим  $x_2$  из ограничений.

$$x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_4 - \frac{1}{3}\left(x_2 + \frac{3}{10}x_3 - \frac{2}{10}x_4\right) = 2 - \frac{1}{3}\left(\frac{3}{2}\right).$$

Система ограничений принимает вид:

$$\begin{cases} x_1 - \frac{1}{10}x_3 + \frac{4}{10}x_4 = \frac{3}{2}, \\ x_2 + \frac{3}{10}x_3 - \frac{2}{10}x_4 = \frac{3}{2}, \end{cases}$$

$$F = 12 + 2x_4.$$

Дальнейшее уменьшение  $F$  невозможно, так как нет переменных с отрицательными коэффициентами.

Таблица 2

$\bar{C}$	Б	$b$	-6	-2		
			$x_1$	$x_2$	$x_3$	$x_4$
	$x_3$	9	2	4	1	0
	$x_4$	6	3	1	0	1
	$-F$	0	-6	-2	0	0
6	$x_3$	5	0	$\frac{10}{3}$	1	$-\frac{2}{3}$
	$x_1$	2	1	$\frac{1}{3}$	0	$\frac{1}{3}$
	$-F$	12	0	$\bar{0}$	0	2
6	$x_1$	$\frac{3}{2}$	0	1	$\frac{3}{10}$	$-\frac{2}{10}$
	$x_2$	$\frac{3}{2}$	1	0	$-\frac{1}{10}$	$\frac{4}{10}$
	$-F$	12	0	0	$\bar{0}$	2

Итак, минимальное значение линейной формы  $F = -12$ , оптимальный базис  $\bar{X}^{opt} = \left\{ \frac{3}{2}; \frac{3}{2}; 0; 0 \right\}$ .

#### 9.4. Определение начального базисного решения

Система ограничений задана в виде неравенств

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b_2, \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m, \end{cases} \quad (10)$$

$$\text{причем } b_j \geq 0, \quad j = \overline{1, m}. \quad (11)$$

Введем в систему дополнительные переменные  $x_{n+1}, x_{n+2}, \dots, x_{n+m}$ , которые в линейную форму войдут с коэффициентами, равными  $M$  ( $M$  - большое положительное число), а в систему ограничений - с коэффициентами 1.

Получим систему  $m$  уравнений с  $(n+m)$  неизвестными (12):

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + x_{n+1} = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n + x_{n+2} = b_2, \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n + x_{n+m} = b_m. \end{cases}$$

Дополнительные переменные принимаем за базисные, переменные  $x_1, x_2, \dots, x_n$  - свободные. Полагаем свободные неизвестные равными нулю, тогда получаем начальное базисное решение

$$\bar{X}^{(0)} = \left\{ \underbrace{0; 0; \dots; 0}_n; b_1; b_2; \dots; b_m \right\}. \quad (13)$$

Число вершин симплекса увеличилось за счет введения новых  $m$  переменных, однако этот прием позволил достаточно легко получить начальный базис.

В общем случае ограничения-неравенства могут иметь вид:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \geq b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \geq b_2, \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \geq b_m. \end{cases} \quad (14)$$

$$b_j \geq 0, \quad j = \overline{1, m}. \quad (15)$$

По-прежнему введем дополнительные переменные  $x_{n+1}, \dots, x_{n+m}$ , получим систему (16):

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n - x_{n+1} = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n - x_{n+2} = b_2, \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n - x_{n+m} = b_m. \end{cases}$$

Если, по аналогии с предыдущим, принять переменные  $x_{n+1}, x_{n+2}, \dots, x_{n+m}$  за базисные, то полученное базисное решение не будет допустимым, так как некоторые из базисных переменных от-

рицательны при нулевых значениях свободных переменных  $x_1, x_2, \dots, x_n$ .

Один из возможных путей преодоления трудностей, возникающих в этом случае, состоит в использовании симплекс-метода для получения допустимого базиса.

### 9.5. Алгоритм получения начального допустимого базиса

1. Изменить ограничения, приводящие к недопустимому базису, путем введения в эти ограничения искусственных переменных. В ограничения, имеющие допустимые базисные переменные, вводить искусственные переменные не следует. Таким образом получим начальный допустимый базис.

2. Сформировать дополнительную линейную форму  $S$ , имеющую смысл суммы введенных искусственных переменных.  $S$  - искусственная целевая функция.

3. Минимизировать симплекс-методом функцию  $S$ :

а) из функции  $S$  исключить искусственные переменные, для чего из  $S$  следует вычесть строки системы ограничений, содержащие эти переменные.

б) к целевой функции  $F$  применить те же преобразования, что и к ограничениям.

На практике вместо двух функций  $F$  и  $S$  часто рассматривают одну целевую функцию

$$F1 = c_1x_1 + c_2x_2 + \dots + c_nx_n + M(x_{n+1} + x_{n+2} + \dots + x_{n+m}), \quad (17)$$

$M$  - достаточно большое положительное число (такое, чтобы в ходе решения исключить искусственные переменные из рассмотрения).

Если система ограничений (16) имеет допустимое решение и, следовательно, базисное допустимое решение, то в результате минимизации функция  $S$  обратится в нуль (вместе со всеми искусственными переменными).

Базисное допустимое решение, минимизирующее функцию  $S$ , может быть использовано как начальное допустимое базисное решение для минимизации функции  $F$ .

#### Замечания.

• По завершении минимизации функции  $S$  целевая функция  $F$  уже приведена к начальному базису.

• Искусственную переменную, выведенную из числа базисных, не имеет смысла вводить в дальнейшем ни в один базис, и, следовательно, ее можно исключить из рассмотрения.

• Если функция  $S$  обратилась в нуль, дальнейшее решение проводится по правилам симплекс-метода.

• В результате минимизации функций  $S$  может оказаться не равной нулю. Этот случай соответствует отсутствию допустимого базисного решения исходной задачи, следовательно, и отсутствию минимального значения целевой функции  $F$ .

Таким же образом решается задача, если ограничения заданы в виде уравнений или/и в смешанном виде (содержат и уравнения, и неравенства).

#### Пример 4.

Минимизировать функцию

$$F = 2x_1 + 6x_2 - 5x_3 + x_4 + 4x_5$$

при ограничениях

$$\begin{cases} x_1 - 4x_2 + 2x_3 - 5x_4 + 9x_5 = 3, \\ x_2 - 3x_3 + 4x_4 - 5x_5 = 6, \\ x_2 - x_3 + x_4 - x_5 = 1, \end{cases}$$

$$x_j \geq 0, \quad j = \overline{1,5}.$$

Решение.

Для получения допустимого начального базиса преобразуем систему ограничений, учитывая, что в первом уравнении есть базисная переменная  $x_1$ . Вводим искусственные переменные  $x_6, x_7$ .

$$\begin{cases} x_1 - x_2 + 2x_3 - 5x_4 + 9x_5 & = 3, \\ x_2 - 3x_3 + 4x_4 - 5x_5 + x_6 & = 6, \\ x_2 - x_3 + x_4 - x_5 + x_7 & = 1, \end{cases}$$

$$x_j \geq 0, \quad j = \overline{1,7}.$$

$$F - 2x_1 - 6x_2 + 5x_3 - x_4 - 4x_5 = 0.$$

Полученный базис  $\bar{X}^0 = \{3; 0; 0; 0; 0; 6; 1\}$  (при нулевых значениях свободных переменных  $x_2, x_3, x_4, x_5$ ) является допустимым (неотрицательным) базисом.

Искусственная линейная форма  $S = x_6 + x_7$ .

Исключаем базисные переменные  $x_1, x_6, x_7$  из функций  $F$  и  $S$ .

Вычитая строки системы ограничений, содержащие эти переменные, из выражений  $F$  и  $S$ , получим



$$F = 6 - 14x_2 + 9x_3 - 11x_4 + 14x_5,$$

$$S = 7 - 2x_2 + 4x_3 - 5x_4 + 6x_5.$$

**Замечания.**

- При указанной форме записи линейная функция  $F$  оптимальна, если все коэффициенты при неизвестных отрицательны.
- Разрешающая переменная соответствует наибольшему по модулю отрицательному значению среди коэффициентов линейной форма  $S$ .

Первая итерация.

В форме  $S$  наибольший по модулю отрицательный коэффициент стоит при переменной  $x_4$ , следовательно  $x_5 = x_4$  - переменная, которую следует ввести в базис. Из базиса выводим  $x_7 = x_7$ , так как в столбце положительных коэффициентов при переменной  $x_4$   $\min_{\substack{a_{i4} > 0, \\ i=1,3}} \frac{b_i}{a_{i4}} = 1$  соответствует  $r=3$ , строке, содержащей базисную

переменную  $x_7$ . Разрешающая строка - третья.

По изложенным выше правилам работы по симплекс-методу переходим к базисным переменным  $x_1, x_6, x_4$ , затем получаем новую каноническую форму

$$\begin{cases} x_1 + x_2 - 3x_3 + 4x_5 = 8, \\ -3x_2 + x_3 - x_5 + x_6 = 2, \\ x_2 - x_3 + x_4 - x_5 = 1, \end{cases}$$

$$F = 17 - 3x_2 - 2x_3 + 3x_5,$$

$$S = 2 + 3x_2 - x_3 + x_5.$$

Вторая итерация.

Так как  $S$  не оптимальна, вводим в базис переменную  $x_3$ , из базиса выводим переменную  $x_6$  ( $x_1, x_3, x_4$  - базисные).

Получаем систему ограничений

$$\begin{cases} x_1 - 8x_2 + x_5 = 14, \\ -3x_2 + x_3 - x_5 = 2, \\ -2x_2 + x_4 - 2x_5 = 3 \end{cases}$$

$$F = 21 - 9x_2 + x_5, \quad S = 0.$$

Искусственная линейная форма минимизирована и равна нулю.

Третья итерация.

Линейная форма  $F$  не оптимальна ( $c_5 = 1$ ). Переменную  $x_5$  вводим в базис, выводим из базиса -  $x_1, (x_5, x_3, x_4$  - базисные).

Каноническая форма

$$\begin{cases} x_1 - 8x_2 + x_5 = 14, \\ x_1 - 11x_2 + x_3 = 16, \\ 2x_1 - 18x_2 + x_4 = 31, \\ F = 7 - x_1 - x_2. \end{cases}$$

Приведем симплекс-таблицу, иллюстрирующую решение задачи.

Таблица 3

Б	В	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$
$x_1$	3	1	-4	2	-5	9		
$x_6$	6		1	-3	4	-5	1	
$x_7$	1		1	-1	1	-1		1
$F$	6		-14	9	-11	14		
$S$	7		-2	4	-5	6		
$x_1$	8	1	1	-3		4		
$x_6$	2		-3	1		-1	1	
$x_4$	1		1	-1	1	-1		
$F$	17		-3	-2		3		
$S$	2		3	-1		1		
$x_1$	14	1	-8			1		
$x_3$	2		-3	1		-1		
$x_4$	3		-2		1	-2		
$F$	21		-9			1		
$S$	0		0			0		
$x_5$	14	1	-8			1		
$x_3$	16	1	-11	1				
$x_4$	31	2	-18		1			
$F$	7	-1	-1					

$x_1 = 0, x_2 = 0, x_3 = 16, x_4 = 31, x_5 = 14$  - оптимальное решение задачи, наименьшее значение линейной формы равно 7.

**Пример 5.**

Минимизировать функцию

$$F = 2x_1 + x_2 + 4x_3 - x_4 \text{ при условиях}$$

$$\begin{cases} 3x_1 + 5x_2 + 2x_3 - 2x_4 = 7, \\ 2x_1 + x_2 - x_3 - 4x_4 = 8, \end{cases}$$

$$x_j \geq 0, \quad j = \overline{1,4}.$$

**Решение.**

Для получения начального базиса введем искусственные переменные  $x_5$  и  $x_6$  в систему ограничений.

$$\begin{cases} 3x_1 + 5x_2 + 2x_3 - 2x_4 + x_5 = 7, \\ 2x_1 + x_2 - x_3 - 4x_4 + x_6 = 8, \end{cases}$$

$$x_j \geq 0, \quad j = \overline{1,6}.$$

$$x_5, x_6 - \text{базисные переменные.}$$

$$F - 2x_1 - x_2 - 4x_3 + x_4 = 0.$$

Искусственная линейная форма  $S = x_5 + x_6$ , после исключения базисных переменных.  $S = 15 - 5x_1 - 6x_2 - x_3 + 6x_4$ .

Дальнейшее решение приведено в таблице 4.

Таблица 4

Б	В	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
$x_5$	7	3	5	2	-2	1	
$x_6$	8	2	1	-1	-4		1
$F$	0	-2	-1	-4	1		
$S$	15	-5	-6	-1	6		
$x_2$	$7/5$	$3/5$	1	$2/5$	$-2/5$		
$x_6$	$33/5$	$7/5$		$7/5$	$-18/5$		1
$F$	$7/5$	$-7/5$		$-18/5$	$3/5$		
$S$	$33/5$	$-7/5$		$-7/5$	$18/5$		
$x_1$	$7/3$	1	$5/3$	$2/3$	$-2/3$		
$x_6$	$10/3$		$-7/3$	$-7/3$	$-8/3$		
$F$	$14/3$		$7/3$	$-8/3$	$-1/3$		
$S$	$10/3$		$7/3$	$7/3$	$8/3$		

Все коэффициенты в строке  $S$  положительны. Но  $S$  в нуль не обратилась: невозможно найти допустимое базисное решение исходной задачи.

Построим О.Д.Р. этой задачи.

Пусть  $x_1, x_2$  - базисные переменные,  $x_3, x_4$  - свободные.

Систему

$$\begin{cases} 3x_1 + 5x_2 + 2x_3 - 2x_4 = 7, \\ 2x_1 + x_2 - x_3 - 4x_4 = 8 \end{cases}$$

приведем к виду

$$\begin{cases} x_3 = \frac{12}{10} - \frac{8}{10}x_1 - \frac{18}{10}x_2, \\ x_4 = -\frac{23}{10} + \frac{7}{10}x_1 + \frac{7}{10}x_2 \end{cases}$$

и построим ОДР в системе координат  $x_1 O x_2$ .

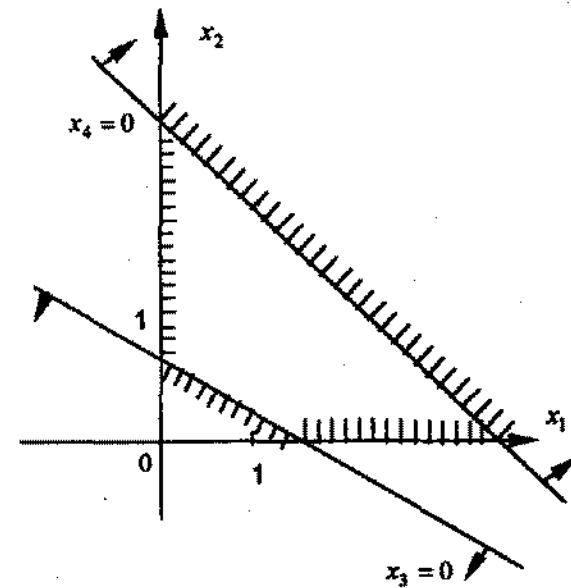


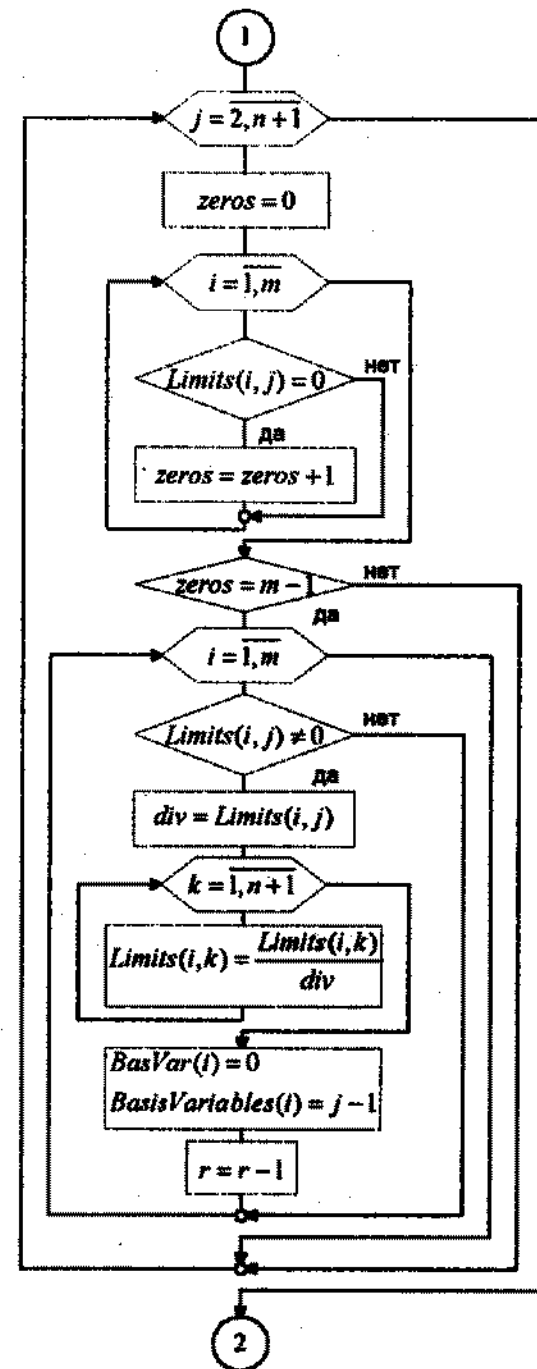
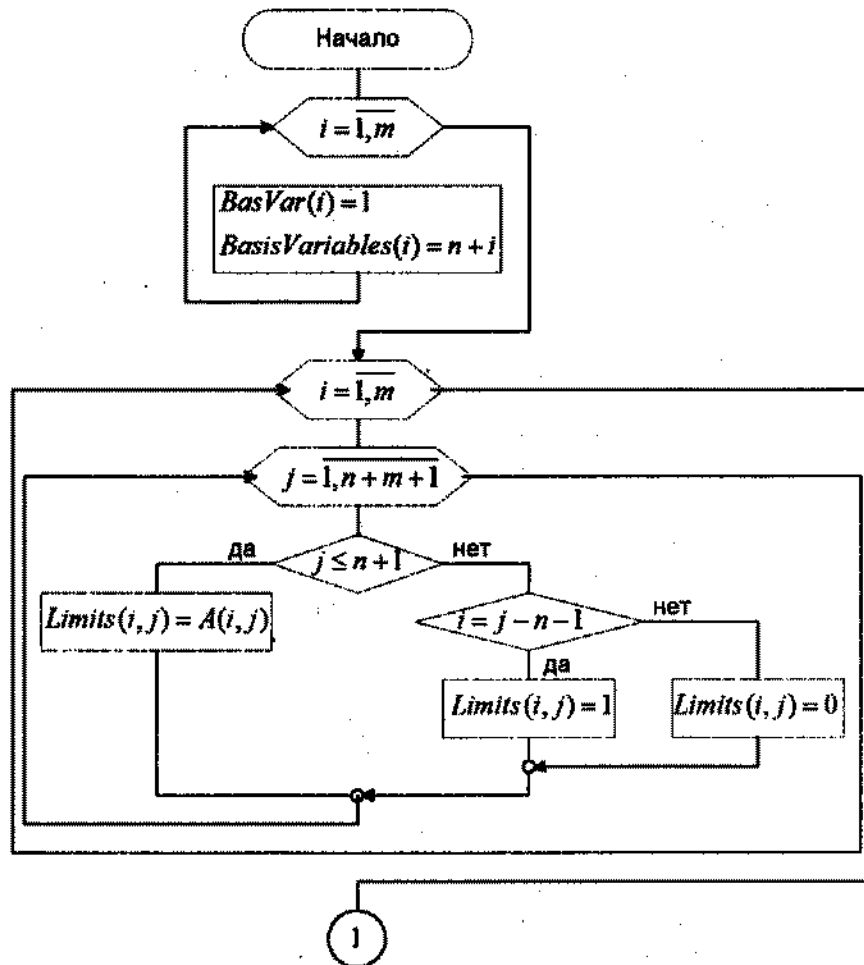
Рис. 9.4.

Система противоречива

Нет ОДР, а значит нет и решения исходной задачи.

### Структурная схема алгоритма симплекс-метода

Входные параметры:  
 $Limits(m, m+n+1)$  – матрица ограничений,  
 $c(n+1)$  – исходная линейная форма,  
 $A(n+1, m)$  – исходная матрица ограничений.  
 Выходные параметры:  
 $BasisSol(n+m)$  – массив базисных решений.



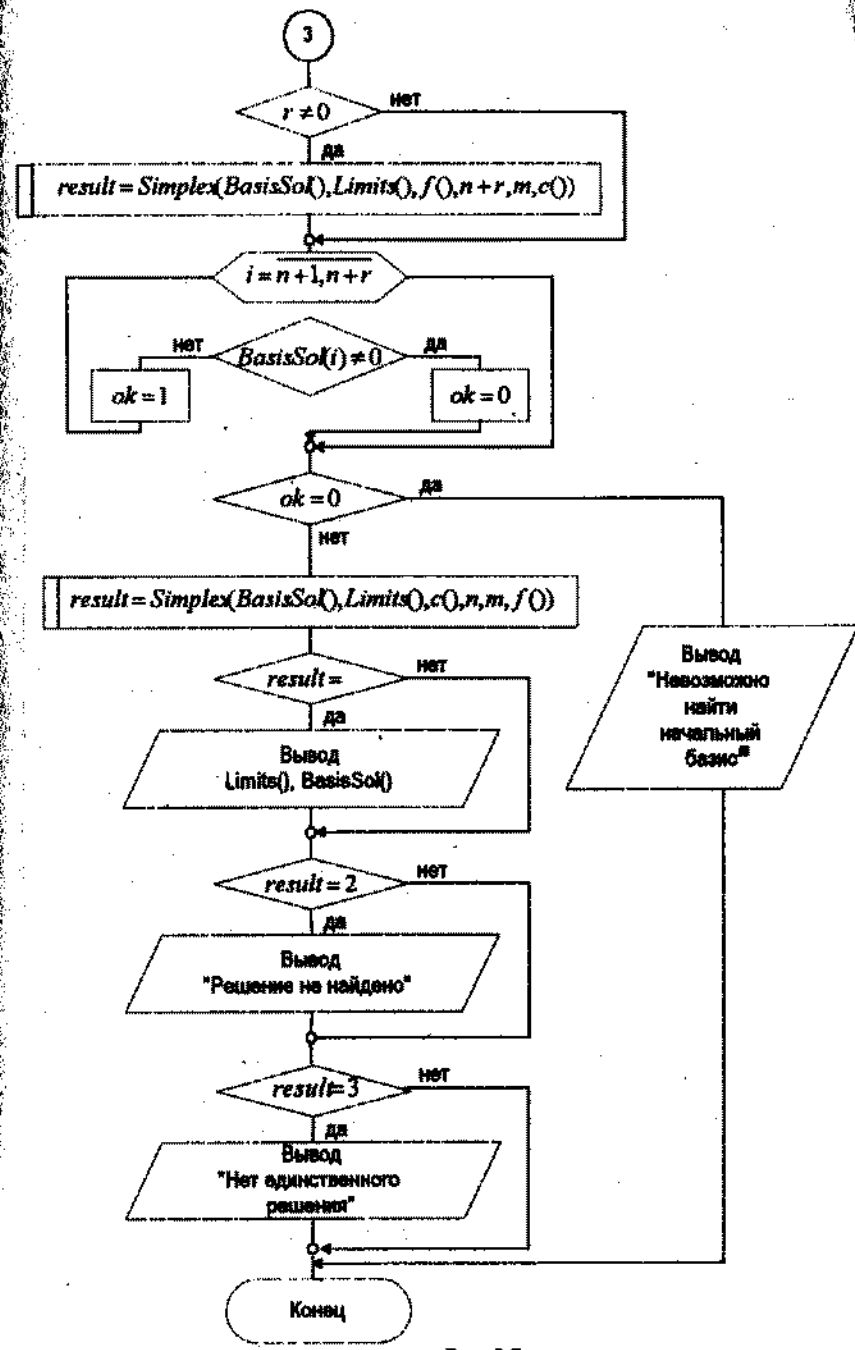
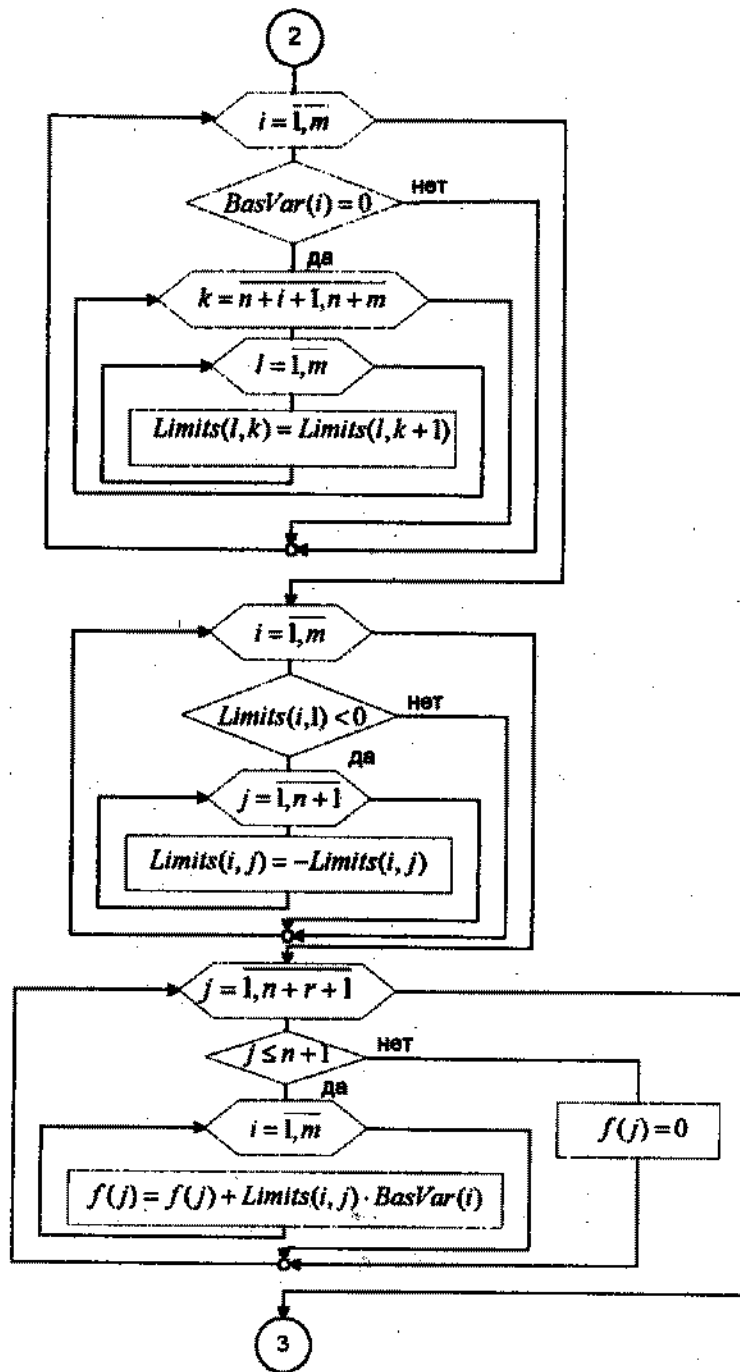


Рис. 9.5

### Процедура построения новой канонической формы

Входные параметры:  
 $Lims()$  – матрица ограничений;  
 $a()$  – столбец с разрешающим элементом;  
 $s$  – количество отрицательных элементов в столбце с разрешающим элементом;  
 $m, n$  – количество переменных и строк в системе ограничений.

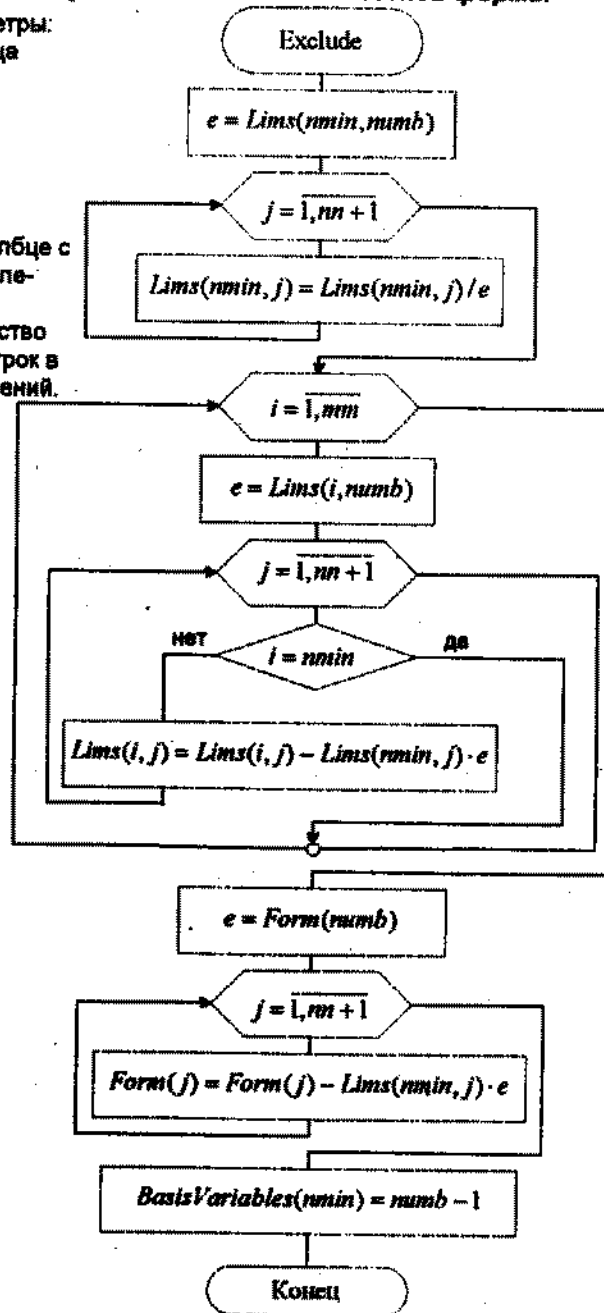


Рис. 9.6

### Процедура поиска переменной для исключения из базиса

Входные параметры:  
 $Lims()$  – матрица ограничений;  
 $a()$  – столбец с разрешающим элементом;  
 $s$  – количество отрицательных элементов в столбце с разрешающим элементом;  
 $m, n$  – количество переменных и строк в системе ограничений.

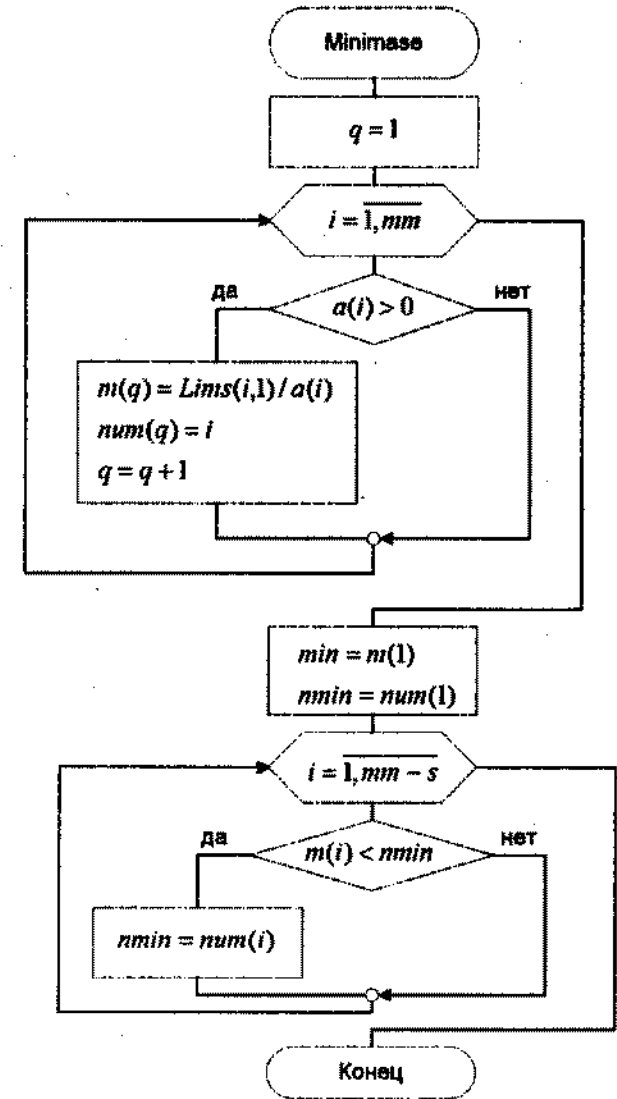
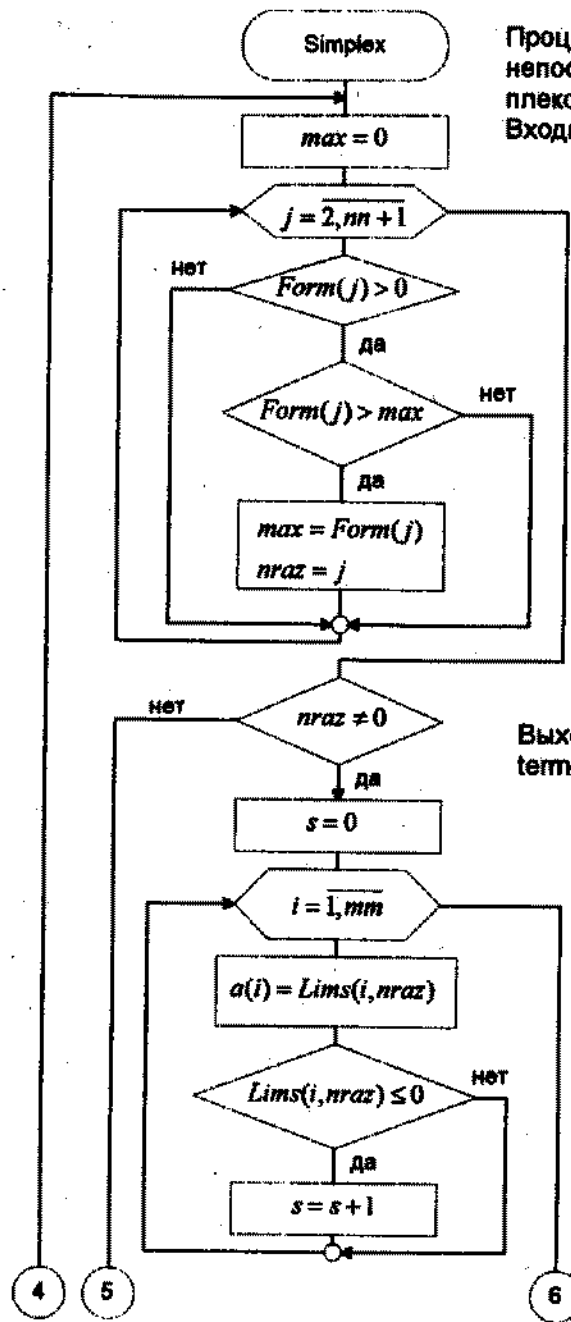
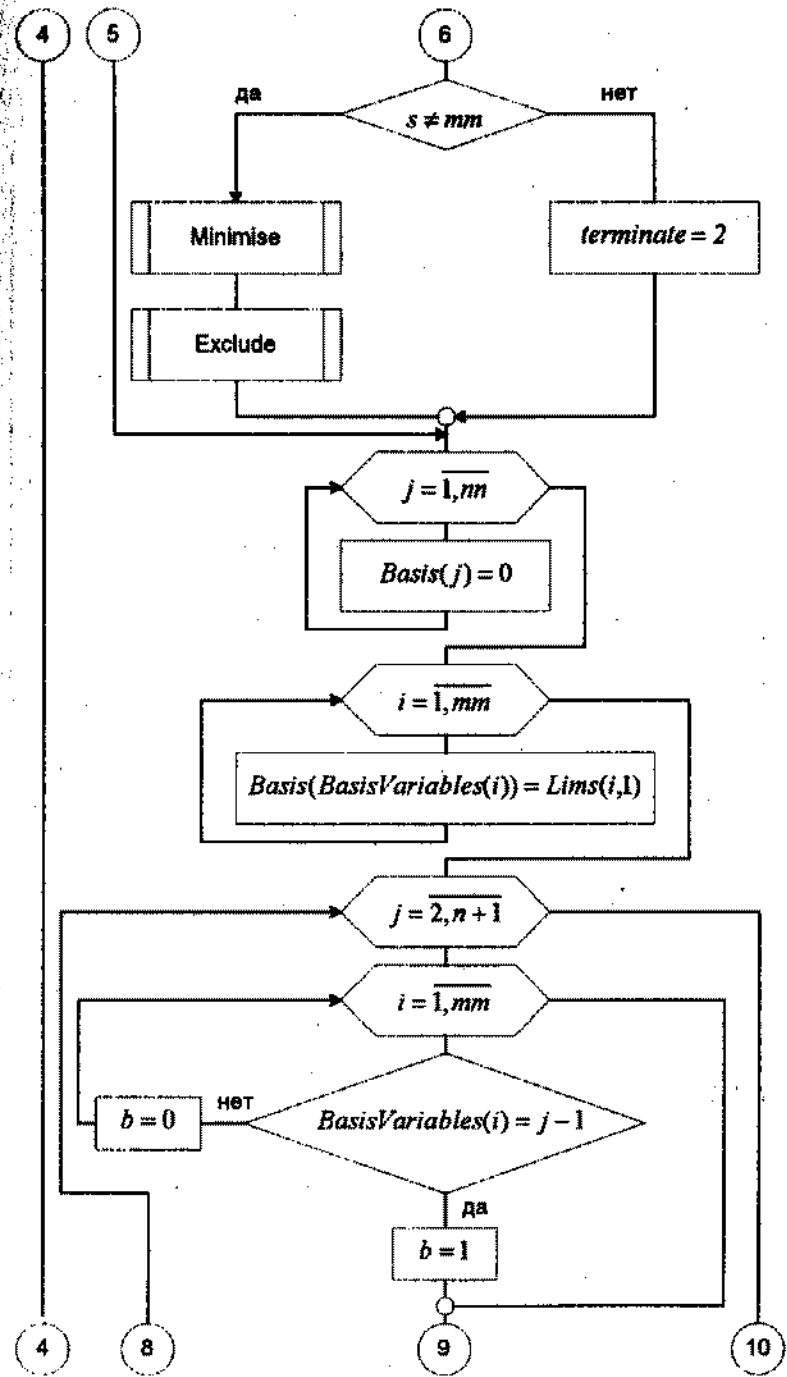


Рис. 9.7



Процедура реализует непосредственно симплекс-метод.  
 Входные параметры:  
 basis ( ) – массив базовых решений;  
 Lims ( ) – матрица ограничений;  
 Forms ( ) – текущая линейная форма;  
 nn – количество переменных в системе ограничений;  
 mm – количество уравнений в системе ограничений.

Выходные параметры:  
 terminate – код выхода.



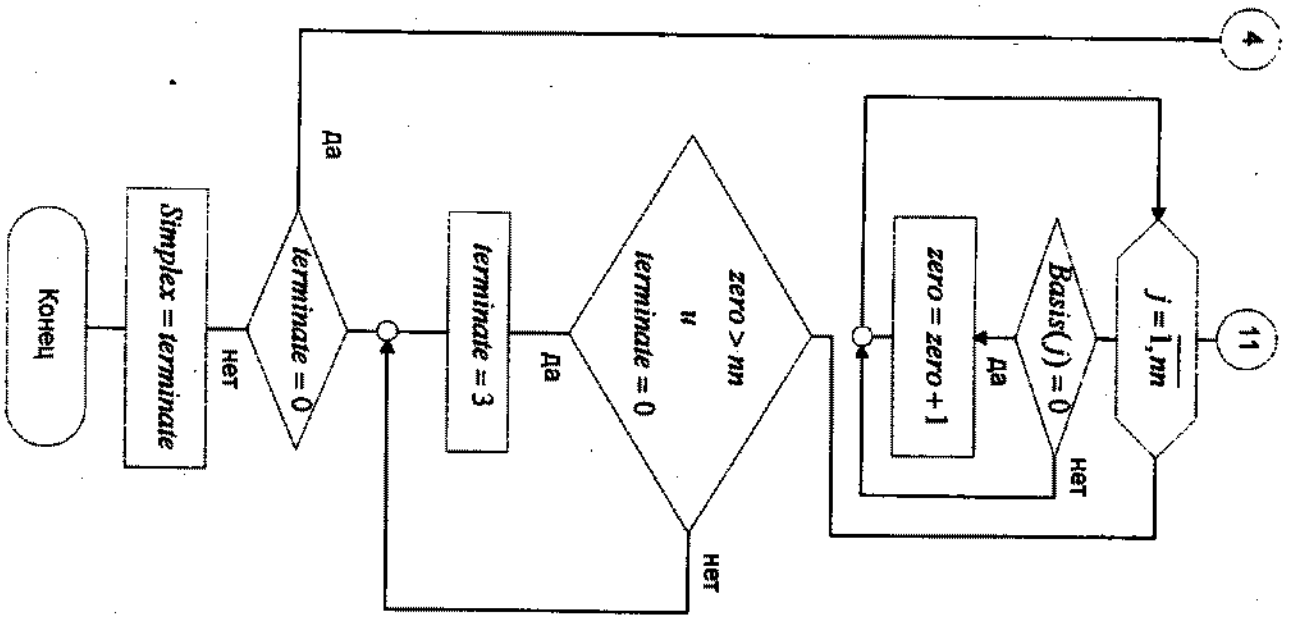
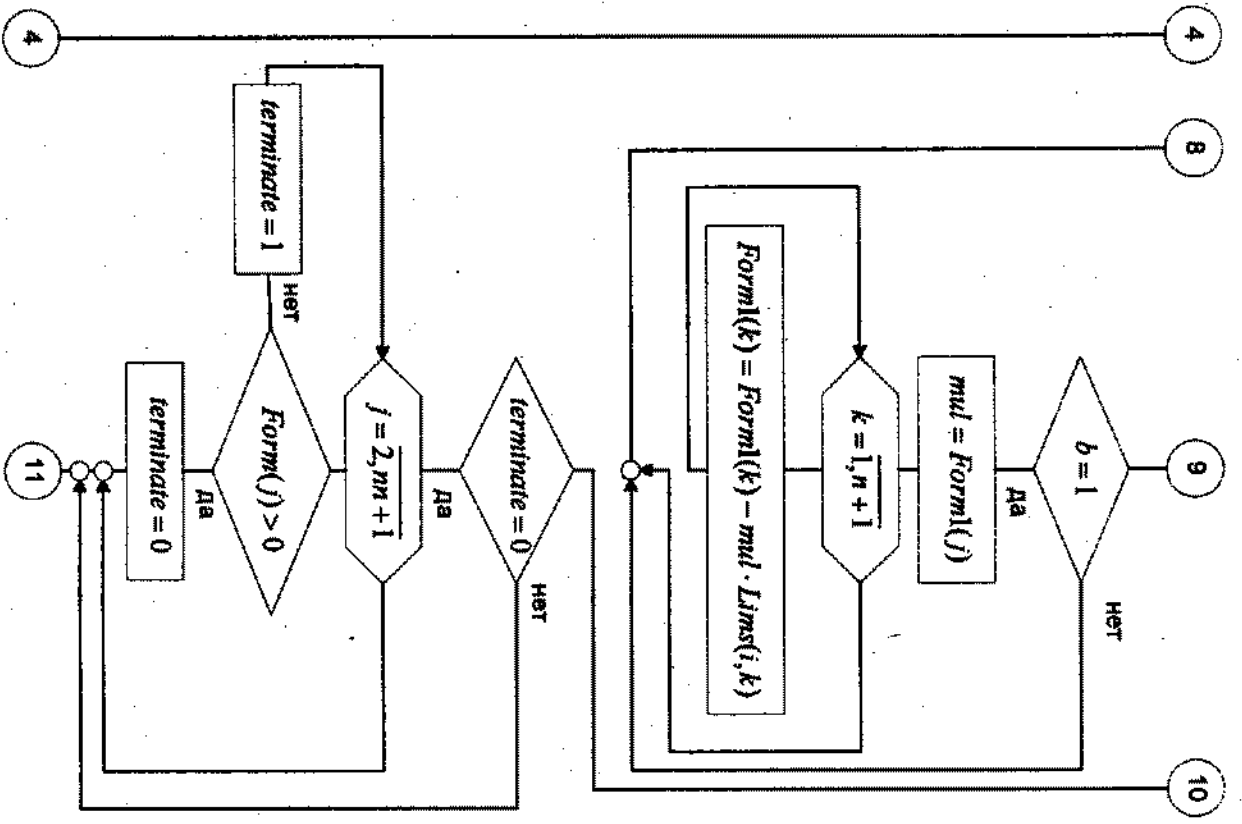


Рис. 9.8

## 9.6. Двойственность в линейном программировании

Понятие двойственности - одно из фундаментальных понятий линейного программирования.

Каждой задаче линейного программирования можно поставить в соответствие другую, вполне определенную задачу, называемую двойственной.

Рассмотрим задачу линейного программирования, в следующей формулировке:

максимизировать функцию

$$F = c_0 + c_1 x_1 + c_2 x_2 + \dots + c_n x_n \quad (3)$$

при ограничениях

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b_2, \\ \dots \dots \dots \dots \dots \dots \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m. \end{cases} \quad (18)$$

$$x_i \geq 0, \quad i = \bar{1}, n. \quad (2)$$

Сформулируем задачу, двойственную к заданной.

Минимизировать функцию:

$$S = b_1 y_1 + b_2 y_2 + \dots + b_m y_m \quad (19)$$

при ограничениях

$$\begin{cases} a_{11}y_1 + a_{21}y_2 + \dots + a_{m1}y_m \geq c_1, \\ a_{12}y_1 + a_{22}y_2 + \dots + a_{m2}y_m \geq c_2, \\ \dots \dots \dots \dots \dots \dots \dots \\ a_{1n}y_1 + a_{2n}y_2 + \dots + a_{mn}y_m \geq c_n. \end{cases} \quad (20)$$

$$y_i \geq 0, \quad i = \bar{1}, m. \quad (21)$$

Рассмотренную пару задач называют взаимно двойственными.

Для любой пары двойственных задач имеют место теоремы двойственности [80], [81].

- Если как исходная, так и двойственная задачи имеют решения, то они имеют и оптимальные решения, при этом экстремальные значения линейных форм равны:  $F_{\max} = S_{\min}$ . (теорема о минимаксе).
- Если линейная форма одной из задач не ограничена (исходная - сверху, двойственная снизу), то система ограничений другой задачи противоречива.
- Симплексный метод, примененный к одной из задач, проводит к решению другой.

Изобразим в таблице симметричную пару задач. Исходная задача читается слева направо, двойственная - сверху вниз.

Таблица 4

	$x_1 \geq 0$	$x_2$	$\dots$	$x_n$		
$y$	$a_{11}$	$a_{12}$	$\dots$	$a_{1n}$	$\leq$	$b_1$
$y$	$a_{21}$	$a_{22}$	$\dots$	$a_{2n}$	$\leq$	$b_2$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$y$	$a_{m1}$	$a_{m2}$	$\dots$	$a_{mn}$	$\leq$	$b_m$
	$\geq$	$\geq$	$\dots$	$\geq$	$S \rightarrow \min$	
	$c_1$	$c_2$	$\dots$	$c_n$	$F \rightarrow \max$	

### 9.6.1. Геометрическая иллюстрация двойственных задач

Если в исходной и двойственной задачах, образующих пару, разность между числом неизвестных и числом базисных неизвестных равно двум, то можно решить задачу геометрически. При этом возможны три случая:

- обе задачи имеют оптимальное решение,
- оптимальное решение имеет только одна задача,
- ни одна из задач не имеет оптимального решения.

#### Пример 6 [81].

Для задачи, состоящей в определении максимального значения  $F = 2x_1 + 7x_2$

при ограничениях

$$\begin{cases} -2x_1 + 3x_2 \leq 14, \\ x_1 + x_2 \leq 8, \end{cases} \quad (23)$$

$$x_1, x_2 \geq 0,$$

составить двойственную задачу и найти решение обеих задач.



**Решение.**

Двойственной задачей по отношению к исходной является задача (таблица 4), состоящая в определении минимального значения  $S = 14y_1 + 8y_2$  при условиях

$$\begin{cases} -2y_1 + y_2 \geq 2, \\ 3y_1 + y_2 \geq 7, \\ y_1, y_2 \geq 0. \end{cases} \quad (24)$$

Воспользуемся алгоритмом построения геометрического решения задачи линейного программирования.

**Геометрическая иллюстрация решения задачи (23).**

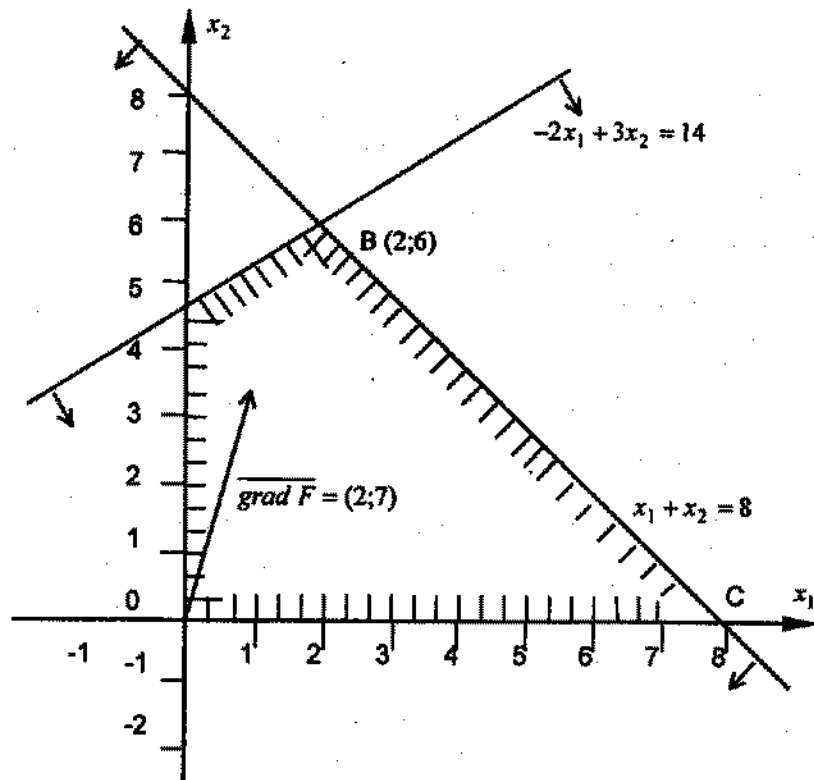


Рис. 9.5

**Геометрическая иллюстрация задачи (24).**

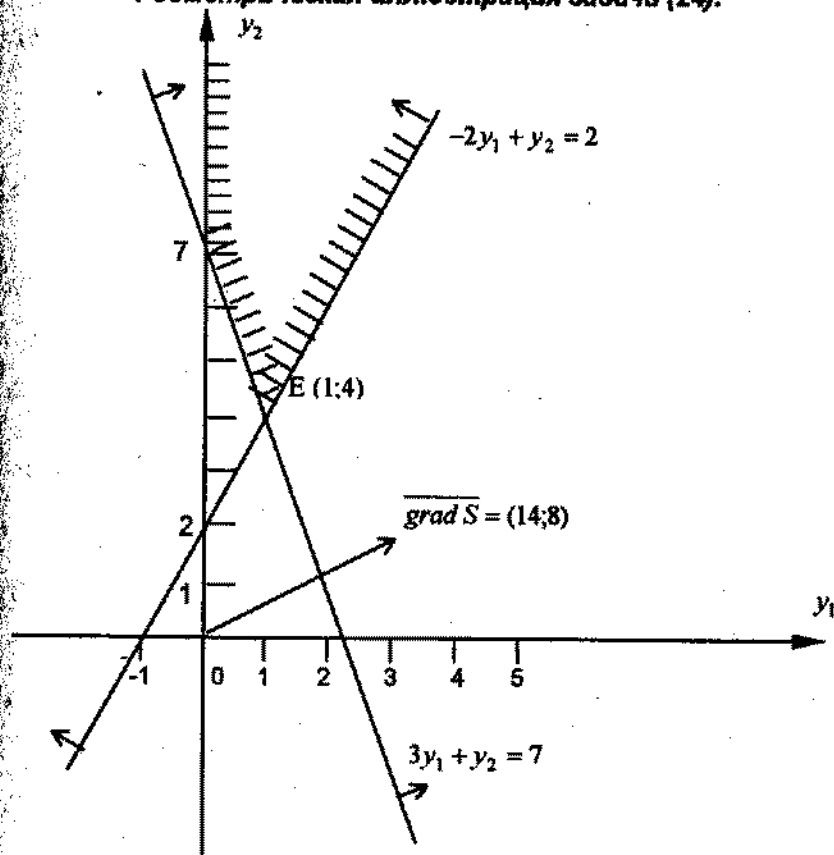


Рис. 9.6

На рис. 5 и 6 видно что задачи имеют оптимальное решение.

$$F_{\max}^* = F(2;6) = 46, \quad S_{\min}^* = S(1;4) = 46.$$

### 9.6.2. Симплексный метод решения двойственной задачи

Запишем задачу (2), (3), (18) и двойственную ей в канонической форме. Введем дополнительные переменные  $x_{n+1}, x_{n+2}, \dots, x_{m+n}$  для исходной задачи, переписываем ограничения в виде равенств.

$$\begin{cases} b_1 - (a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n) = x_{n+1}, \\ \dots \dots \dots \dots \dots \dots \dots \dots \\ b_m - (a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n) = x_{n+m}, \end{cases} \quad (25)$$

$$F_1 = -c_0 - (c_1x_1 + c_2x_2 + \dots + c_nx_n), \quad x_j \geq 0, \quad j = 1, n+m.$$

Для двойственной задачи дополнительные переменные

$$y_{m+1}, y_{m+2}, \dots, y_{m+n}.$$

$$\begin{cases} a_{11}y_1 + a_{21}y_2 + \dots + a_{m1}y_m - c_1 = y_{m+1}, \\ \dots \dots \dots \dots \dots \dots \dots \dots \\ a_{1n}y_1 + a_{2n}y_2 + \dots + a_{mn}y_m - c_n = y_{m+n}, \end{cases} \quad (26)$$

$$S = c_0 + b_1y_1 + \dots + b_my_m.$$

Между неизвестными  $x_1, \dots, x_{n+m}$  исходной задачи и неизвестными  $y_1, \dots, y_{m+n}$  двойственной задачи установим соответствие

$$\begin{array}{cccccccc} x_1 & x_2 & \dots & x_n & x_{n+1} & x_{n+2} & \dots & x_{n+m} \\ \updownarrow & \updownarrow & & \updownarrow & \updownarrow & \updownarrow & & \updownarrow \\ y_{m+1} & y_{m+2} & \dots & y_{m+n} & y_1 & y_2 & \dots & y_m \end{array} \quad (27)$$

так, что базисным неизвестным одной задачи соответствуют свободные неизвестные другой. Базисные неизвестные исходной задачи -  $x_{n+1}, x_{n+2}, \dots, x_{n+m}$ , в двойственной задаче базисные неизвестные -  $y_{m+1}, \dots, y_{m+n}$ .

При установленном соответствии имеют место следующие свойства коэффициентов:

- Базисные неизвестные системы задачи (25) соответствуют свободным неизвестным задачи (26) и наоборот.
- Для любого набора базисных переменных коэффициенты, с которыми свободные неизвестные входят в выражение для базисной неизвестной  $x_{m+j}$ ,  $j = \overline{1, m}$  в системе (25) отличаются лишь знаком от коэффициентов, с которыми свободные неизвестные  $y_j$  (они соответствуют  $x_{n+j}$ ) входят в уравнения системы (26).
- Если некоторый набор свободных неизвестных двойственной задачи определяет ее оптимальное решение, то соответствующий набор базисных и свободных неизвестных исходной задачи определяет оптимальное решение исходной задачи.

Свойства коэффициентов и связь между элементами симплексных таблиц сохраняется при преобразовании базиса.

Если двойственная задача имеет оптимальное решение, то в последней симплексной таблице среди оценочных коэффициентов (коэффициентов линейной формы) не будет ни одного положительного, а в столбце свободных членов ни одного отрицательного (кроме, быть может, значения целевой функции).

Симплексная таблица исходной задачи, в силу указанного соответствия между строками и столбцами, в столбце свободных членов не будет содержать ни одного отрицательного элемента, а в последней строке не будет ни одного положительного оценочного коэффициента. Это означает, что для исходной задачи получено оптимальное решение, причем:  $F_{\max} = S_{\min}$ .

При любых допустимых значениях переменных значение целевой функции исходной задачи не превосходит значения целевой функции  $S$  при произвольных значениях  $y_1$  и  $y_2$ .

### 9.6.3 Двойственный симплексный метод

Для решения некоторой задачи линейного программирования можно составить двойственную ей задачу. Решив последнюю симплексным методом, получим решение и исходной задачи. Такой способ решения называется двойственным симплексным методом.

Преимущества этого подхода в том, что в обычном симплексном методе осуществляется переход от одного допустимого неоптимального решения к другому, сохраняя условия допустимости. В двойственном методе условие допустимости может быть нарушено (что чаще всего и происходит). Кроме того, необходимо в симплексном методе отыскивать разрешающую неизвестную, затем разрешающее уравнение.

В двойственном симплексном методе вначале определяется разрешающее уравнение, как имеющее наименьший отрицательный свободный член

$$h_i = \min_j (h_j < 0), \quad (27)$$

затем разрешающую переменную.

Чтобы указать разрешающую переменную, составляем отношение коэффициентов  $\Delta_j$  (коэффициент линейной формы  $F$  в первой симплексной таблице) к соответствующим отрицательным коэффициентам при неизвестных в выбранном разрешающем уравнении.

$$\text{Если } \frac{\Delta_s}{g_{rs}} = \min_j \left( \frac{\Delta_j}{g_{rj}} \right), \quad g_{rj} < 0, \quad (28)$$

то неизвестную  $x_1$ , принимаем за разрешающую. Заметим, что при переходе к следующей симплексной таблице условие  $\Delta_j \leq 0$  сохраняется.

**Пример 7**

Минимизировать

$$F = 2x_1 + 3x_3 + 2x_4$$

при ограничениях

$$\begin{cases} x_1 - x_2 + 2x_3 + 3x_4 \geq 9, \\ x_1 + 2x_2 - x_3 - 2x_4 \geq 8, \\ x_j \geq 0, \quad j = \overline{1,4}. \end{cases}$$

**Решение**

Введем неизвестные  $x_5$  и  $x_6$ , перепишем систему ограничений в виде

$$\begin{cases} -x_1 + x_2 - 2x_3 - 3x_4 + x_5 = -9, \\ -x_1 - 2x_2 - x_3 + 2x_4 + x_6 = -8, \\ x_j \geq 0, \quad j = \overline{1,6}. \end{cases}$$

Составим таблицу и решим задачу двойственным симплексным методом.

Таблица 5

$\tilde{c}$	Б	b						
			2	3	2			
			$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
	$x_5$	-9	-1	1	-2	-3	1	
	$x_6$	-8	-1	-2	1	2		1
	$F_1$		-2		-3	-2		

Среди коэффициентов  $\Delta_1 = -2, \Delta_2 = 0, \Delta_3 = -3, \Delta_4 = -2, \Delta_5 = 0, \Delta_6 = 0$  (коэффициенты линейной формы первой симплексной таблицы) нет ни одного положительного, поэтому начальное базисное решение  $x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 0, x_5 = -9, x_6 = -8$  удовлетворяет условию оптимальности, но не является допустимым.

Строим вторую таблицу, переходим от одного недопустимого решения к другому, сохраняя условие оптимальности, до тех пор,

либо не получим допустимого оптимального решения, либо не будет доказана неразрешимость задачи.

Разрешающее уравнение – первое, так как оно имеет наибольший по модулю отрицательный свободный член  $\min(-9; -8) = -9 = h_1$ .

Выберем разрешающую переменную.

$$\min_i \left( \frac{\Delta_i}{g_{ij} < 0} \right) = \min \left( \frac{-2}{-1}, \frac{-3}{-2}, \frac{-2}{-3} \right) = \frac{\Delta_4}{g_{14}}$$

$g_{14}$  – отрицательные коэффициенты в выбранном разрешающем уравнении,  $r = 1, x_4$  – разрешающая переменная, вводим  $x_4$  в число базисных, разрешающий элемент  $g_{14} = -3$ . Так как в первом уравнении  $x_5$  – базисная переменная, то переменную  $x_5$  выводим из базиса в свободные. По известным правилам исчисления переходим ко второй симплексной таблице.

Таблица 6

$\tilde{c}$	Б	b	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
2	$x_4$	3	$\frac{1}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$	1	$-\frac{1}{3}$	
	$x_6$	-14	$-\frac{5}{3}$	$-\frac{4}{3}$	$-\frac{1}{3}$		$\frac{2}{3}$	1
	$F$	6	$-\frac{4}{3}$	$-\frac{2}{3}$	$-\frac{5}{3}$		$-\frac{2}{3}$	
2	$x_4$	$\frac{13}{2}$	$\frac{3}{4}$		$\frac{3}{4}$	1	$-\frac{1}{2}$	$-\frac{1}{4}$
	$x_2$	$\frac{21}{2}$	$\frac{5}{4}$	1	$\frac{1}{4}$		$-\frac{1}{2}$	$-\frac{3}{4}$
	$F$	13	$-\frac{1}{2}$		$-\frac{3}{4}$		-1	$-\frac{1}{2}$

Во второй таблице разрешающее уравнение – второе, разрешающая переменная  $-x_2$ .

В третьей симплексной таблице нет ни одного положительного коэффициента в последней строке и ни одного отрицательного свободного члена (столбец  $b$ ). Это означает что базисное решение

$$\bar{x} = \left\{ 0, \frac{21}{2}, 0, \frac{13}{2}, 0, 0 \right\}$$
 является оптимальным. Наименьшее значение

$$\text{линейной формы } F_{\min} = 13, \quad \bar{x}^{opt} = \left\{ 0, \frac{21}{2}, 0, \frac{13}{2}, 0, 0 \right\}.$$

Теперь сформулируем и решим двойственную задачу симплексным методом.

Максимизировать

$$S = 9y_1 + 8y_2$$

при ограничениях

$$\begin{cases} y_1 + y_2 \leq 2, \\ -y_1 + 2y_2 \leq 0, \\ 2y_1 - y_2 \leq 3, \\ 3y_1 - 2y_2 \leq 2, \end{cases} \quad (29)$$

$$y_i \geq 0, \quad i = \overline{1, 2}.$$

Система ограничений с учетом дополнительных переменных

$$\begin{cases} y_1 + y_2 + y_3 = 2, \\ -y_1 + 2y_2 + y_4 = 0, \\ 2y_1 - y_2 + y_5 = 3, \\ 3y_1 - 2y_2 + y_6 = 2, \end{cases}$$

$$y_i \geq 0, \quad i = \overline{1, 6},$$

$$-S = -9y_1 - 8y_2 \rightarrow \min.$$

Взаимное соответствие неизвестных:

$$\begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ y_3 & y_4 & y_5 & y_6 & y_1 & y_2 \end{array}$$

Первая симплексная таблица.

Таблица 7

$\bar{c}$	Б	b	9	8				
			$y_1$	$y_2$	$y_3$	$y_4$	$y_5$	$y_6$
	$y_3$	2	1	1	1			
	$y_4$	0	-1	2		1		
	$y_5$	3	2	-1			1	
	$y_6$	2	3	-2				1
	$-S$		9	8				

$\max(8, 9) = 9$ , следовательно  $y_1$  - разрешающая переменная.

$$\min\left(\frac{2}{1}, \frac{3}{-2}, \frac{2}{3}\right) = \frac{2}{3} \quad (\text{отношение свободных членов столбца } b \text{ к положительным коэффициентам в столбце, содержащим разрешающую}$$

переменную  $y_1$ ), разрешающее уравнение - четвертое. Вводим в

число базисных переменную  $y_1$ , переменную  $y_6$  выводим из базиса в свободные. Продолжим решение.

Таблица 8

$\bar{c}$	Б	b	9	8				
			$y_1$	$y_2$	$y_3$	$y_4$	$y_5$	$y_6$
	$y_3$	$\frac{4}{3}$		$\frac{5}{3}$	1			$-\frac{1}{3}$
	$y_4$	$\frac{2}{3}$		$\frac{4}{3}$		1		$\frac{1}{3}$
	$y_5$	$\frac{5}{3}$		$\frac{1}{3}$			1	$-\frac{2}{3}$
9	$y_1$	$\frac{2}{3}$		$-\frac{2}{3}$				$\frac{1}{3}$
	$-S$	6		14				-3
	$y_3$	$\frac{1}{2}$			1	$-\frac{5}{4}$		$-\frac{3}{4}$
8	$y_2$	$\frac{1}{2}$		1		$\frac{3}{4}$		$\frac{1}{4}$
	$y_5$	$\frac{3}{2}$				$-\frac{1}{4}$	1	$-\frac{3}{4}$
9	$y_1$	1	1			$\frac{1}{2}$		$\frac{1}{2}$
	$-S$	13				$-\frac{21}{2}$		$-\frac{13}{2}$

Таким образом, получаем оптимальное решение

$$y^{opt} = \left(0; 0; 0; -\frac{21}{2}; 0; -\frac{13}{2}\right),$$

$$S_{\max} = 13.$$

Заметим, что полученное базисное решение двойственной задачи соответствует базисному решению исходной задачи.

$$x^{opt} = \left(0; \frac{21}{2}; 0; \frac{13}{2}; 0; 0\right).$$

**Пример 8**

Составить и решить двойственную задачу к задаче

$$\begin{cases} 2x_1 + 3x_2 \leq 19, \\ 2x_1 + x_2 \leq 13, \\ 3x_2 \leq 15, \\ 3x_1 \leq 18, \end{cases} \quad (30)$$

$$x_j \geq 0, \quad j = \overline{1, 4}.$$

$$F = 7x_1 + 5x_2 \rightarrow \max.$$

**Решение**

Запишем задачу (30) в канонической форме.

$$\begin{cases} 19 - (2x_1 + 3x_2) = x_3, \\ 13 - (2x_1 + x_2) = x_4, \\ 15 - 3x_2 = x_5, \\ 18 - 3x_1 = x_6, \end{cases}$$

$$F_1 = -7x_1 - 5x_2 \rightarrow \max.$$

$$x_j \geq 0, \quad j = \overline{1, 6}.$$

 $x_1$  и  $x_2$  - свободные,  $x_3, x_4, x_5, x_6$  - базисные неизвестные.Двойственная задача заключается в отыскании минимума функции  $S = 19y_1 + 13y_2 + 15y_3 + 18y_4 \rightarrow \min$ , при ограничениях

$$\begin{cases} 2y_1 + 2y_2 + 3y_4 - 7 = y_5, \\ 3y_1 + y_2 + 3y_3 - 5 = y_6, \end{cases} \quad (31)$$

 $y_5, y_6$  - дополнительные переменные (базисные переменные для двойственной канонической задачи).

Таблица соответствия неизвестных:

$$\begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ y_5 & y_6 & y_1 & y_2 & y_3 & y_4 \end{array}$$

Пусть  $y_3$  и  $y_4$  - базисные переменные задачи (31).

$$\begin{cases} y_3 = \frac{5}{3} - \left( y_1 + \frac{1}{3}y_2 - \frac{1}{3}y_6 \right), \\ y_4 = \frac{7}{3} - \left( \frac{2}{3}y_1 + \frac{2}{3}y_2 - \frac{1}{3}y_5 \right), \end{cases}$$

$$S = 67 - (8y_1 + 4y_2 - 6y_5 - 5y_6).$$

Тогда в задаче (30)  $x_5$  и  $x_6$  - свободные,  $x_3, x_4, x_1, x_2$  - базисные.

$$\begin{cases} x_3 = -8 - \left( -x_5 - \frac{2}{3}x_6 \right), \\ x_4 = -4 - \left( -\frac{1}{3}x_5 - \frac{2}{3}x_6 \right), \\ x_1 = 6 - \frac{1}{3}x_6, \\ x_2 = 5 - \frac{1}{3}x_5, \\ F = -67 - \left( -\frac{5}{3}x_5 - \frac{7}{3}x_6 \right). \end{cases}$$

Решаем двойственную задачу (31) симплексным методом. Первая симплексная таблица

Таблица 9

$\tilde{c}$	Б	$b$	8	4	-6	-5
			$y_1$	$y_2$	$y_5$	$y_6$
	$y_3$	$\frac{5}{3}$	1	$\frac{1}{3}$	0	$-\frac{1}{3}$
	$y_4$	$\frac{7}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$	0
	$S$	67	8	4	-6	-5

Вторая симплексная таблица

$\tilde{c}$	Б	$b$	$y_1$	$y_4$	$y_5$	$y_6$
	$y_3$	$\frac{1}{2}$	$\frac{2}{3}$	$-\frac{1}{2}$	$\frac{1}{6}$	$-\frac{1}{3}$
4	$y_2$	$\frac{7}{3}$	1	$\frac{3}{2}$	$-\frac{1}{2}$	0
	$S$	53	4	-6	-4	-5

Третья симплексная таблица

$\tilde{c}$	Б	$b$	$y_3$	$y_4$	$y_5$	$y_6$
8	$y_1$	$\frac{3}{4}$	$\frac{3}{2}$	$-\frac{3}{4}$	$\frac{1}{4}$	$-\frac{1}{2}$
4	$y_2$	$\frac{11}{3}$	$\frac{3}{2}$	$\frac{9}{4}$	$-\frac{3}{4}$	$\frac{1}{2}$
	$S$	50	-6	-3	-5	-3

В последней строке таблицы нет положительных коэффициентов, то есть получено оптимальное решение.

$$S_{\min} = 50.$$

Из теоремы о минимаксе следует

$$F_{\max} = S_{\min} = 50.$$

Если воспользоваться таблицей соответствия и свойствами коэффициентов исходной и двойственной задач, можно записать последнюю симплексную таблицу для исходной задачи.

Таблица 10

$\bar{c}$	Б	$b$	$x_3$	$x_4$
	$x_5$	6	$-\frac{3}{2}$	$\frac{3}{2}$
	$x_6$	3	$\frac{3}{4}$	$-\frac{9}{4}$
7	$x_1$	5	$-\frac{1}{4}$	$\frac{3}{4}$
5	$x_2$	3	$\frac{1}{2}$	$-\frac{1}{2}$
	$F_1$	-50	$-\frac{3}{4}$	$-\frac{11}{4}$

Итак, минимум  $F_1$  равен (-50), но  $F = -F_1$ , следовательно

$$F_{\max} = -F_{1\min} = S_{\min} = 50.$$

Оптимальный базис  $\bar{x}^{opt} = (5; 3; 0; 0; 6; 3)$ .

Заметим, что подстановка полученных значений базисных переменных в выражение  $F = 7x_1 + 5x_2$  подтверждает правильность решения задачи  $F = 7 \cdot 5 + 5 \cdot 3 = 50$ .

Решив двойственную задачу, получили решение исходной, не решая последней.

## 10. НЕЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

Задача нелинейного программирования формулируется как задача минимизации функции  $n$  переменных при наличии ограничивающих условий.

$$f(\bar{x}) \rightarrow \min \quad (1)$$

при условиях

$$\begin{cases} \varphi_1(\bar{x}) \leq b_1, \\ \varphi_2(\bar{x}) \leq b_2, \\ \dots\dots\dots \\ \varphi_m(\bar{x}) \leq b_m. \end{cases} \quad (2)$$

$$b_i \geq 0, \quad i = \overline{1, m},$$

где

$f, \varphi_i$  - суть нелинейные функции (хотя бы одна из них),

$f(\bar{x})$  - целевая функция,

$\bar{x} = (x_1, x_2, \dots, x_n)$  - вектор неизвестных.

Требование неотрицательных  $x_j$  входит в условие (2) или

записывается отдельно

$$x_j \geq 0, \quad j = \overline{1, n}. \quad (3)$$

Математические модели многих инженерных и экономических задач представляют собой задачи нелинейного программирования, в которых целевую функцию нужно максимизировать или в системе ограничений встречаются и уравнения и неравенства разных направлений. Такие задачи сводятся к виду (1) - (3) аналогично тому, как это предлагалось для задач линейного программирования. Кроме того, иногда условие (3) бывает удобным считать включенным в условия (2).

Любое неотрицательное решение системы неравенств (2) будем называть допустимым решением задачи нелинейного программирования, а оптимальным решением - такое допустимое решение, при котором целевая функция (1) принимает наименьшее возможное значение.

Множество допустимых решений будем рассматривать как некоторую совокупность точек  $n$ -мерного пространства. Точка  $\bar{x}(x_1, x_2, \dots, x_n)$  называется внутренней точкой допустимой области,

если ее координаты обращают неравенства (2) в строгие неравенства, если же хотя бы одно из условий (2) обращается в равенство, то такую точку будем называть граничной точкой.

### 10.1. Метод множителей Лагранжа

Рассмотрим решение задачи поиска минимума функции (1) в области  $G$  при ограничениях вида:

$$b_i - \varphi_i(\bar{x}) = 0, \quad (5)$$

$$i = \overline{1, m}.$$

Предполагаем, что  $f(\bar{x})$  и  $\varphi_i(\bar{x})$  непрерывны вместе со своими производными. Для решения задачи вводят набор переменных - множителей Лагранжа -  $\lambda_1, \lambda_2, \dots, \lambda_m$  и формируют функцию Лагранжа

$$F(\bar{x}, \bar{\lambda}) = f(\bar{x}) + \sum_{i=1}^m \lambda_i (b_i - \varphi_i(\bar{x})). \quad (6)$$

Необходимые условия безусловного экстремума функции (6) являются также и необходимыми условиями для условного экстремума функции  $f(\bar{x})$ .

Необходимые условия экстремума функции Лагранжа образуют систему  $(n+m)$  уравнений

$$\begin{cases} \frac{\partial f}{\partial x_i} + \sum_{j=1}^m \lambda_j \frac{\partial \varphi_j}{\partial x_i} = 0, \\ \frac{\partial F}{\partial \lambda_j} = b_j - \varphi_j(\bar{x}) = 0, \end{cases} \quad (7)$$

$$j = \overline{1, n}, \quad i = \overline{1, m}.$$

которую, вообще говоря, можно решить относительно неизвестных  $x_1, x_2, \dots, x_n, \lambda_1, \lambda_2, \dots, \lambda_m$ .

Затем необходимо дополнительно определить, в каких из найденных точек целевая функция достигает минимума и сравнить их со значениями функции (1) на границе. Необходимость дополнительных исследований связана с тем фактом, что в решения системы входят точки, в которых возможен локальный минимум.

#### Алгоритм решения задачи методом множителей Лагранжа.

1. Составить функцию Лагранжа.
2. Найти частные производные функции Лагранжа по ее аргументам.

3. Решить систему  $(n+m)$  уравнений, найти критические точки (в которых целевая функция может иметь экстремум).

4. Среди критических точек определить точки, доставляющие минимум целевой функции и вычислить значения функции  $f(\bar{x})$  в этих точках.

5. Определить точку глобального минимума.

#### 10.1.1. Геометрическая интерпретация метода Лагранжа

Относительные локальные  $\min F(\bar{x}, \bar{\lambda})$  при ограничениях  $\varphi(\bar{x})$  могут быть только в точках, где линии уровня функции  $F(\bar{x}, \bar{\lambda})$  касаются кривой  $\varphi(\bar{x})=0$  (точки  $A_1, A_2$ ). Из других точек ( $B$ ), двигаясь вдоль границы  $\varphi(\bar{x})$ , можно уменьшить значение  $F$ .

Таким образом, в точках экстремума (минимума)  $\overline{\text{grad}} F(\bar{x}, \bar{\lambda})$  и  $\overline{\text{grad}} \varphi(\bar{x})$  сонаправлены, т. е. другая форма записи необходимых условий экстремума:

$$\begin{cases} F'(\bar{x}, \bar{\lambda}) = -\lambda \varphi'(\bar{x}) \\ \varphi(\bar{x}) = 0 \end{cases} \quad (8)$$

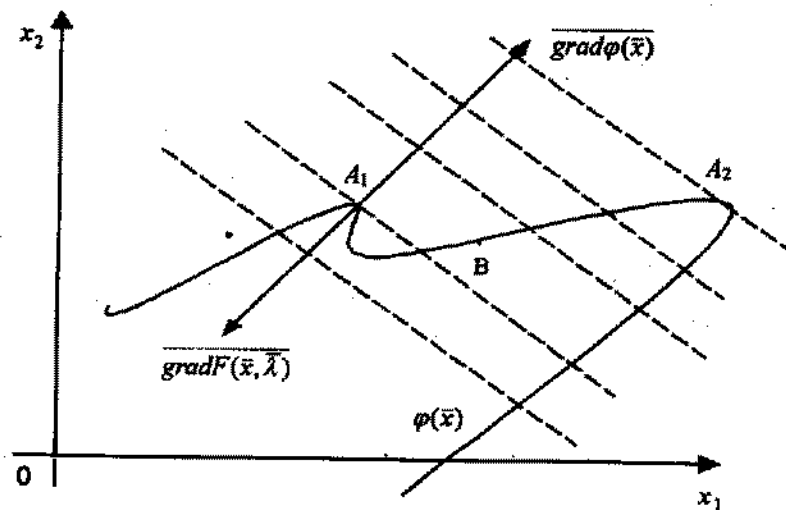


Рис. 10.1

**Пример 1.**

Найти наименьшее значение функции  $f(x_1, x_2) = 6 - 4x_1 - 3x_2$  если переменные  $x_1, x_2$  удовлетворяют условию  $x_1^2 + x_2^2 = 1$ .

**Решение.**

Составляем функцию Лагранжа

$$F(x_1, x_2, \lambda) = 6 - 4x_1 - 3x_2 + \lambda(x_1^2 + x_2^2 - 1).$$

Необходимые условия экстремума образуют систему уравнений

$$\begin{cases} -4 + 2\lambda x_1 = 0, \\ -3 + 2\lambda x_2 = 0, \\ x_1^2 + x_2^2 = 1, \end{cases}$$

решениями которой являются

$$\lambda_1 = \frac{5}{2}, \quad x_1^{(1)} = \frac{4}{5}, \quad x_2^{(1)} = \frac{3}{5}; \quad \lambda_2 = -\frac{5}{2}, \quad x_1^{(2)} = -\frac{4}{5}, \quad x_2^{(2)} = -\frac{3}{5}$$

$$f(x_1^{(1)}, x_2^{(1)}) = 1, \quad f(x_1^{(2)}, x_2^{(2)}) = 11.$$

Таким образом

$$f^* = \min_{x_1, x_2 \in G} f(x_1, x_2) = f(x_1^{(1)}, x_2^{(1)}) = 1.$$

Решим эту задачу, используя геометрическую интерпретацию.

Областью допустимых решений является окружность  $x_1^2 + x_2^2 = 1$ , линии уровня удовлетворяют условию  $6 - 4x_1 - 3x_2 = C$ ,  $C = \text{const}$ .

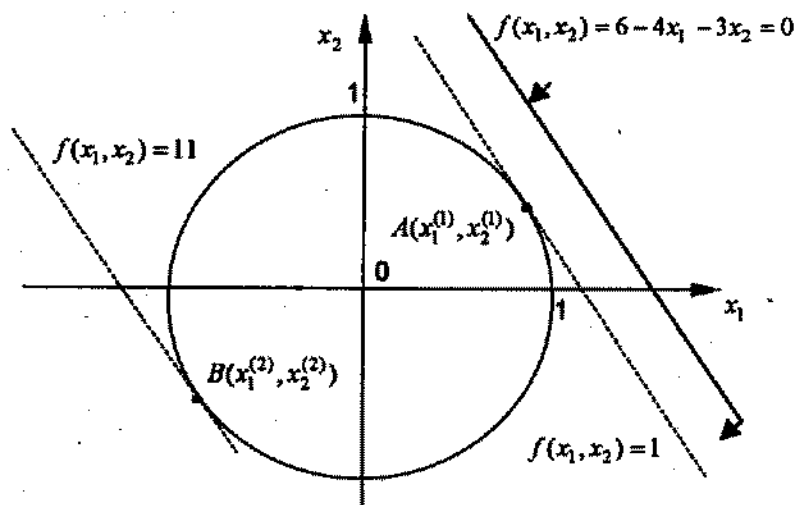


Рис. 10.2

Проводя линии уровня, соответствующих различным значениям константы  $C$ , получим две точки  $A(x_1^{(1)}, x_2^{(1)})$  и  $B(x_1^{(2)}, x_2^{(2)})$ , принадлежащие и области допустимых решений и целевой функции.

Для определения координат этих точек можно воспользоваться тем, что угловые коэффициенты прямых в точках  $A, B$  и касательной к окружности равны. Угловым коэффициентом прямой равен  $(-4/3)$ .

Из уравнения  $x_1^2 + x_2^2 = 1$  следует

$$x_2 = \pm \sqrt{1 - x_1^2},$$

$$x_2' = \mp \frac{x_1}{\sqrt{1 - x_1^2}}.$$

Из условия равенства угловых коэффициентов  $\pm \frac{x_1}{\sqrt{1 - x_1^2}} = -\frac{4}{3}$

находим  $x_1 = \pm \frac{4}{5}$ ,  $x_2 = \pm \frac{3}{5}$ . Получаем две критических точки

$(\frac{4}{5}, \frac{3}{5})$  и  $(-\frac{4}{5}, -\frac{3}{5})$ . Для определения характера критических точек

нужно воспользоваться каким-либо достаточным условием экстремума.

При решении задачи нелинейного программирования в форме (1), (2) функция Лагранжа

$$F(\bar{x}, \bar{\lambda}) = f(\bar{x}) + \sum_{i=1}^m \lambda_i (b_i - \varphi_i(\bar{x})), \quad (8)$$

система необходимых условий экстремума

$$\begin{cases} \frac{\partial f}{\partial x_j} - \sum_{i=1}^m \lambda_i \frac{\partial \varphi_i}{\partial x_j} = 0, \\ \varphi_i(\bar{x}) \leq b_i, \\ \lambda_i (b_i - \varphi_i(\bar{x})) = 0, \\ \lambda_i \geq 0, \end{cases} \quad (9)$$

$$i = \overline{1, m}, \quad j = \overline{1, n}.$$

**Пример 2 [4].**

Найти минимум функции

$$f(\bar{x}) = 3x_1^2 + 4x_1x_2 + 5x_2^2$$

при ограничениях



$$G: \begin{cases} x_1 \geq 0, \\ x_2 \geq 0, \\ x_1 + x_2 \geq 4. \end{cases}$$

Решение.

Составим функцию Лагранжа

$$F(\bar{x}, \bar{\lambda}) = 3x_1^2 + 4x_1x_2 + 5x_2^2 + \lambda_1(-x_1) + \lambda_2(-x_2) + \lambda_3(4 - x_1 - x_2).$$

Система уравнений и неравенств (необходимые условия минимума):

$$\begin{cases} 6x_1 + 4x_2 - \lambda_1 - \lambda_3 = 0, \\ 4x_1 + 10x_2 - \lambda_2 - \lambda_3 = 0, \\ -x_1 \leq 0, \quad -x_2 \leq 0, \\ -x_1 - x_2 \leq -4, \\ \lambda_1 x_1 = \lambda_2 x_2 = 0, \\ \lambda_3(4 - x_1 - x_2) = 0, \\ \lambda_1, \lambda_2, \lambda_3 \geq 0, \end{cases}$$

$$i = \overline{1,3}, \quad j = \overline{1,2}.$$

Можно проверить, что необходимые условия выполняются при  $x_1 = 3, x_2 = 1$ . В точке  $A(3,1)$  функция  $f^* = 44$ . Заметим, что в граничных точках  $B_1(0,4)$  и  $B_2(4,0)$  значения  $f(B_1) = 80, f(B_2) = 48$ .

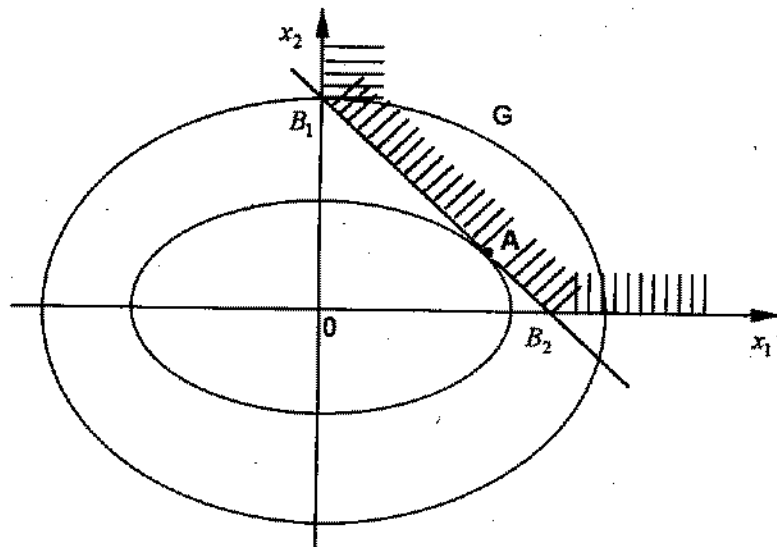


Рис. 10.3

## 10.2. Выпуклое программирование

Задача выпуклого программирования представляет частный случай задачи нелинейного программирования, когда функции  $f$  и  $\varphi_i$  ( $i = \overline{1, m}$ ) являются выпуклыми непрерывно-дифференцируемыми функциями. В этом случае система ограничений определяет некоторую выпуклую область в  $n$ -мерном векторном пространстве, и задача выпуклого программирования состоит в минимизации выпуклой функции в выпуклой области.

Особенность выпуклого программирования в том, что в задаче выпуклого программирования локальный минимум совпадает с глобальным.

В теории выпуклого программирования центральное место занимает так называемая теорема о седловой точке, представляющая собой обобщение классического метода множителей Лагранжа для определения экстремума при наличии ограничивающих условий на случай, когда последние содержат не только уравнения, но и неравенства.

Рассмотрим задачу выпуклого программирования

$$f(\bar{x}) \rightarrow \min, \quad (10)$$

$$b_i - \varphi_i(\bar{x}) \geq 0, \quad i = \overline{1, m}, \quad (11)$$

$$x_j \geq 0, \quad j = \overline{1, n}, \quad (12)$$

$$\bar{x} \in G \subset E_n,$$

$f, \varphi_i$  - выпуклые функции,

$G$  - выпуклое замкнутое множество.

### Определение 1.

Говорят, что множество решений задачи выпуклого программирования удовлетворяет условию регулярности (удовлетворяет условию Слейтера), если существует, по крайней мере, одна точка  $\bar{x}$  множества  $G$ , такая, что условия (11) выполняются как строгие неравенства, т. е. точка  $\bar{x}$  - внутренняя точка области  $G$ .

$$b_i - \varphi_i(\bar{x}) > 0. \quad (13)$$

### Определение 2.

Точка  $(\bar{x}_0, \bar{\lambda}^0)$  называется седловой точкой функции Лагранжа задачи выпуклого программирования, если

$$F(\bar{x}_0, \bar{\lambda}) \leq F(\bar{x}_0, \bar{\lambda}^0) \leq F(\bar{x}, \bar{\lambda}^0). \quad (14)$$

Справедлива теорема Куна-Таккера. [48]

**Теорема 1.**

Вектор  $\bar{x}_0$  является решением задачи выпуклого программирования (10) – (12), удовлетворяющей условию Слейтора, в том и только в том случае, если существует ненулевой вектор  $\bar{\lambda}^0 = (\lambda_1^0, \lambda_2^0, \dots, \lambda_m^0)$  ( $\lambda_i^0 \geq 0, i = \overline{1, m}$ ) такой, что точка  $(\bar{x}_0, \bar{\lambda}^0)$  - седловая точка функции Лагранжа.

Аналитические условия, определяющие седловую точку, в предположении, что функции  $f(\bar{x})$  и  $\varphi_i(\bar{x})$  непрерывно дифференцируемы, имеют вид:

$$\begin{cases} \frac{\partial F_0}{\partial x_j} \geq 0, \\ x_{0j} \frac{\partial F_0}{\partial x_j} = 0, \\ x_{0j} \geq 0, \quad j = \overline{1, n}, \end{cases} \quad (15)$$

$$\frac{\partial F_0}{\partial \lambda_i} \leq 0, \quad (16)$$

$$\lambda_i^0 \frac{\partial F_0}{\partial \lambda_i} = 0, \quad (17)$$

$$\lambda_i^0 \geq 0, \quad i = \overline{1, m}. \quad (18)$$

$$\lambda_i^0 \frac{\partial F_0}{\partial \lambda_i} = 0, \quad (19)$$

$$\lambda_i^0 \geq 0, \quad i = \overline{1, m}. \quad (20)$$

$\frac{\partial F_0}{\partial x_j}$  и  $\frac{\partial F_0}{\partial \lambda_i}$  - значения соответствующих частных производных функции Лагранжа в седловой точке.

В общем случае задачу нелинейного программирования решить не удастся. Теорема Куна-Таккера теоретически позволяет перейти от задачи выпуклого программирования к задаче поиска седловой точки функции Лагранжа  $F(\bar{x}, \bar{\lambda})$  на множестве, заданном простыми ограничениями (включающими условия неотрицательности  $\lambda_i$  и принадлежности  $\bar{x}$  к множеству G).

**Пример 3.**

Используя теорему Куна-Таккера решить задачу

$$f(x_1, x_2) = x_1 + x_2 \rightarrow \min,$$

$$\varphi_1(\bar{x}) = x_1^2 + x_2^2 - 18 \leq 0,$$

$$\varphi_2(\bar{x}) = x_1 \leq 0.$$

**Решение.**

Составим функцию Лагранжа

$$F(x_1, x_2, \lambda_1, \lambda_2) = f(x_1, x_2) + \lambda_1 \varphi_1(x_1, x_2) + \lambda_2 \varphi_2(x_1, x_2).$$

$$F(x_1, x_2, \lambda_1, \lambda_2) = f(x_1, x_2) + \lambda_1 (x_1^2 + x_2^2 - 18) + \lambda_2 x_1.$$

Необходимые условия экстремума (15) – (20)

$$\begin{cases} \frac{\partial F}{\partial x_1} + \lambda_1 \frac{\partial \varphi_1}{\partial x_1} + \lambda_2 \frac{\partial \varphi_2}{\partial x_1} = 0, \\ \frac{\partial F}{\partial x_2} + \lambda_1 \frac{\partial \varphi_1}{\partial x_2} + \lambda_2 \frac{\partial \varphi_2}{\partial x_2} = 0, \end{cases}$$

$$\begin{cases} \varphi_1(x_1, x_2) \leq 0, \\ \varphi_2(x_1, x_2) \leq 0, \\ \lambda_1 \varphi_1(x_1, x_2) = 0, \\ \lambda_2 \varphi_2(x_1, x_2) = 0 \end{cases}$$

$$1 + \lambda_1 2x_1 + \lambda_2 = 0, \quad (1)$$

$$1 + \lambda_1 2x_2 = 0, \quad (2)$$

$$x_1^2 + x_2^2 - 18 \leq 0, \quad (3)$$

$$x_1 \leq 0, \quad (4)$$

$$\lambda_1 (x_1^2 + x_2^2 - 18) = 0, \quad (5)$$

$$\lambda_2 x_1 = 0. \quad (6)$$

или

Пусть  $x_1^2 + x_2^2 = 18$  тогда  $\lambda_1 \geq 0$  (из (5)). Из условий (4) и (6) следует  $\lambda_2 = 0, x_1 < 0$ .

Из (1) и (2) получаем

$$\begin{cases} 1 + \lambda_1 2x_1 = 0, \\ 1 + \lambda_1 2x_2 = 0 \end{cases}$$

$$2\lambda_1 (x_1 - x_2) = 0.$$

или

$$2\lambda_1 (x_1 - x_2) = 0.$$

Так как  $\lambda_1 \neq 0$  ( $\lambda_2 = 0$ ), то  $x_1 = x_2$  и  $2x_1^2 = 18, x_1 = \pm 3, x_2 = \pm 3$ .

С учетом условий (4) и (2)  $x_1 = -3, x_2 = -3, \lambda_1 = \frac{1}{6}$ .

Итак, точка минимума  $x_1^* = x_2^* = -3, f^* = -6$ .

Пара точек  $\bar{x}^* = (-3; -3), \bar{\lambda}^* = (\frac{1}{6}; 0)$  является седловой

точкой функции Лагранжа.

## Геометрическая иллюстрация решения задачи

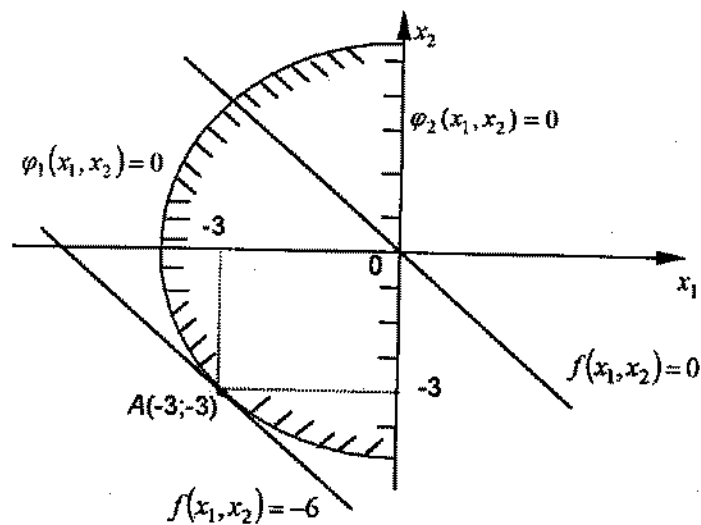


Рис. 10.4

### 10.3. Квадратичное программирование

#### Определение 3.

Квадратичной формой функции  $n$  переменных  $x_1, x_2, \dots, x_n$  называется функция вида

$$A = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j. \quad (21)$$

#### Определение 4.

Квадратичная форма называется положительно определенной, если  $A > 0$  и положительно полуопределенной - если  $A \geq 0$  при всех  $\bar{x} \neq 0$ .

#### Теорема 2 (критерий Сильвестра).

Для того, чтобы квадратичная форма (21) была положительно определенной, необходимо и достаточно, чтобы все «угловые миноры» матрицы  $A$

$$\Delta_1 = a_{11}, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \quad \dots \quad \Delta_n = |A| \quad (22)$$

были положительны.

Достаточные условия минимума определены теоремой.

#### Теорема 3.

Для того, чтобы дважды непрерывно-дифференцируемая функция  $f$  от переменных имела в критической (стационарной) точке  $\bar{x}$  минимум, достаточно, чтобы матрица ее вторых производных (матрица Гессе) была положительно определена.

Задача квадратичного программирования заключается в определении минимального значения выпуклой квадратичной функции на допустимом множестве, заданном линейными ограничениями

$$f(\bar{x}) = \frac{1}{2} (Q\bar{x}, \bar{x}) + (\bar{r}, \bar{x}) \rightarrow \min, \quad (23)$$

$$\sum_{i=1}^m a_{ij} x_j \leq b_i, \quad i = \overline{1, m}, \quad (24)$$

$$x_j \geq 0, \quad j = \overline{1, n}.$$

$$Q = (q_{ij}) = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_n^2} \end{vmatrix}, \quad (25)$$

$\bar{r} = (r_1, r_2, \dots, r_n)$  - заданный вектор.

Согласно теореме Куна-Таккера точка минимума  $\bar{x}$  целевой функции (23) на допустимом множестве, определяемом условиями (24), может быть найдена как решение системы уравнений с дополнительными переменными  $\lambda_i, x_{n+i}, i = \overline{1, m}, \mu_j, j = \overline{1, n}$ :

$$\sum_{i=1}^m q_{ij} x_i + r_j + \sum_{i=1}^m \lambda_i a_{ij} - \mu_j = 0, \quad (26)$$

$$\sum_{j=1}^n a_{ij} x_j + x_{n+i} = b_i, \quad (27)$$

$$\lambda_i x_{n+i} = 0, \quad (28)$$

$$\mu_j x_j = 0, \quad (29)$$

$$x_j \geq 0, \quad (30)$$

$$\lambda_i \geq 0, \quad (31)$$

$$j = \overline{1, n}, \quad i = \overline{1, m}.$$

Система линейных уравнений (26) - (29) в силу специфичности уравнений (28) - (29) не всегда может быть решена методами

решения линейных систем. Решение может быть получено симплексным методом.

Для нахождения начального базисного решения системы уравнений (26) – (27) воспользуемся методом искусственного базиса, причем не следует включать в базисные одновременно переменные  $\lambda_i$  и  $x_{n+i}$  с одним и тем же индексом  $i$  и переменные  $x_j$ ,  $\mu_j$  с индексом  $j$ .

Строим функцию

$$g(x) = x_{n+1} + x_{n+2} + \dots + x_{n+m}.$$

Требуется найти минимум функции  $g(x)$  при ограничениях вида

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n + x_{n+i} = b_i, \\ i = \overline{1, m}, \quad x_k \geq 0, \quad k = \overline{1, n+m}$$

( $a_{ik}$ ,  $b_i$  те же, что и в (27)).

Переменные  $x_{n+1}, x_{n+2}, \dots, x_{n+m}$  можно принять за базисные, точка  $(0, 0, \dots, b_1, b_2, \dots, b_m)$  – начальное базисное решение симплексного метода.

В результате получаем точку

$$\bar{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}, x_{n+1}^{(0)}, \dots, x_{n+m}^{(0)}),$$

минимизирующую  $g(x)$  при ограничениях.

Обозначим  $\alpha = g_{\min}$ .

Справедливо утверждение: если  $\alpha = 0$ , то  $\bar{x}^{(0)}$  – начальное базисное решение для множества  $G$ .

#### Пример 4.

Решить задачу квадратичного программирования

$$f(\bar{x}) = x_1^2 + x_2^2 - 10x_1 - 15x_2 \rightarrow \min,$$

при ограничениях

$$G: \begin{cases} 2x_1 + 3x_2 \leq 13, \\ 2x_1 + x_2 \leq 10, \\ x_1, x_2 \leq 0. \end{cases}$$

Решение.

Матрица  $Q(x_1, x_2) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$  – матрица вторых производных, по-

ложительно - определенная, следовательно  $f(\bar{x})$  – выпуклая.

Составим функцию Лагранжа

$$F(\bar{x}, \bar{\lambda}, \bar{\mu}) = x_1^2 + x_2^2 - 10x_1 - 15x_2 + \lambda_1(2x_1 + 3x_2 - 13) + \\ + \lambda_2(2x_1 + x_2 - 10) - \mu_1x_1 - \mu_2x_2.$$

Система (26) – (31) принимает вид

$$2x_1 + 2\lambda_1 + 2\lambda_2 - \mu_1 = 10, \quad (1)$$

$$2x_2 + 3\lambda_1 + 2\lambda_2 - \mu_2 = 15, \quad (2)$$

$$2x_1 + 3x_2 + x_3 = 13, \quad (3)$$

$$2x_1 + x_2 + x_4 = 10, \quad (4)$$

$$\mu_1x_1 = 0, \quad \mu_2x_2 = 0, \quad \lambda_1x_3 = 0, \quad \lambda_2x_4 = 0. \quad (5)$$

Будем искать решение задачи методом искусственного базиса.

Уравнения (3) и (4) разрешим относительно переменных  $x_3$  и  $x_4$  соответственно. В уравнения (1) и (2) введем вспомогательные переменные  $x_5, x_6$ . Переменные  $x_3, x_4$  – дополнительные.

Начальный базис составляют переменные  $x_3, x_4, x_5, x_6$ . Вспомогательную целевую функцию  $S = x_5 + x_6$  выразим через свободные переменные  $x_1, x_2, \lambda_1, \lambda_2, \mu_1, \mu_2$ .

$$S = 25 - 2x_1 - 2x_2 - 5\lambda_1 - 3\lambda_2 + \mu_1 + \mu_2.$$

Последовательные симплекс-таблицы приведены ниже. Рамками обведены опорные элементы; те свободные переменные, которые на данном шаге нельзя переводить в базисные из-за условий (5), обведены кружками. Самый правый столбец содержит свободные члены условий, последняя строка – коэффициенты минимизируемой функции  $S$ .

Таблица 1

	$x_1$	$x_2$	$\lambda_1$	$\lambda_2$	$\mu_1$	$\mu_2$	
$x_5$	2	0	2	2	-1	0	10
$x_6$	0	2	3	1	0	-1	15
$x_3$	2	3	0	0	0	0	13
$x_4$	2	1	0	0	0	0	10
	-2	-2	-5	-3	1	1	-25

В последней строке таблицы 1 есть отрицательные коэффициенты (при  $x_1$  и  $x_2$ ). Выберем переменную  $x_5 = x_2$ . Для перевода в базисные находим  $\min\left(\frac{10}{1}; \frac{13}{3}; \frac{15}{2}; \infty\right) = \frac{13}{3}$ , следовательно,  $x_1 = x_3$  переводим в свободные переменные.

Правила преобразования элементов таблицы.

- Поменять местами  $x_5$  и  $x_1$ .
- На место опорного элемента поставить обратную ему величину.
- Образовать ведущую строку, записав в нее соответствующие элементы исходной строки, деленные на опорный элемент.
- В ведущем (разрешающем) столбце (содержащим опорный элемент) поставить со знаком минус соответствующие элементы исходного столбца, деленные на опорный элемент.
- Оставшиеся свободные места в новой симплекс таблице заполняются следующим образом: из строки элементов исходной таблицы вычесть произведение ее элемента из разрешающего столбца на уже заполненную разрешающую строку новой таблицы.

Таблица 2

	$x_1$	$x_3$	$\lambda_1$	$\lambda_2$	$\mu_1$	$\mu_2$	
$x_5$	2	0	2	2	-1	0	10
$x_6$	-4/3	-2/3	3	1	0	-1	19/3
$x_2$	2/3	1/3	0	0	0	0	13/3
$x_4$	4/3	2/3	0	0	0	0	17/3
	-2/3	-2/3	-5	-3	1	1	-49/3

В последней строке среди отрицательных коэффициентов выбираем коэффициент (-5), следовательно,  $x_5 = \lambda_1$ . Далее определяем  $x_1 = x_6$ . Формируем следующую таблицу.

Таблица 3

	$x_1$	$x_3$	$x_6$	$\lambda_2$	$\mu_1$		$\mu_2$
$x_5$	26/9	4/9	-2/3	4/3	-1	2/3	52/9
$\lambda_1$	-4/9	-2/9	1/3	1/3	0	-1/3	19/9
$x_2$	2/3	1/3	0	0	0	0	13/3
$x_4$	4/3	2/3	0	0	0	0	17/3
	-26/9	-4/9	5/3	-4/3	-1	-2/3	-52/9

Таблица 4

	$x_5$	$x_3$	$x_6$	$\lambda_2$	$\mu_1$	$\mu_2$	
$x_1$	9/26	2/13		6/13	-9/26	3/13	2
$\lambda_1$	2/13	-2/13		7/13	2/13	-3/13	3
$x_2$	-3/13	3/13		12/39	-3/13	-2/13	3
$x_4$	6/13	6/13		-8/13	-6/13	6/13	3
	1	0		0	0	0	0

В последней строке таблицы 4 нет отрицательных коэффициентов и значение  $S=0$ . Следовательно, минимум достигается в

точке базиса, соответствующего этой таблице. Искомое решение задачи квадратичного программирования имеет вид:

$$\bar{x}^{opt} = (x_1; x_2) = (2; 3), \quad f^{opt} = f(\bar{x}^{opt}) = -52.$$

**Пример 5 [66].**

Решить задачу квадратичного программирования

$$f(\bar{x}) = x_1^2 + 2x_2^2 - 2x_1x_2 - 2x_1 - 6x_2 \rightarrow \min,$$

$$G: \begin{cases} x_1 + x_2 \leq 2, \\ -x_1 + 2x_2 \leq 2, \\ x_1, x_2 \geq 0. \end{cases}$$

**Решение.**

Матрица  $Q(x_1, x_2) = \begin{vmatrix} 2 & -2 \\ -2 & 4 \end{vmatrix}$  - матрица вторых производных,

положительно - определена, следовательно  $f(\bar{x})$  - выпуклая.

Область  $G$  - на рис. 5. Функция Лагранжа

$$F(x, \lambda, \mu) = x_1^2 + 2x_2^2 - 2x_1x_2 - 2x_1 - 6x_2 + \lambda_1(x_1 + x_2 - 2) + \lambda_2(-x_1 + 2x_2 + 2) - \mu_1x_1 - \mu_2x_2.$$

Система (26) - (31) примет вид

$$\begin{cases} 2x_1 - 2x_2 + \lambda_1 - \lambda_2 - \mu_1 = 2, \\ -2x_1 + 4x_2 + \lambda_1 + 2\lambda_2 - \mu_2 = 6, \\ x_1 + x_2 + x_3 = 2, \\ -x_1 + 2x_2 + x_4 = 2, \\ \mu_1x_1 = \mu_2x_2 = \lambda_3x_3 = \lambda_4x_4 = 0, \\ x_1, x_2, x_3, x_4 \geq 0, \quad \mu_1, \mu_2 \geq 0, \quad \lambda_1, \lambda_2 \geq 0. \end{cases} \quad (32)$$

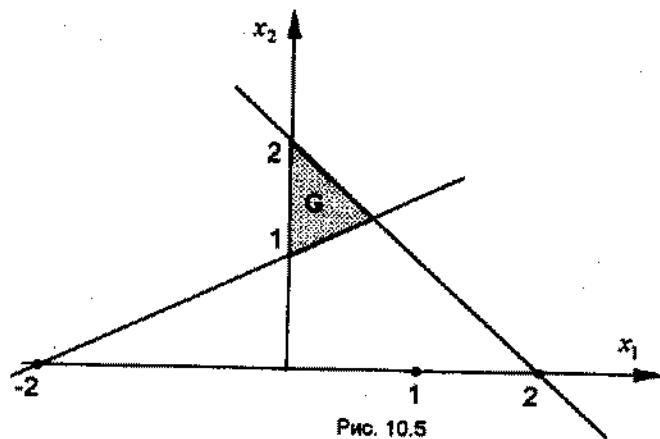


Рис. 10.5

Вспомогательные переменные -  $x_3, x_4$ . Для получения начального базисного решения дополнительные переменные  $x_5, x_6$  входят только в первые два уравнения (32). Базисные переменные -  $x_3, x_4, x_5, x_6$ .

Симплексный метод приводит к решению задачи

$$x_1^* = \frac{4}{5}, \quad x_2^* = \frac{6}{5}, \quad f^*(x_1^*, x_2^*) = \frac{36}{5}.$$

Поскольку  $f(\bar{x})$  - выпуклая функция, то вычисленный минимум функции является глобальным минимумом в рассматриваемой области.

Если задача квадратичного программирования содержит или ограничения в виде равенств, следует их преобразовать к виду (23).

#### Алгоритм решения задачи квадратичного программирования

1. Составить функцию Лагранжа.
2. Составить необходимые и достаточные условия существования седловой точки для функции Лагранжа ((26)-(31)).
3. Используя метод искусственного базиса, либо установить отсутствие седловой точки, либо найти ее координаты.
4. Записать оптимальное решение исходной задачи и вычислить значение целевой функции.

### 10.4. Численные методы решения задачи нелинейного программирования

В практике решения задач нелинейного программирования чаще всего используются: градиентные методы (методы спуска) и методы штрафных и барьерных функций.

#### 10.4.1. Прямые методы

Интерес представляют методы, в которых используются только значения функции - методы прямого поиска. Таким методом является модифицированный метод Хука-Дживса [4], являющийся модификацией метода прямого поиска для учета ограничений. Предлагается проверка того, каждая ли точка, полученная в процессе поиска, принадлежит области ограничений. В случае, если точка «вышла» из области, целевой функции присваивается очень большое значение, что гарантирует осуществление поиска снова в допустимой области. Метод позволяет в некоторых случаях успеш-

но решать задачу, однако с помощью этого метода невозможно проводить поиск вдоль границы области ограничений. В случае выхода на границу алгоритм выдает граничную точку в качестве решения.

Другим методом прямого поиска является итерационный комплексный метод, изложенный в [4]. Комплексный метод применим к достаточно широкому кругу задач с ограничениями. Для выпуклых целевых функций в выпуклой области он дает точку локального минимума. Модификации в формировании начальной точки позволяют найти и точку глобального минимума.

Некоторые из наиболее простых методов прямого поиска основаны на подходе Монте-Карло. Исследуемые точки генерируются датчиком псевдослучайных чисел. После серии исследований допустимых случайных точек определяется минимум исследуемой функции при заданных ограничениях.

Пример такой процедуры – процедура Луус и Якола [74]. Пусть на  $k$ -й итерации известны текущие приближения  $\bar{x}^{(k-1)}$  точки минимума и вектор  $\bar{g}^{(k-1)}$ ,  $i$ -й элемент которого определяет область изменения соответствующей координаты вектора  $\bar{x}$ .

Генерируют  $j$  псевдослучайных чисел  $r_i (i=1, 2, \dots, j)$ , равномерно распределенных на интервале  $(-0.5; 0.5)$ , и среди точек  $\bar{z}^{(i)}$  вида  $\bar{z}^{(i)} = \bar{x}^{(k-1)} + r_i \bar{g}^{(k-1)}$  выбирается допустимая точка  $\bar{z}^{(s)}$ , в которой целевая функция принимает наименьшее значение.

Полагают  $\bar{x}^{(k)} = \bar{z}^{(s)}$ . Новый вектор  $\bar{g}^{(k)}$  вычисляется по формуле  $\bar{g}^{(k)} = (1 - \varepsilon) \bar{g}^{(k-1)}$ ,  $0 < \varepsilon < 1$ .

Метод надежен для локальных минимумов, однако требует большого числа испытаний.

Этот класс методов мало чувствителен к трудностям, возникающим в случаях разрывности производных целевой функции либо ошибки вычислений функции.

Характерными и существенными недостатками этих методов являются:

- сходимость большинства этих методов не доказана;
- поиск, как правило, приводит в достаточно хорошую, но не оптимальную точку;
- проблема критерия останова счета.

Критерий останова чаще всего основывается на результатах счета, вследствие чего, либо проводится бесполезный счет, либо счет может остановиться далеко от решения задачи.

В случаях, если применимы градиентные методы, последние эффективнее.

## 10.4.2. Методы спуска

Метод возможных направлений

Метод предназначен для поиска экстремума (минимума) при наличии ограничений типа неравенств и заключается в построении последовательных приближений к точке минимума  $\bar{x}^*$  целевой функции  $f(\bar{x})$  на допустимом множестве  $G$ :

$$\bar{x}_{k+1} = \bar{x}_k + \alpha_k \bar{e}_k, \quad k = 0, 1, \dots, \quad (33)$$

$$\bar{x}_0 \in G, \quad \alpha_k > 0.$$

Вектор  $\bar{e}_k$  – вектор возможного направления, т.е. такого, что для достаточно малых  $\alpha_k > 0$  точка  $\bar{x}_{k+1} \in G$ ,  $f(\bar{x}_{k+1}) < f(\bar{x}_k)$  – целевая функция убывает.

### Укрупненный алгоритм решения задачи

1. Определить набор  $I_k$  номеров ограничений задачи, для которых в текущей точке  $\bar{x}_k$  выполнены неравенства

$$-\varepsilon_k < \varphi_i(\bar{x}_k) \leq 0; \quad (34)$$

2. Решить задачу линейного программирования  $\min_{z, \bar{e}} z$ ,

$$f'(\bar{x}_k, \bar{e}) \leq z; \quad (35)$$

$$\varphi'_i(\bar{x}_k, \bar{e}) \leq z, \quad i \in I_k$$

$$|\bar{e}^i| \leq 1, \quad i = \overline{1, n}.$$

Если решение  $\bar{z}_k$  меньше  $(-\varepsilon_k)$ , величину  $\varepsilon_{k+1}$  для следующей итерации положить равной  $\varepsilon_k$ , иначе –  $\varepsilon_{k+1} = q\varepsilon_k$ ,  $0 < q < 1$ .

3. Определить точку  $\bar{x}_{k+1} = \bar{x}_k + \alpha_k \bar{e}_k$  в которой достигается минимум функции  $f(\bar{x}_k + \alpha_k \bar{e}_k)$  по неотрицательным  $\alpha$ , удовлетворяющим неравенствам  $\varphi_i(\bar{x}_k + \alpha_k \bar{e}_k) \leq 0$ ,  $i = \overline{1, m}$ .

Алгоритм начинает работу с некоторой допустимой точки  $\bar{x}_0$  и с произвольной  $\varepsilon_0$ . Если функции  $f$  и  $\varphi_i$  выпуклы и выполнено условие Слейтера, то полученное решение и есть искомое.

Условиями окончания поиска минимума могут быть равенство нулю с заданной степенью точности  $E$  градиента функции, либо условие

$$|f(\bar{x}_{k+1}) - f(\bar{x}_k)| < E, \quad (36)$$

$\epsilon > 0$  – достаточное малое число, характеризующее точность численного решения. При этом полагают

$$\bar{x}^* \equiv \bar{x}_k, \quad f^* = f(\bar{x}^*).$$

Детальное описание алгоритма возможных направлений приведено в [48].

## II. Метод симплексных направлений (метод условного градиента)

Требуется найти минимум выпуклой непрерывно дифференцируемой функции  $f(\bar{x})$ .

Для решения этой задачи оказывается полезной идея замены нелинейной целевой функции в некоторой окрестности исследуемой точки линейной функцией и сведение исходной нелинейной задачи к последовательному решению задач линейного программирования.

Пусть  $\bar{x}_k$  – некоторая точка области допустимых решений, причем  $f'(\bar{x}_k) \neq 0$ .

Построим вспомогательную функцию

$$F_k(\bar{x}) = \sum_{j=1}^n c_j x_j, \quad (37)$$

$$c_j = \frac{\partial f(\bar{x}_k)}{\partial x_j}, \quad j = \overline{1, n} \quad (38)$$

и решим задачу минимизации функции (37) на множестве  $G$ .

Эта задача – задача нелинейного программирования. В некоторых случаях ее решение не представляет затруднений.

1) Допустимое множество  $G$  задано линейными ограничениями, тогда задача может быть решена методами линейного программирования;

2) Допустимое множество –  $n$ -мерный параллелепипед:

$$(G = \{ x \in E_n \mid a_j \leq x_j \leq b_j, \quad j = \overline{1, n} \}.$$

$$\bar{x}_j^{(k)} = \begin{cases} a_j, & \text{если } \frac{\partial f(\bar{x}_k)}{\partial x_j} > 0, \\ b_j, & \text{если } \frac{\partial f(\bar{x}_k)}{\partial x_j} < 0, \\ \frac{a_j + b_j}{2}, & \text{если } \frac{\partial f(\bar{x}_k)}{\partial x_j} = 0. \end{cases} \quad (39)$$

3) Допустимое множество – шар радиуса  $R_0$  с центром в точке  $y^{(0)}$

$$G = \left\{ x \in E_n \mid \sum_{j=1}^n (x_j - y_j^{(0)})^2 \leq R_0^2 \right\}$$

Тогда

$$\bar{x}_k = y^{(0)} - R_0 \frac{f'(\bar{x}_k)}{\|f'(\bar{x}_k)\|}. \quad (40)$$

Пусть  $\bar{z}_k$  – решение вспомогательной задачи.

Следующее приближение  $\bar{x}_{k+1}$  к точке минимума найдем по формуле

$$\bar{x}_{k+1} = \bar{x}_k + \alpha_k (\bar{z}_k - \bar{x}_k), \quad \alpha \in (0; 1), \quad (41)$$

причем в силу выпуклости допустимого множества  $\bar{x}_{k+1} \in G$ .

Параметр  $\alpha$  может быть определен различными способами. Укажем два из них:

В качестве параметра  $\alpha$  выбрать наименьший положительный корень уравнения  $\frac{df(\bar{x}_k + \alpha(\bar{z}_k - \bar{x}_k))}{d\alpha} = 0$ , если он не больше едини-

цы; в противном случае положить  $\alpha = 1$ ,  $\bar{x}_{k+1} = \bar{z}_k$ .

Положить  $\alpha = 1$ , проверить условия

$$f(\bar{x}_{k+1}) < f(\bar{x}_k). \quad (42)$$

Если условие нарушено, дроблением  $\alpha$  добиться выполнения условия (42).

В точке  $\bar{x}_{k+1}$  вычислить новую линейную форму и минимизировать ее при тех же ограничениях. Условие окончания поиска минимума – условие (36).

### Алгоритм метода условного градиента

1. Определить исходное допустимое решение ( $\bar{x}_0$ ).
2. Определить вектор градиента функции (38) в точке допустимого решения.
3. Сформировать функцию (37) и решить вспомогательную задачу в области  $G$ .
4. Определить шаг – параметр  $\alpha$ .
5. Определить очередное приближение (41) – компоненты нового допустимого решения.
6. Проверить необходимость перехода к последующему решению. В случае необходимости – переход к п.2.
7. В противном случае положить  $\bar{x}^* = \bar{x}_k$ ,  $f^* = f(\bar{x}^*)$ .



### Пример 6.

Решить задачу нелинейного программирования

$$f(\bar{x}) = x_1^2 + 4x_2^2 - 8x_1 - 8x_2 \rightarrow \min,$$

$$G: \begin{cases} -2 \leq x_1 \leq 2, \\ 0 \leq x_2 \leq 3 \end{cases}$$

методом условного градиента, завершая вычисления, при выполнении условия  $\|\bar{x}^{(k-1)} - \bar{x}^{(k)}\| \leq 0.1$ .

Решение.

Пусть начальное приближение  $\bar{x}^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ,  $\bar{x} \in G$ .

Найдем  $\bar{f}' = \text{grad } f(\bar{x}) = (2x_1 - 8; 8x_2 - 8)$  в точке  $\bar{x}^{(0)}$ ,  $f'_{(\bar{x}^{(0)})} = (-6; 0)$ .

Запишем вспомогательную задачу

$$f_k(\bar{x}) = -6x_1 - 0 \cdot x_2 \rightarrow \min,$$

$$\begin{cases} -2 \leq x_1 \leq 2, \\ 0 \leq x_2 \leq 3 \end{cases}$$

Эта задача линейного программирования может быть решена симплекс-методом.

В данном случае воспользуемся тем, что область допустимых решений позволяет найти  $\tilde{\bar{x}}^{(0)} = (2; 1.5)$ . (из формулы (39)).

$$x_1^{(1)} = 1 + \alpha \cdot (2 - 1), \quad x_2^{(1)} = 1 + \alpha(1.5 - 1) = 1 + 0.5\alpha.$$

Из условия  $\Phi'(\alpha) = 0$  найдем значение  $\alpha = \alpha_0^*$ .

$$\begin{aligned} \Phi(\alpha) &= f[\bar{x}^{(0)} + \alpha(\tilde{\bar{x}}^{(0)} - \bar{x}^{(0)})] = \\ &= (1 + \alpha)^2 + 4(1 + 0.5\alpha)^2 - 8(1 + \alpha) - 8(1 + 0.5\alpha) = 2\alpha^2 - 6\alpha - 11. \end{aligned}$$

$$\alpha_0^* = 1.5.$$

Принимаем  $\alpha_0 = \min(1; 1.5) = 1$ , получим  $x_1^{(1)} = 2$ ,  $x_2^{(1)} = 1.5$ .

Так как  $\|\bar{x}^{(0)} - \bar{x}^{(1)}\| = 1.5 > 0.1$ , требуемая точность не достигнута.

Таблица 5

$k$	$\bar{x}^{(k)}$	$\ \bar{x}^{(k-1)} - \bar{x}^{(k)}\ $	$\tilde{\bar{x}}^{(k)}$	$\alpha_k$
0	(1; 1)		(2; 1.5)	1
1	(2; 1.5)	1.5	(2; 0)	0.333
2	(2; 1.001)	0.499	(2; 0)	$0.99 \cdot 10^{-3}$
3	(2; 1)	0		

$$x^* = (2; 1), \quad f^* = f(x^*) = -16.$$

## 10.4.3. Методы преобразования

Методы преобразования – это методы решения задачи нелинейного программирования – определения

$$\min f(\bar{x}), \quad \bar{x} \in E_n,$$

или ограничениях

$$g_i(\bar{x}) \geq 0, \quad i = 1, m \quad (42)$$

сведением ее к одной или нескольким задачам безусловной оптимизации некоторых вспомогательных функций.

Предполагается, что функции  $f(\bar{x})$ ,  $g_i(\bar{x})$  непрерывно дифференцируемы. Допустимое множество решений задачи обозначается  $G$ , его внутренность  $G_0 = \{\bar{x} : g(\bar{x}) > 0\}$ ,  $G_0 = \{\bar{x} : g(\bar{x}) > 0\}$ .

Общий вид вспомогательной функции

$$W(\bar{x}, \bar{r}) = f(\bar{x}) + v(g(\bar{x}), \bar{r}), \quad (43)$$

где  $\bar{r}$  – вектор управляющих переменных,  $v$  – «штрафующая» функция.

Различные методы преобразования отличаются друг от друга выбором функции  $v$ .

Возможны два подхода, позволяющие найти решение исходной задачи посредством безусловной минимизации  $W(\bar{x}, \bar{r})$  по  $\bar{x}$ .

Определение минимума  $\bar{x}(r)$  методами последовательной безусловной оптимизации.

Выбор такой гладкой функции  $v$ , что при некотором  $\bar{r}^*$  для любого  $\bar{r} \geq \bar{r}^*$  точка  $\bar{x}(\bar{r})$  является локальным минимумом исходной задачи (42). Этот подход детально обсуждается в [74].

Рассмотрим последовательные методы – метод барьерных функций и метод штрафных функций.

### 10.4.3.1. Барьерные функции

В методе барьерных функций (методе внутренних штрафов)  $(g(\bar{x}), r)$  выбираются таким образом, чтобы «построить барьер» на границе допустимого множества, препятствующий нарушению ограничений в процессе безусловной минимизации. Во внутренних точках (множество  $G_0$  должно быть непустым) функции  $W(\bar{x}, \bar{r})$  и  $f(\bar{x})$  мало отличаются, и в то же время при приближении к границе множества функция  $v$  неограниченно возрастает.

Определим функцию барьера

$$v(g(\bar{x}), r) = r \sum_{i=1}^m \varphi_i(g_i(\bar{x})), \quad (44)$$

$r > 0$  – некоторое число.

Обозначим  $\varphi_i(g_i(\bar{x})) = \varphi_i(r)$  – непрерывные при  $r > 0$ ,  $\varphi_i(r) \rightarrow \infty$  при  $r \rightarrow 0$ . Предполагается, что функции  $\varphi_i(r)$  дифференцируемы.

Последовательность функций  $\{\varphi_k(\bar{x})\}$  называется последовательностью барьерных функций множества, если выполняются условия:

- $\lim_{k \rightarrow \infty} \varphi_k(\bar{x}) = 0$  для любой фиксированной внутренней точки  $\bar{x}$  множества  $G$ ,
- $\lim_{r \rightarrow \infty} \varphi_k(\bar{x}(r)) = \infty$  для любой последовательности  $\{\bar{x}(r)\}$  внутренних точек  $G_0$ , сходящейся к какой-либо граничной точке этого множества.

Наиболее распространенными функциями барьера являются

$$\varphi_i(g_i(\bar{x})) = (g_i(\bar{x}))^{-1} \quad (45)$$

и

$$\varphi_i(g_i(\bar{x})) = -\ln(g_i(\bar{x})). \quad (46)$$

Вспомогательная функция

$$B(\bar{x}, r) = f(\bar{x}) + r \sum_{i=1}^m \varphi_i(g_i(\bar{x})), \quad r > 0 \quad (47)$$

определена в  $G_0$  и неограниченно возрастает, если  $g_i(\bar{x}) \rightarrow 0$  для какого-нибудь  $i$

**Алгоритм метода барьерных функций [74]**

1. Положить  $k = 0$ .
2. Задать точку  $\bar{x}^{(0)} \in G_0$  (внутреннюю точку множества).
3. Выбрать монотонно убывающую, сходящуюся к нулю последовательность  $\{r^{(k)}\}$ . Например,  $r^{(k)} = 10^{1-k}$ ,  $k = 1, 2, \dots$
4. Положить  $k = k + 1$ .
5. Начиная с точки  $\bar{x}^{(k-1)}$ , найти безусловный минимум для  $B(\bar{x}, r^{(k)})$  по  $\bar{x}$  и положить  $\bar{x}^{(k)} = \bar{x}(r^{(k)})$ .
6. Проверить выполнение условий останова.

Если условия выполнены – конец алгоритма.

В противном случае – возврат к п.2

Доказательство сходимости метода внутренних штрафных функций приведено в [46].

Достоинства и недостатки методов барьерных функций.

1) При условии непрерывности функции и существования ее локальных минимумов в замыкании множества  $G_0$  выполнены следующие соотношения:

$$a) \lim_{k \rightarrow \infty} r^{(k)} \sum_{i=1}^m \varphi_i(g_i(\bar{x}^{(k)})) = 0,$$

$$b) \lim_{k \rightarrow \infty} f(\bar{x}^{(k)}) = f(\bar{x}^*) \text{ и } f(\bar{x}^{(k)}) \text{ монотонно убывает,}$$

$$в) \lim_{k \rightarrow \infty} B(\bar{x}^{(k)}, r^{(k)}) = f(\bar{x}^*),$$

$$г) \left\{ \sum_{i=1}^m \varphi_i(g_i(\bar{x}^{(k)})) \right\} \text{ монотонно возрастает.}$$

2) Методами барьерных функций удается находить решения, не удовлетворяющие условиям Куна-Таккера.

3) Методы барьерных функций позволяют вычислять оценки вектора множителей Лагранжа  $\bar{\lambda}^*$ .

4) Оценка отклонения от оптимума

$$f(\bar{x}(r)) - f(\bar{x}^*) \leq \sum_{i=1}^m \lambda_i(r) g_i(\bar{x}(r)), \quad (48)$$

$$\lambda_i(r) = -r \frac{d\varphi_i(g_i(\bar{x}(r)))}{dg_i}, \quad i = \overline{1, m}. \quad (49)$$

5) Алгоритм обладает некоторой устойчивостью по отношению к ошибкам округления  $\bar{x}(r)$  в случае, если барьерная функция – логарифмическая (46).

К недостаткам методов барьерных функций можно отнести:

1) Методы неприменимы для отыскания точки минимума, реализующейся в изолированной точке допустимого множества.

2) Нет конструктивной идеи (правила) для построения последовательности  $\{\bar{x}^{(k)}\}$ .

Практически используют формулу

$$r^{(k+1)} = r^{(k)} \cdot 10^{-1}, \quad k = 1, 2, \dots \quad (50)$$

$r^{(1)}$  выбирают, как правило, так, чтобы

$$f(\bar{x}^{(0)}) \approx r^{(1)} \sum_{i=1}^m \varphi_i(g_i(\bar{x}^{(0)})). \quad (51)$$

3) Плохая обусловленность матрицы Гессе, и, как следствие, медленная сходимость в окрестности точки  $\bar{x}^*$ . В окрестности

границы допустимого множества (если  $x^*$  принадлежит границе) барьерные функции имеют «овражный» характер.

4) Применение обычных методов безусловной оптимизации дает удовлетворительные результаты лишь при начальных шагах, когда точка достаточно далеко от границы.

5) Нужна специальная процедура спуска по направлению, включающая проверку соблюдения ограничений задачи.

6) Возникают трудности с аппроксимацией барьерных функций, учитывающей особенности поведения барьерной функции на границе допустимой области.

7) При уменьшении  $r$  барьерная функция приобретает явно выраженную овражную структуру, и норма вектора градиента в окрестности  $\bar{x}(r)$  достаточно велика, что делает невозможным получить точку с близкими к нулю величинами частных производных, что во многих методах безусловной оптимизации является критерием остановки алгоритма.

### Пример 7.

Методом барьерных функций решить задачу нелинейного программирования.

$$f(\bar{x}) = \frac{1}{3}x_1^3 + x_2,$$

$$g_1(\bar{x}) = x_1 \geq 0,$$

$$g_2(\bar{x}) = x_2 \geq 0.$$

**Решение.**

Используем последовательность барьерных функций (45), вспомогательная функция

$$B(\bar{x}, r) = \frac{1}{3}x_1^3 + x_2 + \frac{r}{x_1} + \frac{r}{x_2}.$$

В данном случае (в силу линейности ограничений) можно найти минимум из условий, определяющих стационарную точку:

$$\begin{cases} \frac{\partial B}{\partial x_1} = 0, \\ \frac{\partial B}{\partial x_2} = 0. \end{cases} \quad \begin{cases} x_1^2 - \frac{r}{x_1^2} = 0, \\ 1 - \frac{r}{x_2^2} = 0, \end{cases} \quad \text{откуда следует}$$

$$\begin{cases} x_1 = \sqrt[3]{r}, \\ x_2 = \sqrt{r}. \end{cases}$$

Последовательные значения  $(x_1, x_2)$  при различных значениях параметра  $r$  приведены в таблице.

$r$	1	$10^{-2}$	$10^{-4}$	$10^{-6}$
$(x_1^{(k)}, x_2^{(k)})$	(1; 1)	(0.316; 0.1)	(0.1; 0.01)	( $10^{-2}$ ; $10^{-4}$ )

Полагаем  $\bar{x}^* = (0; 0)$ ,  $f^*(\bar{x}^*) = 0$ .

### Пример 8.

Вычислить три итерации решения задачи нелинейного программирования

$$f(x_1, x_2) = x_2^2 - 2x_2 - x_1 \rightarrow \min,$$

$$G: 4 - x_1^2 - x_2^2 \geq 0$$

методом барьерных функций.

**Решение.**

Примем  $x_1^{(0)} = 1.5$ ,  $x_2^{(0)} = 0.5$ , точка  $\bar{x}^{(0)} = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \end{pmatrix}$  - внутренняя

точка области  $G$ .

Штрафная вспомогательная функция

$$B(x_1, x_2, r) = x_2^2 - 2x_2 - x_1 + r(4 - x_1^2 - x_2^2)^{-1}.$$

Задачу безусловной минимизации функции  $B(\bar{x}, r)$  решаем методом градиентного спуска с дроблением шага. На каждой итерации необходимо проверять принадлежность полученной точки области  $G$ .

$$\begin{cases} x_1^{(k+1)} = x_1^{(k)} - \lambda \cdot \frac{\partial B}{\partial x_1}, \\ x_2^{(k+2)} = x_2^{(k)} - \lambda \cdot \frac{\partial B}{\partial x_2}, \end{cases} \quad k = 0, 1, \dots$$

$$\frac{\partial B}{\partial x_1} = \frac{-2x_1 \cdot r}{(4 - x_1^2 - x_2^2)^2} - 1, \quad \frac{\partial B}{\partial x_2} = \frac{-2x_2 \cdot r}{(4 - x_1^2 - x_2^2)^2} + 2x_2 - 2.$$

В точке начального приближения  $\frac{\partial B}{\partial x_1} = -2.2$ ,  $\frac{\partial B}{\partial x_2} = -1.44$ .

Пусть  $\lambda = 0.25$ , тогда  $\begin{cases} x_1^{(1)} = 1.5 - 0.25(-2.2) = 2.05, \\ x_2^{(1)} = 0.5 - 0.25(-1.44) = 0.86. \end{cases}$

Полученная точка  $\begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \end{pmatrix}$  не удовлетворяет ограничению

$4 - x_1^2 - x_2^2 \geq 0$ , следовательно нужно уменьшить шаг  $\lambda$ . В результа-

те получим при  $\lambda = 0.1$  точку с координатами  $\begin{cases} x_1^{(0)} = 1.72, \\ x_2^{(0)} = 0.644, \end{cases}$  принадлежащую области  $G$ .

Решение приведено в таблице 7.

Таблица 7

$k$	$x_1$	$x_2$	$\frac{\partial B}{\partial x_1}$	$\frac{\partial B}{\partial x_2}$	$\lambda$	$r$	$f(x_1, x_2)$
0	1.5	0.5	-2.2	-1.44	0.1	1	-2.25
1	1.72	0.644	-1.8754	-0.0308	0.05	0.1	-2.5933
2	1.813	0.696	-1.6933	-0.8742	0.015	0.01	-2.7206
3	1.8384	0.709					-2.7537

### 10.4.3.2. Штрафные функции

Последовательность функций  $\{\varphi_k(\bar{x})\}$ , определенных на множестве  $E_n$  и обладающих свойством

$$\lim_{k \rightarrow \infty} \varphi_k(\bar{x}) = \begin{cases} 0, & \text{если } \bar{x} \in G, \\ +\infty, & \text{если } \bar{x} \notin G, \end{cases}$$

называется последовательностью штрафных функций на множестве  $G$ ,

$$\varphi_k(\bar{x}) = k\varphi(\bar{x}) = \frac{1}{r}\varphi(\bar{x}), \quad k = 1, 2, \dots \quad (52)$$

Примеры штрафных функций:

- Ограничения типа равенств:
  - квадратичная функция штрафа

$$\varphi_i(g_i(\bar{x})) = (\min(0, g_i(\bar{x})))^2, \quad (53)$$

б) функция Зангвилла

$$\varphi_i(g_i(\bar{x})) = -\min(0, g_i(\bar{x})). \quad (54)$$

- Ограничения типа неравенств:

$$\text{а) } \varphi_i(g_i(\bar{x})) = (g_i(\bar{x}))^2, \quad (55)$$

$$\text{б) } \varphi_i(g_i(\bar{x})) = |g_i(\bar{x})|. \quad (56)$$

Штрафная вспомогательная функция имеет вид

$$P(\bar{x}, r) = f(\bar{x}) + \frac{1}{r} \sum_{i=1}^m \varphi_i(g_i(\bar{x})), \quad r > 0. \quad (57)$$

При достаточно малых  $r$  ( $k \rightarrow \infty$ ) полагают

$$\bar{x}^* \equiv \bar{x}^{(k)}, \quad f^* = f(\bar{x}^*),$$

$\bar{x}^{(k)}$  - последовательность решений задач безусловной минимизации функции (57).

Критерием достижения требуемой точности  $E$ , решения задачи (42) может служить неравенство

$$\|\bar{x}^{(k)} - \bar{x}^{(k/2)}\| \leq E, \quad (58)$$

$$k = 2l, \quad l \in Z.$$

Алгоритм решения задачи такой же, как в методе барьерных функций. Отличие в формировании вспомогательной функции и выборе начальной точки.

В случае, если  $f(\bar{x})$  - выпуклая квадратичная функция, ограничения - линейные функции, решение задачи минимизации функции (57) можно найти решением системы уравнений, определяющих стационарную точку.

Достоинства и недостатки методов штрафных и барьерных функций аналогичны.

Сравнение методов.

1) В методах барьерных функций (внутренних штрафных функций) гарантировано соблюдение ограничений задачи на протяжении всего процесса ее решения.

2) В методах штрафных (внешних) функций начальное приближение - любая точка вне области допустимых решений, в методах барьерных функций - только внутренняя точка.

3) В методах барьерных функций решение может быть получено в любой момент времени и получено допустимое решение задачи, что невозможно при использовании методов внешних штрафных функций.

4) Применяя какой-либо градиентный метод для поиска минимума внутренней штрафной функции, на каждой его итерации нужно вычислять производные для всех ограничений; при минимизации вспомогательной функции методом внешнего штрафа требуется вычисление производных только для нарушенных в текущей точке ограничений.

5) Если задача допускает решение любым из рассмотренных методов, предпочтительнее пользоваться методами внешних штрафных функций, прежде всего в силу п.2. Доказательство сходимости последовательности точек минимума вспомогательных барьерных и штрафных функций к глобальному решению задачи условной оптимизации приведено в [11], [48].

**Пример 9.**

Решить задачу нелинейного программирования

$$f(\bar{x}) = x_1^2 + x_2^2 - 20x_1 - 30x_2 \rightarrow \min,$$

$$2x_1 + 3x_2 - 13 \leq 0,$$

$$2x_1 + x_2 - 10 \leq 0$$

методом внешних штрафных функций.

*Решение.*

Проверим, что целевая функция является выпуклой квадратичной функцией. Запишем матрицу вторых производных

$$f''(\bar{x}) = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix}.$$

Угловые миноры матрицы положительны при  $\bar{x} \in E_2$ , следовательно, функция  $f(\bar{x})$  - выпуклая.

Перепишем ограничения в виде (42)

$$g_1(x_1, x_2) = 13 - 2x_1 - 3x_2 \geq 0,$$

$$g_2(x_1, x_2) = 10 - 2x_1 - x_2 \geq 0.$$

Решение задачи минимизации вспомогательной функции

$$P(\bar{x}, r) = x_1^2 + x_2^2 - 20x_1 - 30x_2 + \frac{1}{r}(-2x_1 - 3x_2 + 13)^2 + \frac{1}{r}(-2x_1 - x_2 + 10)^2$$

может быть найдено из условия  $P'(\bar{x}) = 0$ .

Получаем систему уравнений

$$\begin{cases} \frac{\partial P}{\partial x_1} = 2x_1 - 20 + \frac{4}{r}(2x_1 + 3x_2 - 13) + \frac{4}{r}(2x_1 + x_2 - 10) = 0, \\ \frac{\partial P}{\partial x_2} = 2x_2 - 30 + \frac{6}{r}(2x_1 + 3x_2 - 13) + \frac{2}{r}(2x_1 + x_2 - 10) = 0. \end{cases}$$

Пусть  $k = \frac{1}{r}$ , перепишем систему  $\begin{cases} (1+8k)x_1 + 8kx_2 = 46k+10, \\ 8kx_2 + (1+10k)x_2 = 49k+15. \end{cases}$

Получаем

$$\begin{cases} x_1 = \frac{68k^2 + 26k + 10}{16k^2 + 18k + 1}, \\ x_2 = \frac{24k^2 + 89k + 15}{16k^2 + 18k + 1}, \end{cases} \quad k = 1, 2, \dots$$

$$g_1(x_1, x_2) = -(2x_1 + 3x_2 - 13) = \frac{-85k - 52}{16k^2 + 18k + 1} < 0,$$

$$g_2(x_1, x_2) = -(2x_1 + x_2 - 10) = \frac{+39k - 25}{16k^2 + 18k + 1} > 0.$$

Вспомогательная штрафная функция (т.к.  $g_2(x_1, x_2) > 0$ )

$$P(\bar{x}, r) = x_1^2 + x_2^2 - 20x_1 - 30x_2 + k(2x_1 + 3x_2 - 13)^2.$$

Решая систему

$$\begin{cases} \frac{\partial P}{\partial x_1} = 0 = x_1(1+4k) + 6kx_2 - 10 - 26k, \\ \frac{\partial P}{\partial x_2} = 0 = 6kx_1 + (1+9k)x_2 - 15 - 39k, \end{cases} \quad \text{получим}$$

$$\begin{cases} x_1 = \frac{26k+10}{13k+1}, \\ x_2 = \frac{39k+15}{13k+1}. \end{cases} \quad g_1(\bar{x}) = \frac{32}{13k+1} > 0 \quad \text{при } k = 1, 2, \dots$$

Найденные  $x_1$  и  $x_2$  определяют точку безусловного минимума  $\bar{x}^{(k)}$  вспомогательной функции  $P(\bar{x}, r)$ .

Определяем точку минимума функции исходной задачи:

Определяем точку минимума функции исходной задачи:

$$x_1^* = \lim_{k \rightarrow \infty} \frac{26k+10}{13k+1} = 2,$$

$$x_2^* = \lim_{k \rightarrow \infty} \frac{39k+15}{13k+1} = 3.$$

Окончательно получаем  $\bar{x}^* = (2; 3)$ ,  $f^* = -117$ .

#### 10.4.4. Метод Фиакко-Маккормика решения задачи оптимизации с ограничениями

Метод предназначен для минимизации функции

$$P(\bar{x}, r) = f(\bar{x}) + \frac{1}{r} \sum_{i=1}^m \varphi_i(g_i(\bar{x})), \quad r > 0. \quad (59)$$

##### Алгоритм метода Фиакко-Маккормика [4]

1. Задать начальные данные и начальную допустимую точку

$$\bar{x} = \bar{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_N^{(0)}).$$

2. Вычислить очередное приближение  $\bar{x}^{(k)}$  к точке минимума методом ДФП.

3. Если  $\bar{x}^{(k)}$  - искомое решение, конец алгоритма.

В противном случае положить  $r_{k+1} = \frac{r_k}{c}$ ,  $k = k+1$ ,  $\bar{x} = \bar{x}^{(k)}$  и

переход к п. 2.

Задавая начальное значение  $r = r_0$ , можно найти минимум

$P(\bar{x}, r)$  методом ДФП. Уменьшая значение  $r$ , полагают  $r_1 = \frac{r_0}{c}$ ,

где  $c > 1$  — некоторая произвольная константа. К примеру, полагая  $c = 10$  или  $c = 12$  вновь найти минимум  $P(\bar{x}, r)$ .

На  $k$ -ом шаге найдем минимальное значение  $P(\bar{x}, r)$ , в точке  $\bar{x}_k^*$ , которая может быть использована в качестве первой точки в итерационной процедуре поиска минимума функции

$P(\bar{x}, r_{k+1})$ ,  $r_{k+1} = \frac{r_k}{c}$ . Последовательность  $r_k$  убывает и стремится к

нулю при  $k \rightarrow \infty$ , т.е. последовательность минимумов будет сходиться к решению задачи с ограничениями.

Некоторые пояснения к рассматриваемой процедуре минимизации.

- Метод предполагает наличие начальной допустимой точки.
- Метод ДФП, применяемый для безусловной одномерной оптимизации, использует кубическую интерполяцию целевой функции. Необходимо, чтобы в процессе одномерного поиска точки оставались в допустимой области.

Применяют следующий прием. Пусть имеется точка  $\bar{p}$  и направление поиска  $\bar{d} = -H\bar{g}$  (см. метод ДФП).

Следующая точка для кубической интерполяции  $\bar{q} = \bar{p} + \lambda\bar{d}$ . Полагают  $\lambda = 2$  и проверяют, выполняются ли ограничения задачи в точке  $\bar{q}$ . Если ограничения не выполняются, полагают  $\lambda = \frac{\lambda}{a}$  и вновь находят точку  $\bar{q}$  и проверяют ограничения (в противном случае —  $\lambda = 2$ ). В конечном счете, точка  $\bar{q}$  будет найдена. В алгоритме необходимо предусмотреть принятие таких мер, чтобы в процессе одномерной оптимизации минимум находился между двумя допустимыми точками. Параметр  $a$  ( $a = 1.05$ ) выбран эвристическим способом.

- Условие окончания поиска минимума формируют исходя из заданной точности.
- Необходимо дополнить алгоритм минимизации процедурами, задающими минимизируемую функцию, условия ограничения и компоненты градиента целевой функции.

## 11. СОБСТВЕННЫЕ ЗНАЧЕНИЯ И СОБСТВЕННЫЕ ВЕКТОРЫ

Задачи нахождения собственных значений и соответствующих им собственных векторов возникают в самых разных научных областях. Например, при анализе динамических систем собственные значения определяют частоты колебаний, а собственные векторы характеризуют их форму.

В электро-радиотехнических устройствах собственные значения матриц определяют характеристические постоянные времени и периоды работы этих устройств.

Рассмотрим квадратную матрицу  $n$ -го порядка:

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

Собственные значения  $\lambda_i$  ( $i = \overline{1, n}$ ) квадратной матрицы  $A$  — действительные и/или комплексные числа, удовлетворяющие условию:

$$(A - \lambda E)\bar{X} = 0, \quad (1)$$

$E$  — единичная матрица,

$\bar{X} = \{x_1, x_2, \dots, x_n\}$  — собственный вектор матрицы  $A$ , соответствующий некоторому собственному значению  $\lambda$ .

Определитель матрицы

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

является многочленом  $n$ -ой степени относительно  $\lambda$  и называется характеристическим многочленом.

Корни этого многочлена — собственные значения.

Некоторые свойства собственных значений и векторов:

• Все  $n$  собственных значений любой симметричной матрицы  $(a_{ij} = a_{ji}; i, j = 1, 2, \dots, n)$  вещественны.

• Собственные векторы, отвечающие различным собственным значениям симметричной матрицы, ортогональны:

$$\begin{aligned} (\bar{X}^{(i)}, \bar{X}^{(j)}) &= 0, \text{ при } i \neq j, \\ (\bar{X}^{(i)}, \bar{X}^{(i)}) &= 1, \text{ при } i = j. \end{aligned} \quad (3)$$

• Собственный вектор матрицы, умноженный на произвольное число, также является собственным вектором.

• Подобные матрицы

$$A \text{ и } B = P^{-1}AP, \quad (4)$$

где  $P$  — неособая матрица, имеют одинаковые собственные значения, их собственные вектора связаны соотношением:

$$\bar{X} = P\bar{Y}. \quad (5)$$

Характеристическое уравнение

$$\lambda^n - p_1\lambda^{n-1} - p_2\lambda^{n-2} - \dots - p_n = 0 \quad (6)$$

решается ранее изложенными методами решения нелинейных уравнений. Однако задача осложняется тем, что среди собственных значений часто встречаются кратные. Кроме того, для произвольной матрицы непросто вычислить сами коэффициенты характеристического многочлена.

Ряд задач требует нахождения только наибольшего или наименьшего собственных значений. В общем случае ставится задача о нахождении всех собственных значений и собственных векторов, то есть полная проблема собственных значений.

Алгоритмы решения можно разбить на два вида:

• итерационные, которые просты в применении и хорошо приспособлены для отыскания наибольших и наименьших собственных чисел,

• алгоритмы, реализующие методы преобразования для решения полной проблемы собственных значений.

### 11.1. Метод итерации

Для решения частичной проблемы собственных значений (отыскания наибольших и наименьших собственных чисел), применяется метод простой итерации решения системы уравнений

$$\lambda \bar{X} = A \bar{X}.$$

Итерационный процесс удовлетворяет соотношению

$$\lambda^{(k+1)} \cdot \bar{X}^{(k+1)} = A \cdot \bar{X}^{(k)}, \quad k = 0, 1, \dots, \quad (7)$$

номер итерации,

- нормированный собственный вектор

Выбранный произвольный нормированный вектор  $\bar{X}^{(0)}$  подставляем в (7) и получаем после умножения слева на  $A$  некоторый вектор.

После нормировки этот вектор представится в виде

$$\lambda^{(1)} \cdot \bar{X}^{(1)}, \text{ где } \bar{X}^{(1)} - \text{нормированный вектор. Вновь подставим}$$

в (7), найдем  $\lambda^{(2)}$  и  $\bar{X}^{(2)}$ .

Итерационный процесс продолжается до установления определенных с некоторой точностью значений  $\lambda$  и  $\bar{X}$ .

Собственный вектор, отвечающий наибольшему по модулю собственному значению данной матрицы  $A$ , определяется из равенства

$$\bar{X}_1 \approx \bar{Y}^{(k+1)}. \quad (8)$$

#### Алгоритм метода итерации

1. Выбор начального нормированного собственного вектора

2.  $(\bar{X}^{(0)})^T = (x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)})$ ,  $n$  - размерность вектора.

3.  $\bar{Y}^{(0)} = \bar{X}^{(0)}$ .

4. Вычисление

$$\bar{Y}^{(k+1)} = A \cdot \bar{Y}^{(k)}, \quad k = 0, 1, \dots, \quad (9)$$

- номер итерации.

5. Вычисление  $\lambda_i^{(k+1)} = \frac{y_i^{(k+1)}}{y_i^{(k)}}$ ,  $(i = 1, 2, \dots, n)$  (10)

$y_i^{(k+1)}$  и  $y_i^{(k)}$  - одноименные координаты двух последовательных векторов.

6. Если  $|\lambda_m^{(k+1)} - \lambda_{\text{max}}^{(k+1)}| > E$ ,  $(m = 2, 3, \dots, n)$ , (11)

переход к пункту 3.

7. В противном случае положить:

$$\lambda = \lambda_1^{(k+1)},$$

собственный вектор  $\bar{X}_1$ ,

нормированный вектор -  $\bar{Y}^{(k+1)}$ .

$$\bar{X}_1^T = (x_1^{(1)}, x_2^{(1)}, \dots, x_n^{(1)}),$$

$$x_i^{(1)} = \frac{y_i^{(k+1)}}{y_{\max}}, \quad y_{\max} = \max_i |y_i^{(k+1)}|.$$

Скорость сходимости процесса итераций (7) зависит от близости начального вектора  $\bar{X}^{(0)}$  к истинному.

В случае, если необходимо найти наименьшее значение  $\lambda$ , систему  $\lambda \bar{X} = A \bar{X}$  приводят к виду  $\frac{1}{\lambda} \bar{X} = A^{-1} \bar{X}$ .

Затем методом итерации находят собственное значение  $\frac{1}{\lambda}$  (что будет достигнуто при наименьшем  $\lambda$ ).

Заметим, что собственные значения матриц  $A^{-1}$  и  $A$  обратны друг другу.

### Пример 1.

Используя метод итерации, определить первое (наибольшее) собственное значение и первый собственный вектор матрицы

$$A = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Решение.

$$\text{Выбираем } \bar{Y}^{(0)} = \bar{X}^{(0)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

$$\text{Вычисляем } \bar{Y}^{(1)} = A \cdot \bar{Y}^{(0)},$$

$$\bar{Y}^{(1)} = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 4 \\ 2 \end{pmatrix}.$$

$$\lambda_1^{(1)} = \frac{5}{1} = 5, \quad \lambda_2^{(1)} = \frac{4}{1} = 4, \quad \lambda_3^{(1)} = \frac{2}{1} = 2.$$

$$\bar{Y}^{(2)} = A \cdot \bar{Y}^{(1)},$$

$$\bar{Y}^{(2)} = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 24 \\ 15 \\ 6 \end{pmatrix}.$$

$$\lambda_1^{(2)} = \frac{24}{5} = 4.8, \quad \lambda_2^{(2)} = \frac{15}{4} = 3.75, \quad \lambda_3^{(2)} = \frac{6}{2} = 3.$$

Дальнейшие вычисления сведены в таблицу 1.

Таблица 1

	4	1	0	$y_1^{(k+1)}$	$y_2^{(k+1)}$	$y_3^{(k+1)}$
$A$	1	2	1	$y_1^{(k)}$	$y_2^{(k)}$	$y_3^{(k)}$
	0	1	1			
$\lambda_1$	1	1	1			
$\lambda_2$	5	4	2	4.8	3.75	3
$\lambda_3$	24	15	6	4.625	4.00	3.5
$\lambda_4$	111	60	21	4.540	4.20	3.86
$\lambda_5$	504	252	81	4.500	4.32	4.111
$\lambda_6$	2268	1089	333	4.480	4.388	4.270
$\lambda_7$	10161	4779	1422	4.470	4.424	4.361
$\lambda_8$	45423	21141	6201	4.465	4.443	4.409
$\lambda_9$	202833	93936	27342	4.463	4.450	4.434
$\lambda_{10}$	905238	417987	121248	4.462	4.456	4.447
$\lambda_{11}$	4038939	1862460	539235	4.461	4.458	4.454
$\lambda_{12}$	18018220	8303094	2401695	4.461	4.459	4.457
$\lambda_{13}$	80375960	37026100	10704790	4.460	4.460	4.459
$\lambda_{14}$	358553000	165132900	47730890	4.460	4.460	4.460
$\lambda_{15}$	1599253000	736526700	212863800			

Итак,  $\lambda_1 = 4.460$ . Определим собственный вектор.

$$\bar{X}_1^T = \bar{Y}_{14}^T = (1599253000; 736526700; 212863800).$$

Нормированный собственный вектор  $\bar{X}_1^T = (1; 0.46; 0.13)$ .

### Замечание.

Чтобы избежать потери точности при вычислении координат векторов  $\bar{Y}$ , рекомендуется нормировать вектора на каждой итерации. Структурная схема алгоритма метода итерации на рис. 1.



Укажем модификацию метода итерации.

В [23] показано, что вектор  $\bar{Y}^{(m)} = A^m \cdot \bar{Y}$  приближенно представляет собой собственный вектор матрицы  $A$ , соответствующий наибольшему собственному значению  $\lambda_1$ . Начальный вектор  $\bar{Y}^{(0)}$  выбирается произвольно.

Справедливы соотношения

$$\bar{Y}^{(m)} = A^m \bar{Y}, \quad \bar{Y}^{(m+1)} = A \bar{Y}^{(m)}, \quad (12)$$

где  $m$  - степень матрицы.

$$\lambda_1 \approx \frac{y_i^{(m+1)}}{y_i^{(m)}} \quad (m=1, 2, \dots, n). \quad (13)$$

**Пример 2** [23].

Найти наибольшее собственное значение и соответствующий ему собственный вектор матрицы

$$A = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

**Решение.**

Выбираем начальный вектор  $\bar{Y}^{(0)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ .

Составляем таблицу

Таблица 2

$y$	$Ay$	$A^2y$	$A^3y$	$A^4y$	$A^5y$	$A^6y$	$A^7y$	$A^8y$	$A^9y$	$A^{10}y$
1	5	24	111	504	2268	10161	45433	202833	905238	4038939
1	4	15	60	252	1089	4779	21141	93906	417987	1862460
1	2	6	21	81	333	1422	6201	27342	121248	539235

Остановившись на итерациях  $A^9 \bar{Y} = \bar{Y}^{(9)}$  и  $A^{10} \bar{Y} = \bar{Y}^{(10)}$ , получаем

$$\frac{y_1^{(10)}}{y_1^{(9)}} = \frac{4038939}{905238} = 4.462, \quad \frac{y_2^{(10)}}{y_2^{(9)}} = \frac{1862460}{417987} = 4.456,$$

$$\frac{y_3^{(10)}}{y_3^{(9)}} = \frac{539235}{121248} = 4.447.$$

Принимаем  $\lambda_1 \approx \frac{1}{3}(4.462 + 4.456 + 4.447) \approx 4.46$ .

Собственный вектор  $\bar{X}_1^T = (4038939; 1862460; 539235)$ .

Нормированный вектор  $\bar{X}_1^T = (1; 0.46; 0.13)$ .

Более быстрый итерационный процесс основан на приближенном равенстве

$$\lambda_1 \approx \frac{SpA^{m+1}}{SpA^m}, \quad (14)$$

где  $SpA^{m+1}$  - "след" матрицы  $A^{m+1}$ ,  $SpA = \sum_{i=1}^n a_{ii}$ .

определим

$$A^2 = \begin{pmatrix} 17 & 6 & 1 \\ 6 & 6 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad A^4 = \begin{pmatrix} 326 & 141 & 37 \\ 141 & 81 & 30 \\ 37 & 30 & 14 \end{pmatrix},$$

$$A^6 = \begin{pmatrix} 127526 & 58497 & 16810 \\ 58497 & 27342 & 8067 \\ 16810 & 8067 & 2465 \end{pmatrix},$$

$$A^9 = \begin{pmatrix} 127526 & 58497 & 16810 \\ 58497 & 27342 & 8067 \\ 16810 & 8067 & 2465 \end{pmatrix} \cdot \begin{pmatrix} 4 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 568601 & 261330 & 75307 \\ 261330 & 121248 & 35409 \\ 75307 & 35409 & 10532 \end{pmatrix}.$$

$$SpA^9 = 568601 + 121248 + 10532 = 700381,$$

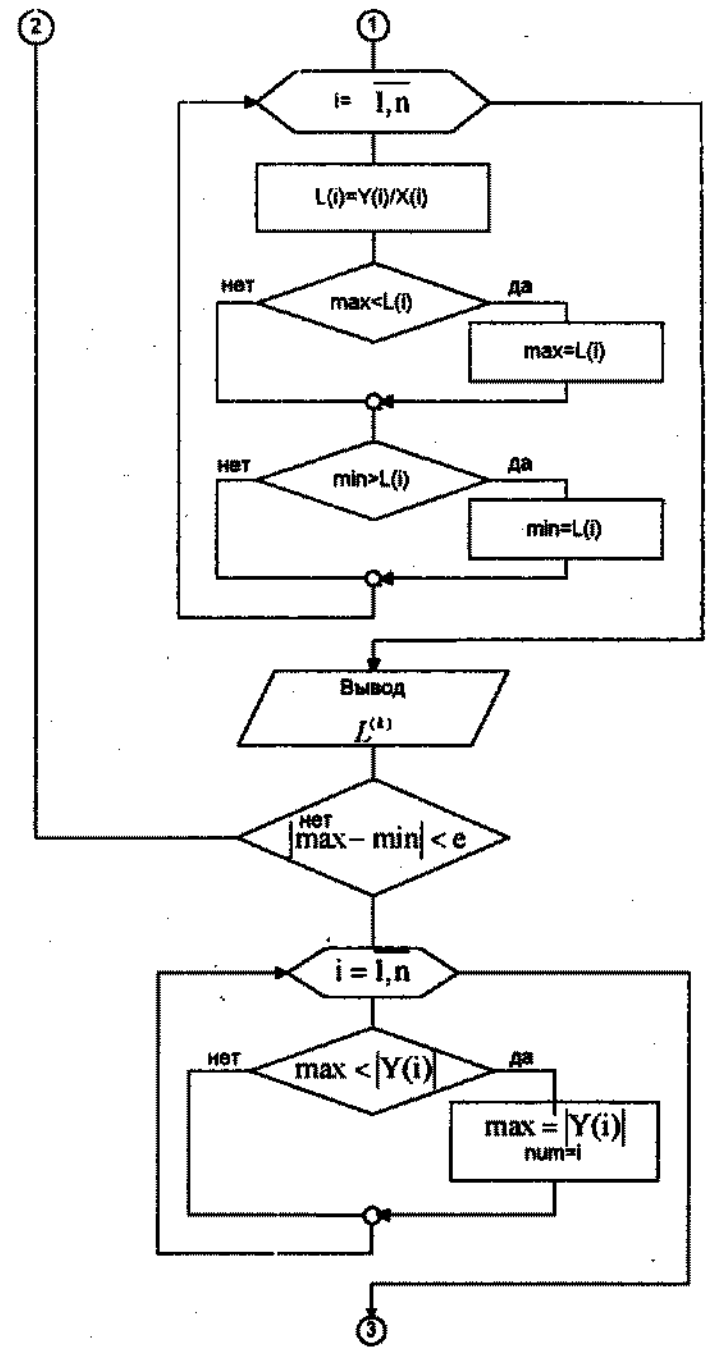
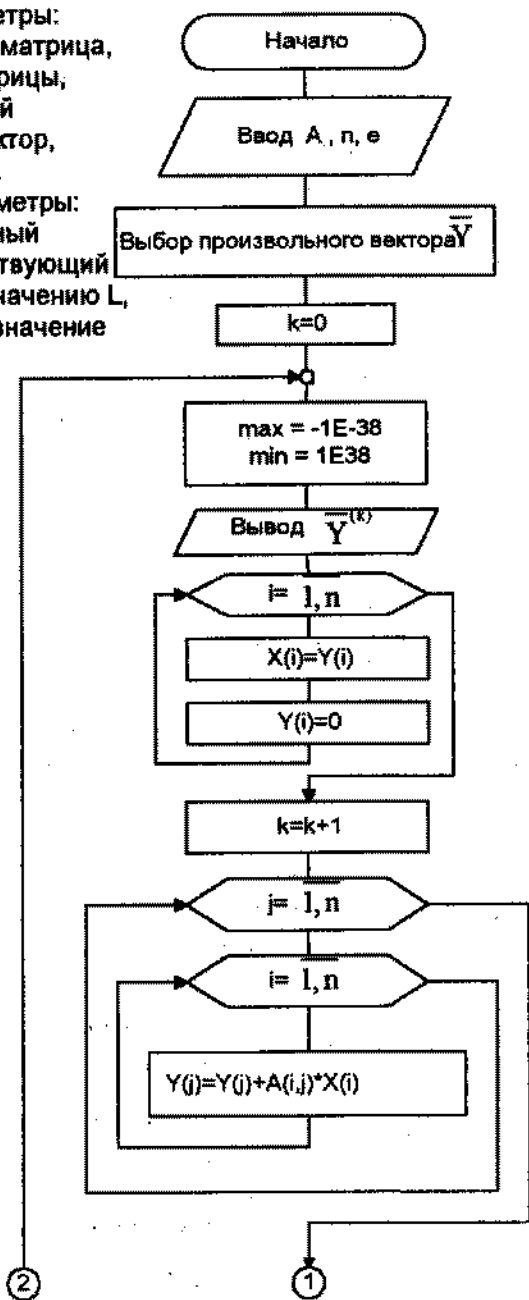
$$SpA^8 = 127526 + 27342 + 2465 = 157333,$$

$$\lambda_1 = \frac{SpA^9}{SpA^8} \approx 4.45.$$

Структурная схема алгоритма модифицированного метода итерации на рис. 2.

### Структурная схема алгоритма метода итерации

Входные параметры:  
 $A()$  – исходная матрица,  
 $n$  – порядок матрицы,  
 $Y()$  – начальный собственный вектор,  
 $e$  – погрешность.  
 Выходные параметры:  
 $X1()$  – собственный вектор, соответствующий  
 собственному значению  $L$ ,  
 $L$  – собственное значение матрицы  $A$ .



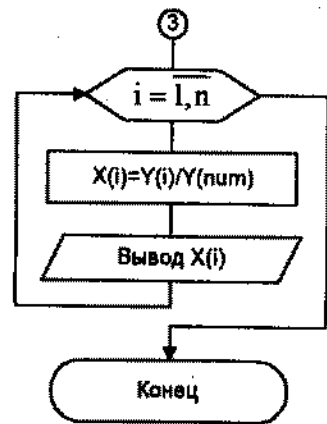


Рис. 11.1

**Структурная схема алгоритма модифицированного метода итерации**

Входные параметры:  
 $A()$  – исходная матрица;  
 $n$  – порядок матрицы;  
 $\epsilon$  – погрешность.  
 Выходные параметры:  
 $L$  – собственное значение матрицы

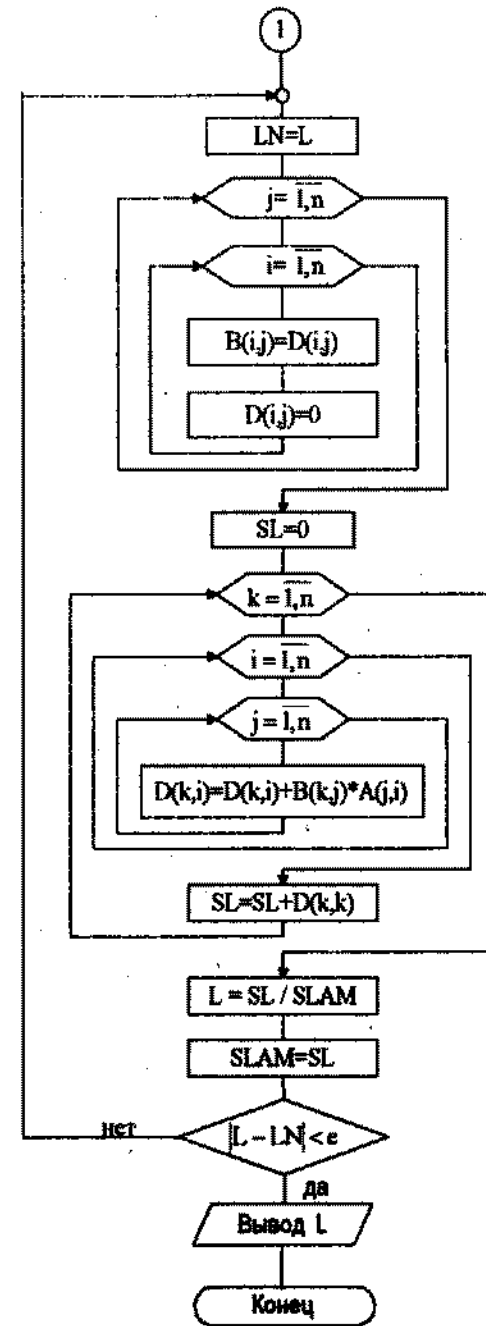
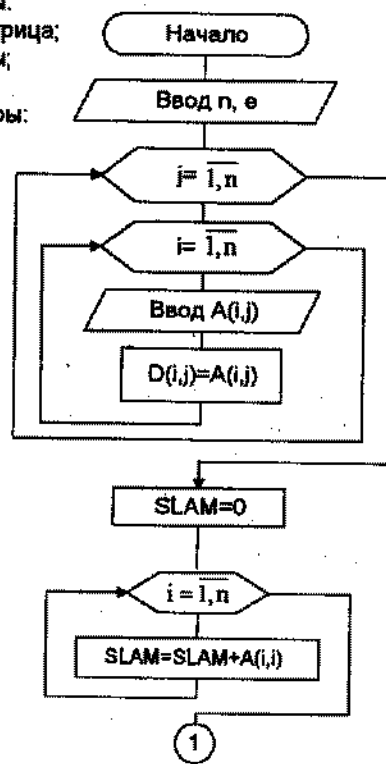


Рис. 11.2

## 11.2. Метод Левеерье-Фаддеева

В основе методов преобразования матрицы лежит стремление получить из исходной матрицы  $A$  некоторую другую, более простую, но с теми же собственными значениями. Оптимальным упрощением является приведение  $A$  к диагональному виду. Однако большинство методов позволяют привести матрицу  $A$  к трехдиагональной форме.

Коэффициенты характеристического уравнения

$$\lambda^n - p_1 \lambda^{n-1} - p_2 \lambda^{n-2} - \dots - p_n = 0 \quad (6)$$

вычисляются посредством построения последовательности матриц:

$$A_1 = A, \quad p_1 = \text{Sp}A_1, \quad B_1 = A_1 - p_1 E, \quad (15)$$

$$A_2 = AB_1, \quad p_2 = \frac{\text{Sp}A_2}{2}, \quad B_2 = A_2 - p_2 E,$$

$$\dots \dots \dots$$

$$A_n = AB_{n-1}, \quad p_n = \frac{\text{Sp}A_n}{n},$$

где  $\text{Sp}A = \sum_{i=1}^n a_{ii}$  - "след" матрицы.  $\lambda_i (i=\overline{1, n})$  - решение уравнения (6).

Собственные векторы  $\overline{X}_i$ , определяют по формуле

$$\overline{X}_i = \lambda_i^{n-1} \overline{e} + \lambda_i^{n-2} \overline{b}_1 + \dots + \lambda_i \overline{b}_{n-2} + \overline{b}_{n-1}, \quad (16)$$

$\overline{e}$  - какой-либо единичный вектор,

$\overline{b}_1, \overline{b}_2, \dots, \overline{b}_{n-1}$  - одноименные с  $\overline{e}$  вектора матриц  $B_1, B_2, \dots, B_{n-1}$ .

Последовательно вычисляем вектора для  $\lambda = \lambda_1, \lambda_2, \dots, \lambda_n$ .

$$\overline{u}_{0i} = \overline{e}, \quad \overline{u}_{1i} = \lambda \overline{u}_{0i} + \overline{b}_1, \quad \overline{u}_{2i} = \lambda \overline{u}_{1i} + \overline{b}_2, \dots,$$

$$\overline{u}_{n-1,i} = \overline{X}_i = \lambda \overline{u}_{i-2,i} + \overline{b}_{n-1}. \quad (17)$$

**Основные пункты алгоритма метода Левеерье-Фаддеева**

1. Ввод исходной матрицы  $A_{n \times n}$ ,  $n$  - размерность матрицы.
2. Вычисление коэффициентов уравнения (6)  $p_i, i=\overline{1, n}$ .
3. Решение характеристического уравнения (определение  $\lambda_i, i=\overline{1, n}$ ).
4. Задание начальных единичных векторов  $\overline{U}$ .
5. Вычисление собственного вектора  $\overline{X}_i$ , соответствующего  $\lambda_i$ .

нормировка вектора  $\overline{X}_i$ .

6. Конец алгоритма.

п. 1 - 6 повторяются для  $i=\overline{1, n}$ .

Структурная схема алгоритма метода Левеерье-Фаддеева на рис. (3) - (8).

**Пример 3.**

Используя метод Левеерье-Фаддеева, найти собственные числа и собственные векторы матрицы

$$A = \begin{pmatrix} 2.2 & 1 & 0.5 & 2 \\ 1 & 1.3 & 2 & 1 \\ 0.5 & 2 & 0.5 & 1.6 \\ 2 & 1 & 1.6 & 2 \end{pmatrix}$$

**Решение.**

Коэффициенты уравнения:  $\lambda^4 - p_1 \lambda^3 - p_2 \lambda^2 - p_3 \lambda - p_4 = 0$

определяем по формулам (15):

$$A_1 = \begin{pmatrix} 2.2 & 1 & 0.5 & 2 \\ 1 & 1.3 & 2 & 1 \\ 0.5 & 2 & 0.5 & 1.6 \\ 2 & 1 & 1.6 & 2 \end{pmatrix}, \quad p_1 = 2.2 + 1.3 + 0.5 + 2 = 6,$$

$$B_1 = \begin{pmatrix} 2.2 & 1 & 0.5 & 2 \\ 1 & 1.3 & 2 & 1 \\ 0.5 & 2 & 0.5 & 1.6 \\ 2 & 1 & 1.6 & 2 \end{pmatrix} - \begin{pmatrix} 6 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix} = \begin{pmatrix} -3.8 & 1 & 0.5 & 2 \\ 1 & -4.7 & 2 & 1 \\ 0.5 & 2 & -5.5 & 1.6 \\ 2 & 1 & 1.6 & -4 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} 2.2 & 1 & 0.5 & 2 \\ 1 & 1.3 & 2 & 1 \\ 0.5 & 2 & 0.5 & 1.6 \\ 2 & 1 & 1.6 & 2 \end{pmatrix} \begin{pmatrix} -3.8 & 1 & 0.5 & 2 \\ 1 & -4.7 & 2 & 1 \\ 0.5 & 2 & -5.5 & 1.6 \\ 2 & 1 & 1.6 & -4 \end{pmatrix} =$$

$$= \begin{pmatrix} -3.11 & 0.5 & 3.55 & -1.8 \\ 0.5 & -0.11 & -6.3 & 2.5 \\ 3.55 & -6.3 & 4.06 & -2.6 \\ -1.8 & 2.5 & -2.6 & -0.44 \end{pmatrix}$$

Таблица 3

$A_i$				$B_i$			
2.2	1	0.5	2	-3.8	1	0.5	2
1	1.3	2	1	1	-4.7	2	1
0.5	2	0.5	1.6	0.5	2	-5.5	1.6
2	1	1.6	2	2	1	1.6	-4

 $p_1 = 6$ 

-3.11				-3.31	0.5	3.55	-1.8
	-0.11			0.5	-0.31	-6.3	2.5
		4.06		3.55	-6.3	3.86	-2.6
			-0.44	-1.8	2.5	-2.6	-0.64

 $p_2 = 0.2$ 

-8.607				4.128	2.64	-1.76	-4.04
	-10.003			2.64	2.732	0.48	-4.39
		-13.055		-1.76	0.48	-0.32	1.776
			-8.54	-4.04	-4.39	1.776	6.195

 $p_3 = -12.735$ 

2.7616			
	2.7616		
		2.7616	
			2.7616

 $p_4 = 2.7616$ 

Получили уравнение

$$\lambda^4 - 6\lambda^3 - 0.2\lambda^2 + 12.735\lambda - 2.7616 = 0,$$

корнями которого являются:

$$\lambda_1 = 5.652, \lambda_2 = -1.545, \lambda_3 = -1.420, \lambda_4 = 0.2226.$$

Вычислим векторы  $\bar{X}_i$ , следуя (17):

$$\lambda_1 = 5.652.$$

Начальный вектор:

$$\bar{u}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \bar{u}_1 = \begin{pmatrix} \lambda \cdot u_{01} + b_{11} \\ \lambda \cdot u_{02} + b_{12} \\ \lambda \cdot u_{03} + b_{13} \\ \lambda \cdot u_{04} + b_{14} \end{pmatrix} = \begin{pmatrix} 5.652 \cdot 1 - 3.8 \\ 5.652 \cdot 0 + 1 \\ 5.652 \cdot 0 + 0.5 \\ 5.652 \cdot 0 + 2 \end{pmatrix} = \begin{pmatrix} 1.852 \\ 1 \\ 0.5 \\ 2 \end{pmatrix},$$

$$\bar{u}_2 = \begin{pmatrix} \lambda \cdot u_{11} + b_{21} \\ \lambda \cdot u_{12} + b_{22} \\ \lambda \cdot u_{13} + b_{23} \\ \lambda \cdot u_{14} + b_{24} \end{pmatrix} = \begin{pmatrix} 5.652 \cdot 1.852 - 3.31 \\ 5.652 \cdot 1 + 0.5 \\ 5.652 \cdot 0.5 + 3.55 \\ 5.652 \cdot 2 - 1.8 \end{pmatrix} = \begin{pmatrix} 7.1575 \\ 6.152 \\ 6.376 \\ 9.504 \end{pmatrix},$$

$$\bar{u}_3 = \begin{pmatrix} \lambda \cdot u_{21} + b_{31} \\ \lambda \cdot u_{22} + b_{32} \\ \lambda \cdot u_{23} + b_{33} \\ \lambda \cdot u_{24} + b_{34} \end{pmatrix} = \begin{pmatrix} 5.652 \cdot 7.1575 + 4.128 \\ 5.652 \cdot 6.152 + 2.64 \\ 5.652 \cdot 6.376 - 1.76 \\ 5.652 \cdot 9.504 - 4.04 \end{pmatrix} = \begin{pmatrix} 44.5822 \\ 37.4111 \\ 34.2772 \\ 49.6766 \end{pmatrix}.$$

Аналогично вычисляются вектора при  $\lambda_2 = -1.545$ ,  $\lambda_3 = -1.420$ ,  $\lambda_4 = 0.2226$ .

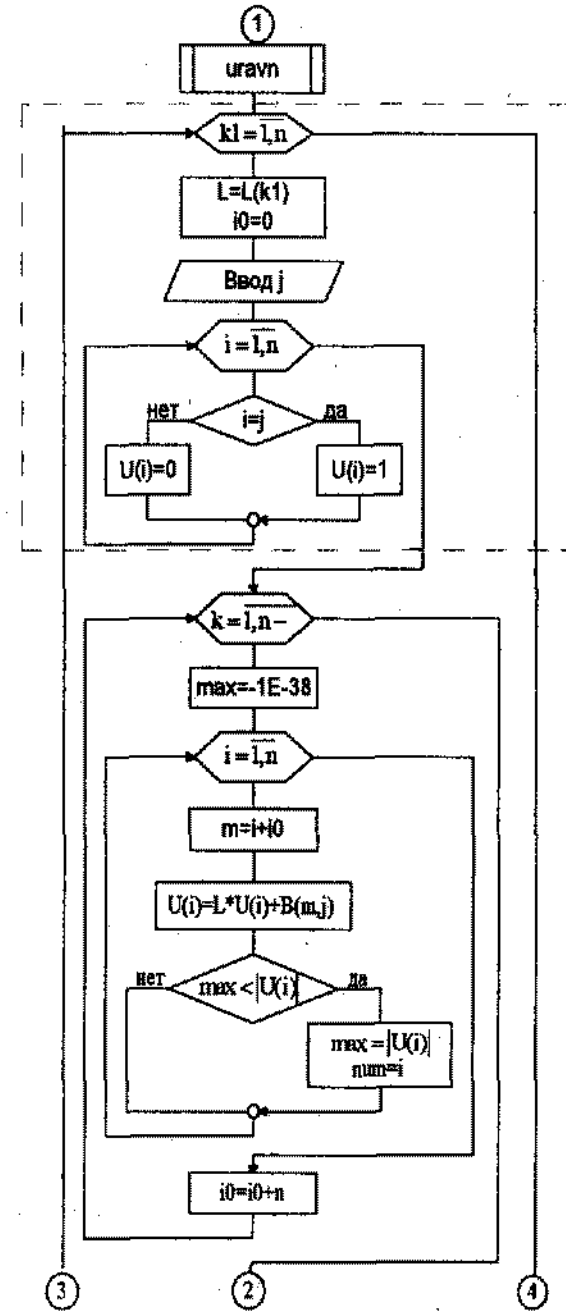
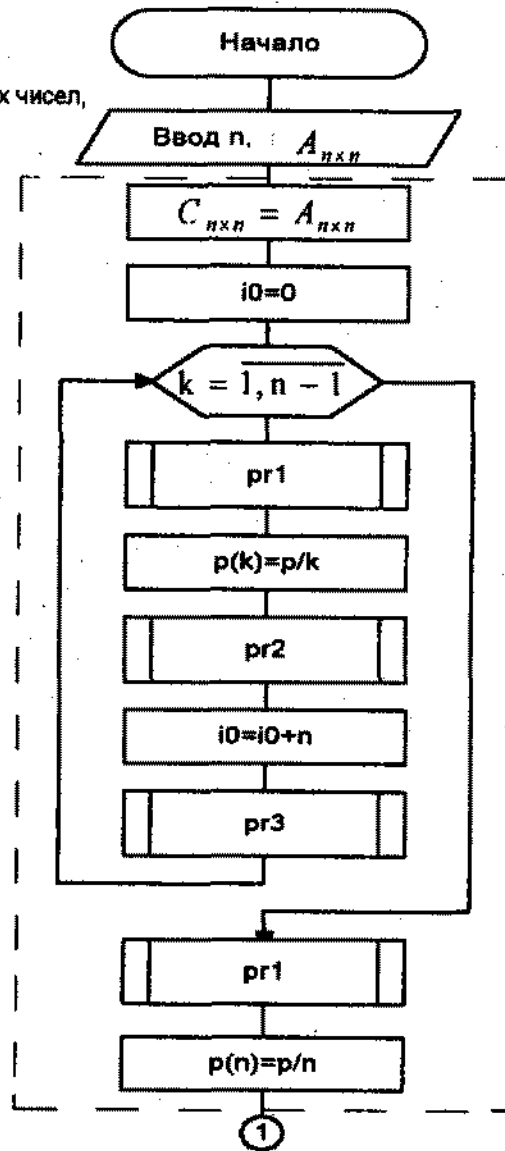
Результаты в таблице 4.

Таблица 4

$\lambda_i$	$\bar{u}_{0i}$	$\bar{u}_{1i}$	$\bar{u}_{2i}$	$\bar{u}_{3i} = \bar{X}_i$	$\bar{X}_i$
	1	1.852	7.1575	44.5822	0.897
	0	1	6.152	37.4111	0.753
	0	0.5	6.3760	34.2772	0.690
5.652	0	2	9.504	49.6766	1
	1	-2.255	-6.7940	-6.3687	1
	0	1	2.045	5.7995	-0.911
	0	0.5	4.3225	4.9183	-0.772
-1.545	0	2	1.290	-2.0470	0.321
	0	1	-0.92	3.9484	0.293
	1	-6.12	8.3804	-9.1682	-0.681
	0	2	-9.14	13.4588	1
-1.420	0	1	1.08	-5.9236	-0.440
	0	1	0.7226	2.8009	-0.740
	1	-4.4775	-1.3067	2.4411	-0.645
	0	2	-5.8548	-0.8233	0.218
0.2226	0	1	2.7226	-3.7839	1

Структурная схема алгоритма метода Ливерье-Фаддеева

Входные параметры:  
 $A()$  – исходная матрица,  
 $n$  – порядок матрицы,  
 $E$  – погрешность.  
 Выходные параметры:  
 $L()$  – матрица собственных чисел,  
 $X()$  – матрица собственных векторов.



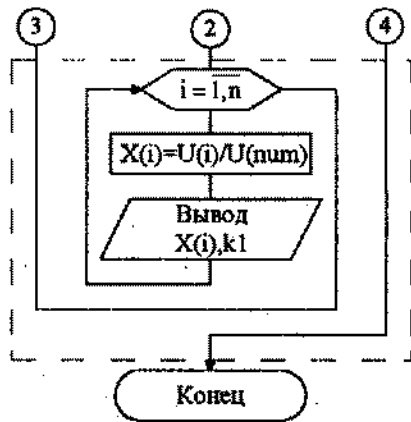


Рис. 11.3

**Процедура формирования матрицы A**

Выходные параметры:  
A ( ) – формируемая матрица.

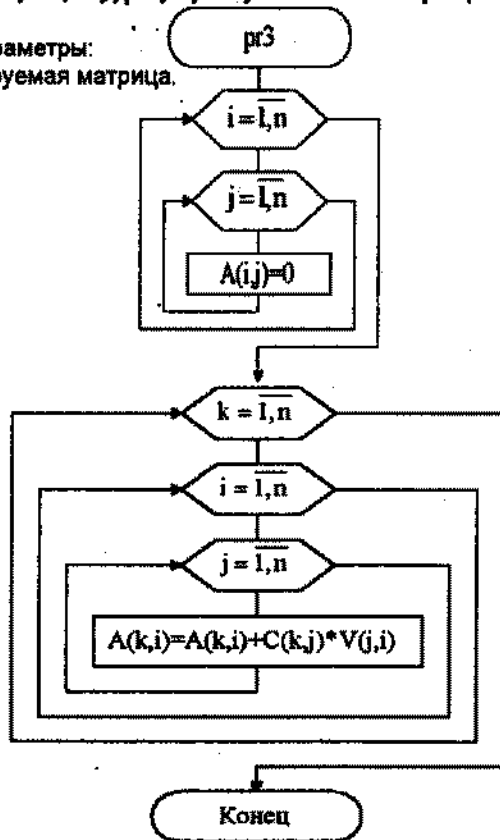


Рис. 11.4

**Процедура формирования «следа» матрицы A**

Выходной параметр:  
p – «след» матрицы A.

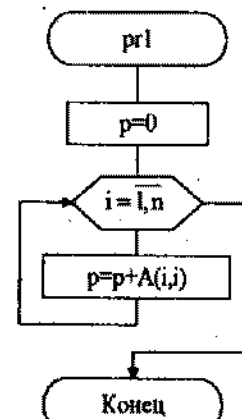


Рис. 11.5

**Процедура формирования матрицы V**

Выходной параметр:  
V ( ) – матрица  
собственных векторов.

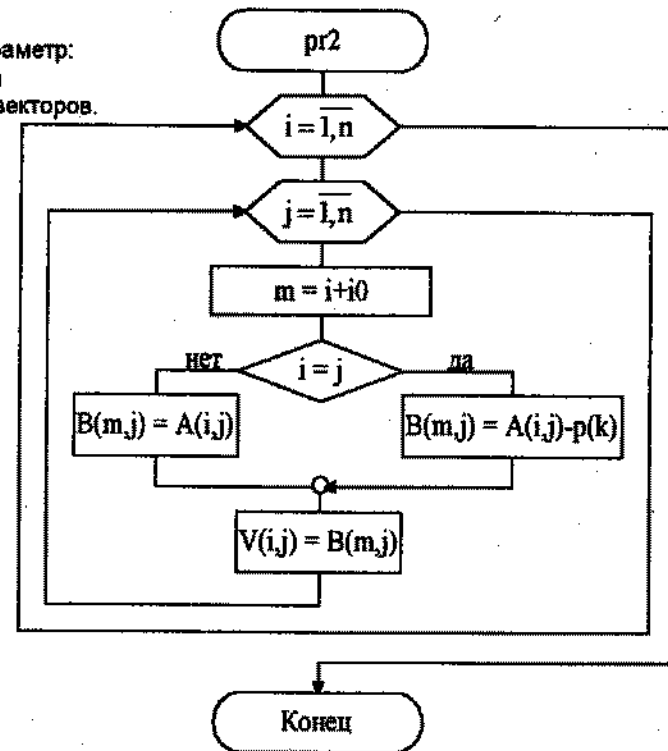


Рис. 11.6

**Процедура нахождения корня характеристического уравнения**

Выходной параметр:  
L ( ) - массив корней.

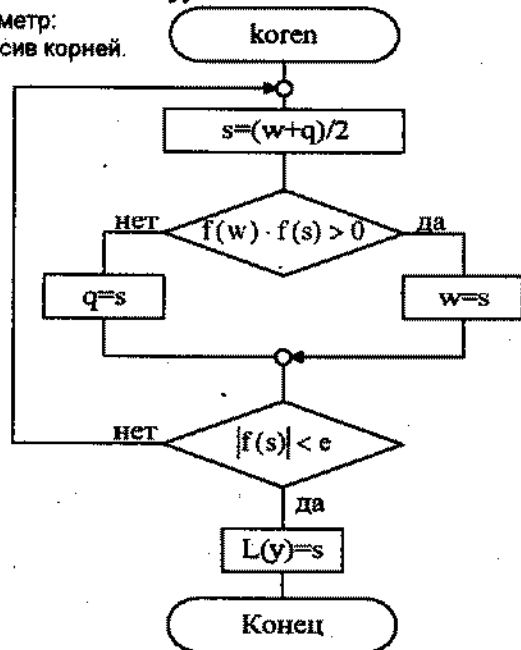


Рис. 11.7

**Процедура определения границ действительного корня**

Выходные параметры:  
w, q - границы  
действительного корня.

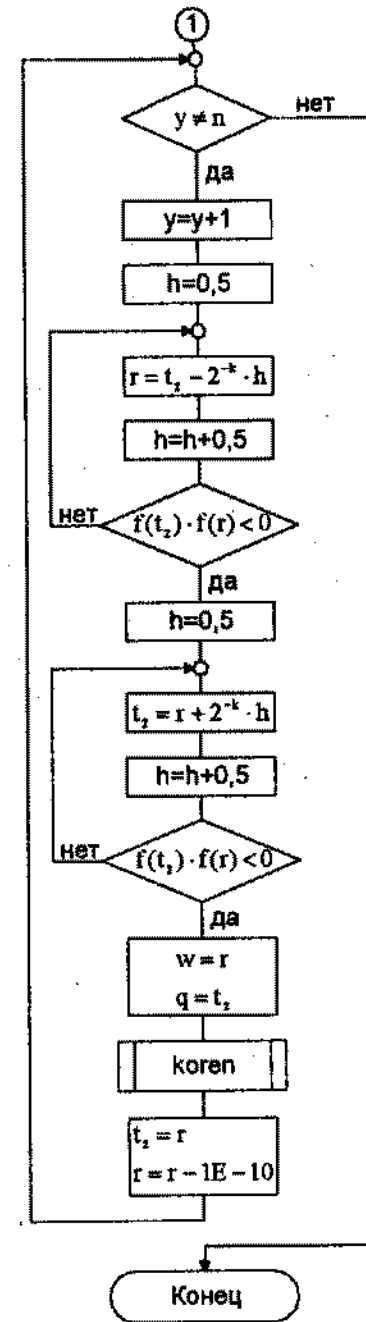
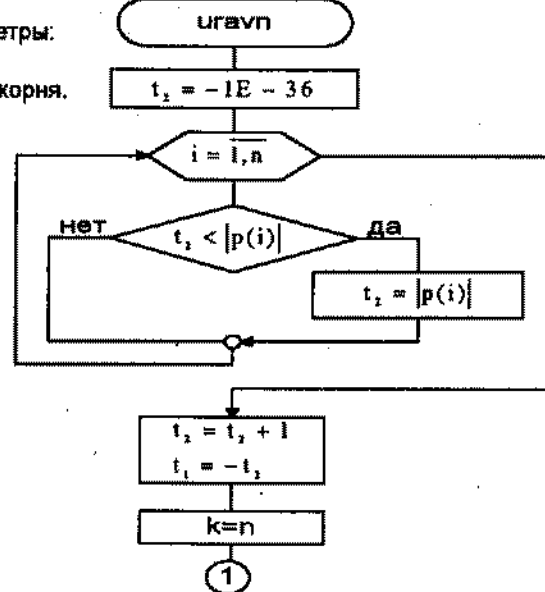


Рис. 11.8



## 11.1. Метод Данилевского

Для вычисления коэффициентов  $p_1, p_2, \dots, p_n$  характеристического уравнения матрицы

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

рассматривается переход от матрицы  $A$  к матрице

$$P = \begin{pmatrix} p_1 & p_2 & \dots & p_{n-1} & p_n \\ 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

Этот переход осуществляется с помощью  $(n-1)$  преобразования подобия, последовательно преобразующих строки матрицы  $A$ , начиная с последней, в соответствующие строки матрицы  $P$ .

В предположении, что  $a_{n,n-1} \neq 0$ , проводим преобразования:

$$1) \quad m_{n-1,i} = -\frac{a_{ni}}{a_{n,n-1}} \quad \text{при } i \neq n-1, \quad (18)$$

$$m_{n-1,n-1} = \frac{1}{a_{n,n-1}} \quad \text{при } i = n-1.$$

$$2) \quad b_j = a_j + a_{i,n-1} m_{n-1,j} \quad \text{при } 1 \leq i \leq n; \quad j \neq n-1, \quad (19)$$

$$b_{i,n-1} = a_{i,n-1} \cdot m_{n-1,n-1} \quad \text{при } 1 \leq i \leq n.$$

$$3) \quad c_j = b_j \quad \text{при } 1 \leq i \leq n-2, \quad (20)$$

$$c_{n-1,j} = \sum_{k=1}^n a_{nk} b_{kj} \quad \text{при } 1 \leq j \leq n.$$

П.п. 1-3 повторяются с матрицей  $C$ , взяв за основную ее  $(n-2)$  строку и так далее.

Если на каком-то этапе окажется  $a_{k,k-1} = 0$ , то возможны два случая:

- Пусть какой-то элемент матрицы  $D$ , стоящий левее нулевого элемента  $d_{k,k-1} = 0$  отличен от нуля, то есть  $d_{k,l} \neq 0$ , где  $l < k-1$ . Тогда этот элемент выдвигаем на место  $d_{k,k-1}$ , то есть переставляем  $(k-1)$ -й и  $l$ -й столбцы, одновременно переставляя ее  $(k-1)$ -ю и  $l$ -ю строки.

К новой матрице применяем метод Данилевского.

- Пусть  $d_{k,l} = 0$  ( $l=1,2,\dots,k-1$ ), тогда матрица  $D$  имеет вид

$$D = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,k-1} & a_{1,k} & \dots & a_{1,n-1} & \dots & a_{1,n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ c_{k-1,1} & c_{k-1,2} & \dots & c_{k-1,k-1} & c_{k-1,k} & \dots & c_{k-1,n-1} & \dots & c_{k-1,n} \\ 0 & 0 & \dots & 0 & c_{kk} & \dots & c_{k,n-1} & \dots & c_{kn} \\ 0 & 0 & \dots & 0 & 1 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 & \dots & 0 \end{pmatrix} = \begin{pmatrix} D_1 & L \\ 0 & D_2 \end{pmatrix} \quad (21)$$

$$\det(D - \lambda E) = \det(D_1 - \lambda E) \cdot \det(D_2 - \lambda E).$$

Остается применить метод Данилевского к матрице  $D_1$ .

Структурная схема алгоритма метода Данилевского на рисунках (9) – (12).

### Пример 4.

Найти коэффициенты характеристического уравнения матрицы

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 2 & 3 \\ 3 & 2 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

Решение.

Матрица  $A$  – невырожденная. Преобразуем, начиная с последней строки, матрицу  $A$ . В соответствии с приведенным алгоритмом необходимы три преобразования ( $n=4$ ). Так как  $a_{43} = 2 \neq 0$ , по формулам (18) – (20) получим матрицу

$$C_1 = \begin{pmatrix} -5 & -2.5 & 1.5 & 2.5 \\ -2 & -2 & 1 & 2 \\ -24 & -15 & 11 & 19 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Последовательно преобразуя строки, получим

$$C_2 = \begin{pmatrix} -0.999999 & 0.166667 & -0.333334 & -0.666667 \\ 5.999994 & 5 & 34 & 24 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$C_3 = \begin{pmatrix} 4 & 40 & 56 & 20 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Коэффициенты характеристического уравнения матрицы  $A$  расположены в первой строке матрицы  $C_3$ .

$$p_1 = 4, \quad p_2 = 40, \quad p_3 = 56, \quad p_4 = 20.$$

Согласно (6), характеристическое уравнение имеет вид:

$$\lambda^4 - 4\lambda^3 - 40\lambda^2 - 56\lambda - 20 = 0.$$

### Пример 5.

Вычислить коэффициенты характеристического уравнения матрицы

$$A = \begin{pmatrix} 1 & 1 & 4 & 3 \\ 1 & 1 & 2 & 1 \\ 0 & 0 & 4 & 3 \\ 0 & 0 & 2 & 1 \end{pmatrix}$$

*Решение.*

Преобразуем  $n-1=3$  строку исходной матрицы по формулам (18) – (20).

$$C = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Элемент  $a_{32} = 0$ . В соответствии с алгоритмом получаем матрицу (21).

$$\begin{pmatrix} 2 & 0 & 2 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} D_1 & L \\ 0 & D_2 \end{pmatrix}$$

Так как  $\det D_1 = 0$ , коэффициенты характеристического многочлена матрицы равны:

$$\text{для первого уравнения } p_1 = 2, \quad p_2 = 0,$$

$$\text{для второго уравнения } p_1 = 5, \quad p_2 = 2.$$

## 11.2. Метод Якоби

Метод предназначен для приведения матрицы к диагональному виду. Это бесконечный процесс, поэтому практически метод Якоби рассматривается как итерационная процедура, преобразующая исходную матрицу к матрице, достаточно близкой к диагональной.

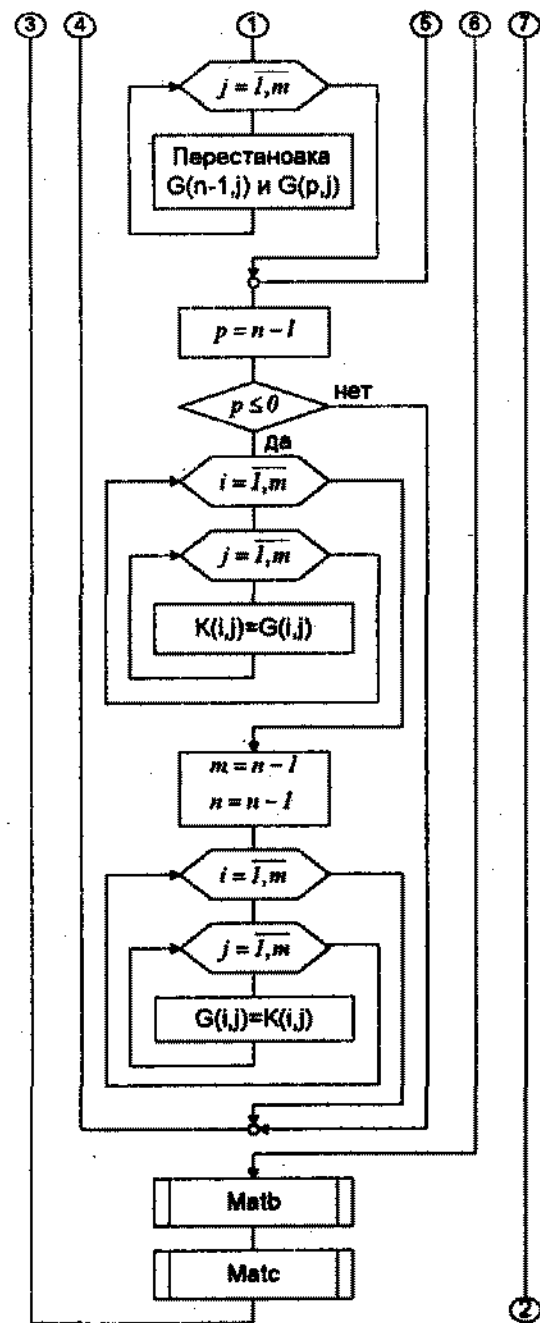
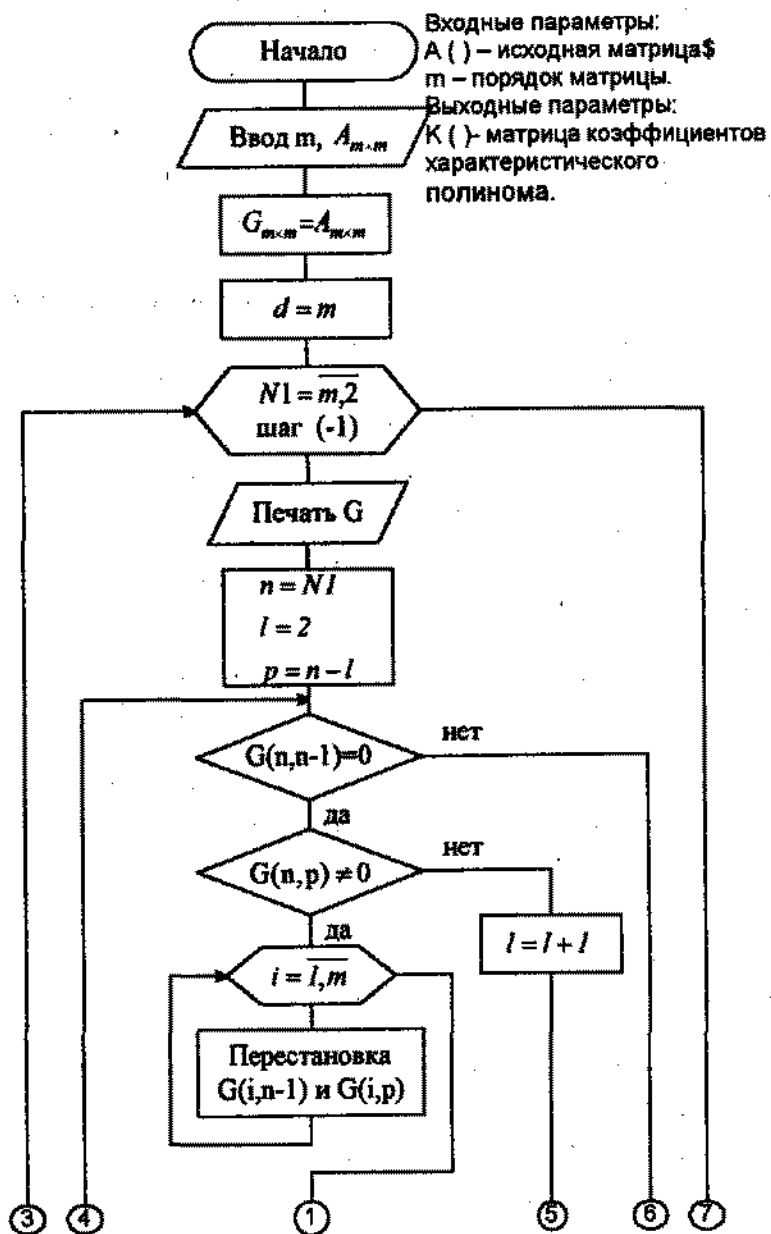
Схема вычислений такова: матрицу  $A_1 = P_1 \cdot A \cdot P_1^T$  получают с помощью ортогональной матрицы  $P_1$  так, чтобы получить нулевой внедиагональный элемент, затем получают аналогично  $A_2 = P_2 \cdot A_1 \cdot P_2^T$  и так далее. На каждом шаге наибольший по модулю внедиагональный элемент обращается в нуль.

Процедура продолжается до тех пор, пока наибольший по модулю внедиагональный элемент не окажется настолько малым по сравнению с диагональным элементом, что его можно принять равным нулю.

Если  $a_{kl}$  - наибольший по модулю элемент матрицы  $A_{m-1}$ , то матрица  $P_m$  задается так:

$$P_{kk} = P_{ll} = \cos \theta, \quad P_{ij} = 1, \quad i \neq k, l, \quad P_{kl} = -P_{lk} = \sin \theta, \\ \text{остальные } P_{ij} = 0. \quad (22)$$

### Структурная схема алгоритма метода Данилевского



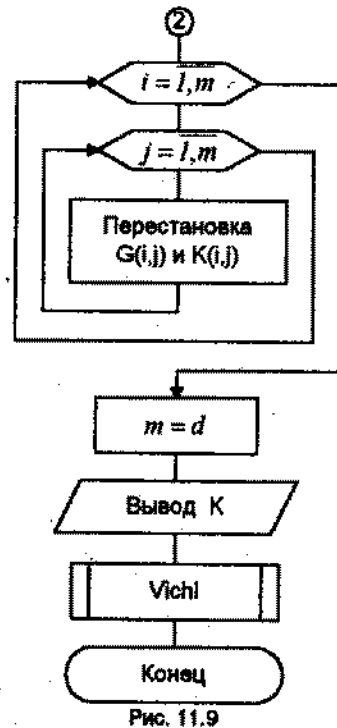


Рис. 11.9

**Процедура формирования промежуточной матрицы B**

Выходные параметры:  
B( ) – матрица промежуточных вычислений.

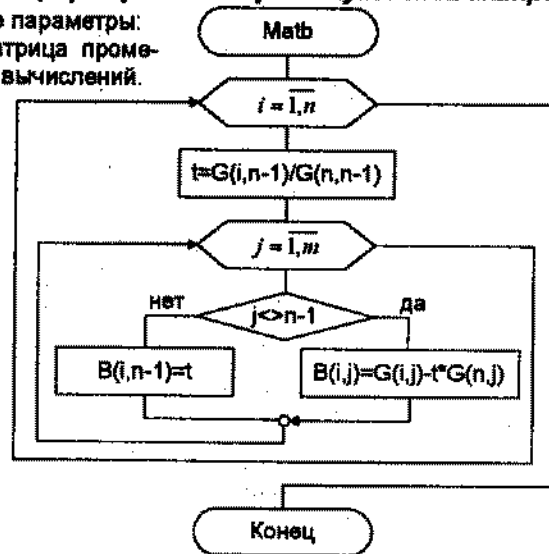


Рис. 11.10

**Процедура вывода коэффициентов характеристического многочлена**

Выходные параметры:  
B( ) – матрица промежуточных вычислений.

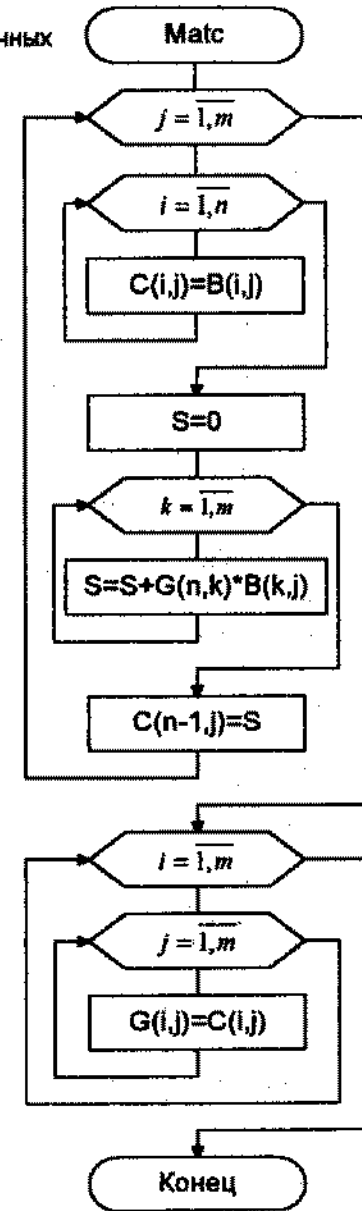
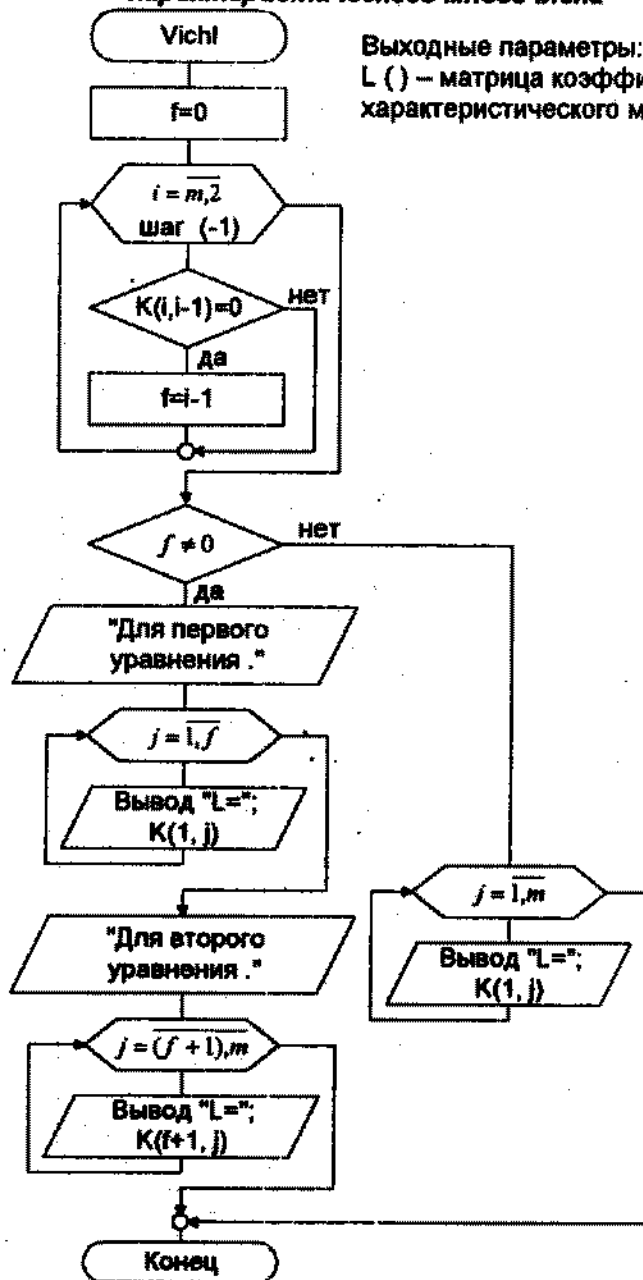


Рис. 11.11

**Процедура вывода коэффициентов  
характеристического многочлена**



Выходные параметры:  
L ( ) – матрица коэффициентов  
характеристического многочлена.

Рис. 11.12

Чтобы элемент  $a_{kl}^{(m)}$  оказался равным нулю, величину  $\theta$  выбирают из условий

$$\operatorname{tg} 2\theta = 2a_{kl}^{(m-1)} / (a_{kk}^{(m-1)} - a_{ll}^{(m-1)}),$$

$$-\frac{\pi}{4} \leq \theta \leq \frac{\pi}{4}. \quad (23)$$

Матрица  $A_m$  отличается от  $A_{m-1}$  только  $k$  строкой и  $l$  столбцом. Новые элементы этих строк и столбцов можно найти из соотношений:

$$\begin{aligned} a_{kk}^{(m)} &= a_{kk}^{(m-1)} \cdot \cos \theta + a_{ll}^{(m-1)} \cdot \sin \theta \quad \text{при } i \neq k, l, \\ a_{ll}^{(m)} &= a_{ll}^{(m-1)} \cdot \cos \theta - a_{kk}^{(m-1)} \cdot \sin \theta \quad \text{при } i \neq k, l, \\ a_{kk}^{(m)} &= a_{kk}^{(m-1)} \cdot \cos^2 \theta + 2a_{kl}^{(m-1)} \cdot \cos \theta \cdot \sin \theta + a_{ll}^{(m-1)} \cdot \sin^2 \theta, \\ a_{ll}^{(m)} &= a_{ll}^{(m-1)} \cdot \sin^2 \theta - 2a_{kl}^{(m-1)} \cdot \cos \theta \cdot \sin \theta + a_{kk}^{(m-1)} \cdot \cos^2 \theta, \\ a_{kl}^{(m)} &= (a_{kl}^{(m-1)} - a_{kk}^{(m-1)}) \cdot \cos \theta \cdot \sin \theta + a_{ll}^{(m-1)} \cdot (\cos^2 \theta - \sin^2 \theta). \end{aligned} \quad (24)$$

$$\cos \theta = \sqrt{\frac{z+y}{2z}}, \quad \sin \theta = \operatorname{sign} \sqrt{\frac{z-y}{2z}}, \quad (25)$$

$$x = 2a_{kl}, \quad y = a_{kk} - a_{ll}, \quad z = \sqrt{x^2 + y^2}, \quad (26)$$

$$\operatorname{sign} = \frac{x/y}{|x/y|}.$$

В результате получаем диагональную матрицу  $(P_f \dots P_2 \cdot P_1) A_0 (P_1^T \cdot P_2^T \dots P_f^T)$ .

Собственные числа — диагональные элементы. Собственные векторы матрицы  $A_0 = A$  есть столбцы матрицы  $R_f = P_1^T \cdot P_2^T \dots P_f^T$ . Исходная матрица  $R_1$  - единичная.

У матрицы  $R$  имеется только  $k$ -я строка и  $l$ -ый столбец:

$$r_{kk}^{(m)} = r_{kk}^{(m-1)} \cdot \cos \theta + r_{ll}^{(m-1)} \cdot \sin \theta, \quad (27)$$

$$r_{ll}^{(m)} = -r_{kk}^{(m-1)} \cdot \sin \theta + r_{ll}^{(m-1)} \cdot \cos \theta, \quad i = \overline{1, n}.$$

### 11.3. QR-метод

QR-метод - метод преобразования матрицы общего вида  $A$  к блочно-диагональной матрице. Метод позволяет найти все собственные значения, включая комплексные.

Строим последовательность матриц  $A_m$  следующим образом.

Матрицу  $A_m$  представляем в виде:

$$A_m = Q_{m+1} \cdot R_{m+1}, \quad (28)$$

где

$Q_{m+1}$  - ортогональная матрица,

$R_{m+1}$  - верхне-гессенбергова матрица.

Полагаем

$$A_{m+1} = R_{m+1} \cdot Q_{m+1}. \quad (29)$$

Так как  $A_{m+1} = Q_{m+1}^{-1} \cdot A_m \cdot Q_{m+1}$ , то все матрицы  $A_m$  подобны между собой и подобны матрице  $A$ .

Рекуррентная формула:

$$A_{m+1} = R_m \cdot Q_m. \quad (30)$$

Матрица  $A$  преобразуется к блочно-диагональной матрице:

$$\begin{pmatrix} x_1 & * & * & \dots & * \\ & x_2 & * & \dots & * \\ & & x_3 & \dots & * \\ & & & \dots & * \\ & & & & \dots \\ & & & & & \dots \end{pmatrix}. \quad (31)$$

Блоки  $x_m$  — матрицы  $2 \times 2$ .

Собственные числа блоков  $x_m$  совпадают с собственными числами матрицы  $A$ .

Определители второго порядка блоков  $x_m$  позволяют найти комплексные собственные числа без использования комплексных элементов в преобразованной матрице.

Если же собственные числа исходной матрицы вещественны, то преобразованная матрица имеет чисто диагональную форму с искомыми собственными числами на диагонали.

Ускорение сходимости QR-алгоритма [79] может быть достигнуто сдвигом

$$A_m - k_m E = Q_{m+1} R_{m+1},$$

$k_m$  — сдвиг,  $E$  - единичная матрица.

$$A_m = R_m Q_m + k_m E.$$

Использование двухступенчатой процедуры сдвига на  $k_m E$  и  $k_{m+1} E$ , где  $k_m$  и  $k_{m+1}$  — либо два вещественных числа, либо пара комплексно-сопряженных чисел, позволяет избежать комплексной арифметики.

В методе используется матрица в форме Гессенберга. Если матрица имеет вид:

$$\begin{pmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ 0 & * & * & * & * & * \\ 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & * & * \end{pmatrix}$$

е называют ниже-гессенберговой, матрицу

$$\begin{pmatrix} * & * & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 \\ * & * & * & * & 0 & 0 \\ * & * & * & * & * & 0 \\ * & * & * & * & * & * \\ * & * & * & * & * & * \end{pmatrix}$$

верхне-гессенберговой.

Рассмотрим способ приведения матрицы  $A$  к форме Гессенберга. Ищем матрицу преобразования  $N$  такую, что

$$N^{-1} A N = H,$$

где  $H$  — верхне-гессенбергова матрица.

Матричное уравнение для нахождения матрицы  $N$ :

$$A N = N H.$$

**Алгоритм приведения матрицы  $A$  к форме Гессенберга**

1. Вычислить при  $i = \overline{1, r}$  величины

$$h_{ir} = a_{ir} + \sum_{k=r+1}^n a_{ik} n_{kr} - \sum_{k=2}^{r-1} n_{ik} h_{kr}.$$

Все величины  $h_{ir}$  можно записывать на место  $a_{ir}$ .

2. Вычислить при  $i = \overline{r+1, n}$  величины

$$l_i = a_{ir} + \sum_{k=r+1}^n a_{ik} n_{kr} - \sum_{k=2}^r n_{ik} h_{kr}.$$

Найти из  $l_i$  максимальную по модулю, зафиксировать ее номер

$$i = i_{\max}$$

3. Поменять местами строки с номерами  $r+1$  и  $i_{\max}$  (включая  $n_j$  и  $l_i$ ) и затем поменять местами столбцы с номерами  $r+1$  и  $i_{\max}$ . Цель перестановок - перемещение наибольшего осевого элемента на место  $h_{r+1,r}$  для обеспечения максимальной точности и исключения нулевых элементов.

4. Записать величины  $l_{r+1}$  на место  $h_{r+1,r}$ .

5. Получить при всех  $i = \overline{r+2, n}$  величину  $H_{i,r+1} = \frac{l_i}{h_{r+1,r}}$  и за-

писать ее на место  $a_{ir}$ .

#### Замечание.

- При перестановке двух строк собственные значения матрицы не изменяются лишь в том случае, когда соответствующие столбцы тоже переставляются.
- Две матрицы  $N$  и  $H$  можно хранить в памяти компьютера на месте исходной матрицы  $A$ , если сдвинуть матрицу  $N$  на один столбец влево и помнить, что первый столбец матрицы  $N$  почти весь состоит из нулей, а все диагональные элементы  $N$  равны единице.

## 12. ИНТЕГРАЛЬНЫЕ УРАВНЕНИЯ

Область применения интегральных уравнений в задачах теории связи весьма обширна. Традиционной является область применения в задачах статистической динамики: определение корреляционной функции стационарного случайного процесса, определение динамических характеристик системы. В таких разделах, как статистическая радиотехника, проблемы передачи информации, синтез антенн и других также широко применяются методы составления и решения интегральных уравнений.

В общем случае интегральное уравнение имеет вид

$$h(x)\varphi(x) - \lambda \cdot \int_V K(x,s)\varphi(s)ds = f(x), \quad (1)$$

$K(x,s)$  - ядро интегрального уравнения,

$\varphi(x)$  - искомая функция,

$V$  - область интегрирования,

$\lambda$  - числовой множитель, параметр,

$f(x), h(x)$  - заданные функции.

Интегральное уравнение называется линейным, если неизвестная функция входит в уравнение линейно. Ограничимся рассмотрением лишь линейных уравнений. Один из наиболее важных приложений классов линейных интегральных уравнений — уравнения Фредгольма.

Общий вид уравнения Фредгольма 2-го рода

$$\varphi(x) = \lambda \cdot \int_a^b K(x,s)\varphi(s)ds + f(x), \quad (2)$$

$\varphi(x)$  - неизвестная функция, пределы интегрирования могут быть как конечными, так и бесконечными. Полагаем, что переменная  $x$  меняется в том же промежутке, что и переменная интегрирования.

Предполагаем также, что ядро  $K(x,s)$  и свободный член  $f(x)$  либо непрерывны, либо удовлетворяют условиям:

$$\bullet \int_a^b \int_a^b |K(x,s)|^2 dx ds < +\infty, \quad (3)$$

$$\bullet \int_a^b |f(x)|^2 dx < +\infty. \quad (4)$$

Уравнение Фредгольма 1-го рода

$$\int_a^b K(x,s)\varphi(s)ds = f(x), \quad (5)$$

где ядро  $K(x,s)$  и функция  $f(x)$  удовлетворяют сформулированным выше условиям.

Линейное интегральное уравнение Вальтерры 2-го рода имеет вид

$$\varphi(x) = \int_a^x K(x,s)\varphi(s)ds + f(x), \quad a \leq x < b, \quad (6)$$

$\varphi(x)$  - неизвестная функция, ядро  $K(x,s)$  и свободный член  $f(x)$  - известные функции.

Уравнение Вальтерры 1-го рода

$$\int_a^x K(x,s)\varphi(s)ds = f(x) \quad (7)$$

Если ввести ядро

$$\Phi(x,s) = \begin{cases} K(x,s), & s \leq x, \\ 0, & s > x, \end{cases} \quad (8)$$

то уравнение Вальтерры можно рассматривать как частный случай уравнения Фредгольма.

Если  $f(x) \equiv 0$ , то уравнение называется однородным, в противном случае неоднородным. При решении интегральных уравнений Фредгольма приходится встречаться с решением двух задач:

1) отыскание решения неоднородного интегрального уравнения при заданном значении параметра и заданном свободном члене  $f(x)$ ;

2) отыскание собственных значений и собственных функций ядра  $K(x,s)$ , то есть отыскание такого параметра  $\lambda$ , при котором

уравнение  $\varphi(x) - \lambda \int_a^b K(x,s)\varphi(s)ds = 0$  имеет нетривиальное решение  $\varphi(x)$ . Эти значения  $\lambda$  и нетривиальные решения  $\varphi(x)$  называются, соответственно, собственными значениями и собственными функциями ядра.

## 12.1. Линейные интегральные уравнения Фредгольма 2-го рода

Численное решение уравнения Фредгольма осложняется тем, что это уравнение не всегда разрешимо. В случае, если решение возможно, интегральное уравнение приближенно представляется с помощью конечно-разностных выражений.

### 12.1.1. Метод квадратур для уравнений Фредгольма 2-го рода

В уравнении (2) интеграл заменяется той или иной квадратурой

$$\int_a^b F(s) ds = \sum_{j=1}^n A_j F(s_j),$$

$s_1, s_2, \dots, s_n$  принадлежат отрезку  $[a,b]$  изменения переменной  $s$ , коэффициенты  $A_j$  не зависят от выбора функции  $F(x)$ . Полагая в (2)  $x = s_1, s_2, \dots, s_n$  получим систему  $n$  линейных алгебраических уравнений относительно  $n$  неизвестных  $\varphi(s_j)$ .

$$\varphi(s_i) - \lambda \cdot \sum_{j=1}^n K(s_i, s_j) \varphi(s_j) \delta = f(s_i) \quad (9)$$

или

$$\varphi_i - \lambda \cdot \sum_{j=1}^n K_{ij} \varphi_j \delta = f_i$$

$$\varphi_i = \varphi(s_i), \quad K_{ij} = K(s_i, s_j), \quad f_i = f(s_i), \quad \delta = \frac{b-a}{n}$$

Предполагается, что  $\varphi(s)$  и  $f(s)$  сохраняют в  $i$ -м интервале постоянные значения  $\varphi(s_i)$  и  $f(s_i)$ , а ядро  $K(x,s) = K(s_i, s_j)$  в каждом частичном квадрате с индексами  $i$  и  $j$ . Определитель системы

$$D_n(\lambda) = \begin{vmatrix} 1 - \lambda K(s_1, s_1) \delta & -\lambda K(s_1, s_2) \delta & \dots & -\lambda K(s_1, s_n) \delta \\ -\lambda K(s_2, s_1) \delta & 1 - \lambda K(s_2, s_2) \delta & \dots & -\lambda K(s_2, s_n) \delta \\ \dots & \dots & \dots & \dots \\ -\lambda K(s_n, s_1) \delta & -\lambda K(s_n, s_2) \delta & \dots & 1 - \lambda K(s_n, s_n) \delta \end{vmatrix} \quad (10)$$

есть многочлен относительно  $\lambda$ .

Если  $D(\lambda) \neq 0$ , то система имеет единственное решение, которое позволяет получить для любого  $x \in [a,b]$  решение интегрального уравнения (2) через известные (найденные) значения  $\varphi_1, \varphi_2, \dots, \varphi_n$ , к примеру интерполированием.

$$\varphi(x) = f(x) + \sum_{j=1}^n A_j K(x, s_j) \varphi(s_j). \quad (11)$$

Оценка погрешности приведена в [8].



Для уравнения Фредгольма 1-го рода система алгебраических уравнений такова:

$$\lambda \cdot \sum_{j=1}^n K_{ij} \varphi_j \delta = f_i, \quad i = \overline{1, n}. \quad (12)$$

Приведем значения коэффициентов  $A_j$  (11) в зависимости от типа квадратурной формулы.

1. Метод прямоугольников:

$$s_1 = a, \quad s_2 = a+h, \quad \dots, \quad s_n = a+(n-1)h = b, \quad h = \frac{b-a}{n}. \quad (13)$$

2. Метод трапеций:

$$s_1 = a, \quad s_2 = a+h, \quad \dots, \quad s_n = a+(n-1)h = b, \\ A_1 = A_n = \frac{h}{2}, \quad A_2 = A_3 = \dots = A_{n-1} = h. \quad (14)$$

3. Метод Симпсона:

$$n = 2m+1, \quad h = \frac{b-a}{2m}, \quad (15)$$

$$s_1 = a, \quad s_2 = a+h, \quad \dots, \quad s_{2m+1} = a+2m \cdot h = b; \\ A_1 = A_{2m+1} = \frac{h}{3}, \quad A_2 = A_4 = \dots = A_{2m} = \frac{4h}{3}.$$

Выбор квадратурной формулы зависит от свойств функций  $K(x, s)$ ,  $f(x)$ . При достаточной гладкости этих функций можно брать и более точные квадратурные формулы.

**Пример 1** [8].

Найти приближенное решение уравнения

$$\varphi(x) + \int_0^1 x(e^{xs} - 1)\varphi(s) ds = e^x - x.$$

**Решение.**

Воспользуемся квадратурной формулой Симпсона при  $m=1$ , то есть  $h=0.5$ ,  $x=0; 0.5; 1$ .

$$\int_0^1 F(x) dx = \frac{h}{3}(F(0) + 4F(0.5) + F(1)).$$

Получаем систему

$$\begin{cases} \varphi_1 = 1, \\ \frac{1}{3}(e^{0.25} + 2)\varphi_2 + \frac{1}{12}(e^{0.5} - 1)\varphi_3 = e^{0.5} - 0.5, \\ \frac{2}{3}(e^{0.5} - 1)\varphi_2 + \frac{1}{6}(e + 5)\varphi_3 = e - 1, \end{cases}$$

$$\begin{cases} \varphi_1 = 1, \\ 1.0947\varphi_2 + 0.0541\varphi_3 = 1.1487, \\ 0.4325\varphi_2 + 1.2864\varphi_3 = 1.7183. \end{cases}$$

Решение системы

$$\varphi_1 = 1; \quad \varphi_2 = 0.9999; \quad \varphi_3 = 0.9996.$$

Точное решение интегрального уравнения -  $\varphi(x) \equiv 1$ .

### 12.1.2. Решение интегральных уравнений с помощью резольвенты

Доказано, что решение (2) можно представить в виде

$$\varphi(x) = f(x) + \lambda \cdot \int_a^b R(x, s, \lambda) f(s) ds. \quad (16)$$

$$R(x, s, \lambda) = \frac{D(x, s, \lambda)}{D(\lambda)} - \text{резольвента Фредгольма}, \quad (17)$$

$$D(\lambda) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} C_m \lambda^m - \text{определитель Фредгольма}, \quad (18)$$

$$C_m = \int_a^b \dots \int_a^b \begin{vmatrix} K(\alpha_1, \alpha_1) & K(\alpha_1, \alpha_2) & \dots & K(\alpha_1, \alpha_m) \\ K(\alpha_2, \alpha_1) & K(\alpha_2, \alpha_2) & \dots & K(\alpha_2, \alpha_m) \\ \dots & \dots & \dots & \dots \\ K(\alpha_m, \alpha_1) & K(\alpha_m, \alpha_2) & \dots & K(\alpha_m, \alpha_m) \end{vmatrix} d\alpha_1 \dots d\alpha_m, \quad (19)$$

$$C_0 = 1,$$

$$D(x, s, \lambda) = \sum_{m=0}^{\infty} (-1)^m \frac{\lambda^m}{m!} B_m(x, s), \quad (20)$$

$$B_0(x, s) = K(x, s). \quad (21)$$

$$B_m(x, s) = \int_a^b \dots \int_a^b \begin{vmatrix} K(x, s) & K(x, \alpha_1) & \dots & K(x, \alpha_m) \\ K(\alpha_1, s) & K(\alpha_1, \alpha_1) & \dots & K(\alpha_1, \alpha_m) \\ \dots & \dots & \dots & \dots \\ K(\alpha_m, s) & K(\alpha_m, \alpha_1) & \dots & K(\alpha_m, \alpha_m) \end{vmatrix} d\alpha_1 \dots d\alpha_m, \quad (22)$$

Отметим, что

$$C_n = \int_a^b B_{n-1}(x, x) dx, \quad n > 0, \quad (C_0 = 1). \quad (23)$$

Таким образом возможно получить резольвенту (разрешающее ядро) интегрального уравнения посредством соотношений (17) - (23).

Значения  $\lambda$ , для которых существует резольвента, называются регулярными, а значения  $\lambda$ , для которых резольвента не существует ( $D(\lambda) = 0$ ) — характеристическими.

Итак, если значение  $\lambda$  - регулярное, то интегральное уравнение

$$\varphi(x) = \lambda \int_a^b R(x, s, \lambda) f(s) ds + f(x)$$

с непрерывным ядром  $(x, s)$  и правой частью  $f(x)$  имеет единственное непрерывное решение

$$\varphi(x) = f(x) + \lambda \int_a^b R(x, s, \lambda) f(s) ds.$$

Если  $\lambda$  - регулярное, то однородное уравнение

$$\varphi(x) = \lambda \int_a^b K(x, s) \varphi(s) ds$$

имеет только тривиальное решение  $\varphi(x) \equiv 0$ .

Однородное уравнение имеет нетривиальное решение, в том и только в том случае, если  $\lambda$  - характеристическое число. Решение называют собственной или фундаментальной функциями ядра  $K(x, s)$ .

#### Пример 2 [39].

Построить резольвенту ядра

$$K(x, s) = e^{x-s}, \quad 0 \leq x, s \leq 1.$$

Решение.

$$C_0 = 1, \quad B_0 = e^{x-s}.$$

Согласно (22) и (23)  $C_1 = \int_0^1 B_0(\alpha_1, \alpha_1) d\alpha_1, \quad C_1 = 1$ .

$$B_1(x, s) = \int_0^1 \begin{vmatrix} e^x e^{-s} & e^x e^{-\alpha_1} \\ e^{\alpha_1} e^{-s} & e^{\alpha_1} e^{-\alpha_1} \end{vmatrix} d\alpha_1 = e^x \int_0^1 \begin{vmatrix} e^{-s} & e^{-\alpha_1} \\ e^{-s} & e^{-\alpha_1} \end{vmatrix} d\alpha_1 = 0.$$

$$C_2 = \int_0^1 B_1(\alpha_1, \alpha_1) d\alpha_1 = 0, \quad B_k(x, s) = 0, \quad C_k = 0, \quad k \geq 2.$$

$$D(\lambda) = \sum_{m=0}^1 \frac{(-1)^m}{m!} C_m \lambda^m = 1 - \lambda,$$

$$D(x, s, \lambda) = \sum_{m=0}^1 (-1)^m \frac{\lambda^m}{m!} B_m(x, s) = e^{x-s}.$$

Таким образом, резольвента ядра  $K(x, s) = e^{x-s}$  равна

$$R(x, s, \lambda) = \frac{e^{x-s}}{1-\lambda}.$$

Для всех  $\lambda \neq 1$  интегральное уравнение

$$\varphi(x) = \lambda \int_0^1 e^{x-s} \varphi(s) ds + f(x)$$

имеет единственное непрерывное решение (при непрерывной на  $[0; 1]$  функции  $f(x)$ ).

$$\varphi(x) = f(x) + \frac{\lambda}{1-\lambda} \int_0^1 e^{x-s} f(s) ds.$$

При  $\lambda = 1$  однородное уравнение  $\varphi(x) = \int_0^1 e^{x-s} \varphi(s) ds$  имеет нену-

левое решение  $\varphi(x) = e^x$  - собственную функцию ядра  $K(x, s) = e^{x-s}$ .

Указанный общий способ построения резольвенты для любого значения  $\lambda$  сложен для практического использования из-за кратных интегралов (22), определяющих коэффициенты рядов (18) и (20).

### 12.1.3. Метод последовательных приближений

Перепишем уравнение (2) в виде

$$\varphi(x) = f(x) + \lambda \int_a^b K(x, s) \varphi(s) ds.$$

Пусть начальное приближение  $\varphi_0(x) = f(x)$ . Тогда для  $n$ -го приближения имеем

$$\varphi_n(x) = f(x) + \sum_{j=1}^n \lambda^j \int_a^b K_j(x, s) f(s) ds = f(x) + \lambda \int_a^b \left( \sum_{j=1}^n \lambda^{j-1} K_j(x, s) \right) f(s) ds, \quad (24)$$

где  $n \in N$ .

Итерированные ядра определяются соотношениями

$$K_1(x, s) = K(x, s), \quad K_j(x, s) = \int_a^b K(x, s) K_{j-1}(x, s) ds, \quad j = 2, 3, \dots \quad (25)$$

$$\text{При } |\lambda| < \frac{1}{B_k} = \left( \int_a^b \int_a^b |K(x, s)|^2 dx ds \right)^{-\frac{1}{2}} \quad (26)$$

ряд Неймана  $\sum_{j=1}^{\infty} \lambda^{j-1} K_j(x, s)$  сходится к резольвенте.

Таким образом при  $n \rightarrow \infty$

$$\varphi(x) = f(x) + \lambda \int_a^b R(x, s, \lambda) f(s) ds, \quad (27)$$

$$R(x, s, \lambda) = \sum_{j=1}^{\infty} \lambda^{j-1} K_j(x, s), \quad (28)$$

что позволяет выразить решение уравнения через резольвенту. Отметим важность выполнения условия (26).

**Пример 3 [66].**

Найти резольвенту и решение интегрального уравнения

$$\varphi(x) - \frac{1}{\ln 2} \int_0^1 \frac{x}{1+s^2} \varphi(s) dx = 1+x^2$$

методом итерированных ядер.

*Решение.*

$$K(x, s) = \frac{x}{1+s^2}, \quad K_1(x, s) = K(x, s) = \frac{x}{1+s^2}.$$

$$\text{В силу (25)} \quad K_2(x, s) = \int_0^1 \frac{x}{1+t^2} \cdot \frac{t}{1+s^2} dt = \frac{\ln 2}{2} \cdot \frac{x}{1+s^2},$$

$$K_3(x, s) = \int_0^1 K(x, t) K_2(t, s) dt = \frac{\ln 2}{2} \int_0^1 \frac{x}{1+t^2} \cdot \frac{t}{1+s^2} dt = \left(\frac{\ln 2}{2}\right)^2 \frac{x}{1+s^2},$$

$$\dots \dots \dots$$

$$K_j(x, s) = \left(\frac{\ln 2}{2}\right)^{j-1} \frac{x}{1+s^2}.$$

Поэтому резольвента ядра равна

$$R(x, s, \lambda) = \sum_{j=1}^{\infty} \lambda^{j-1} K_j(x, s) = \sum_{j=1}^{\infty} \left(\frac{\ln 2}{2} \lambda\right)^{j-1} \frac{x}{1+s^2} = \frac{1}{1 - \frac{\ln 2}{2} \lambda} \cdot \frac{x}{1+s^2},$$

причем ряд сходится в области  $|\lambda| < \frac{2}{\ln 2}$ .

В рассматриваемом случае

$$B_k^2 = \int_0^1 \int_0^1 |K(x, s)|^2 dx ds = \int_0^1 \int_0^1 \frac{x^2}{(1+s^2)^2} dx ds = \frac{\pi+2}{24}.$$

Условие (26) приводит к неравенству  $|\lambda| < 2\sqrt{\frac{6}{\pi+2}}$ , но  $2\sqrt{\frac{6}{\pi+2}} < \frac{2}{\ln 2}$ ,

то есть область сходимости ряда Неймана для резольвенты шире.

$$R(x, s, \frac{1}{\ln 2}) = 2 \frac{x}{1+s^2} \quad (\text{из формулы (28) при } \lambda = \frac{1}{\ln 2}).$$

Решение интегрального уравнения

$$\varphi(x) = 1+x^2 + \frac{1}{\ln 2} \int_0^1 2 \frac{x}{1+s^2} (1+x^2) ds = 1 + \frac{2}{\ln 2} x + x^2.$$

**12.1.4. Метод вырожденных ядер**

Ядро  $K(x, s)$  называется вырожденным, если оно имеет вид

$$K(x, s) = \sum_{i=1}^m \alpha_i(x) \beta_i(s), \quad (29)$$

где  $\alpha_i(x)$  и  $\beta_i(x)$ ,  $i = \overline{1, m}$  - системы линейно-независимых между собой функций.

Уравнение (2) принимает вид

$$\varphi(x) - \lambda \sum_{i=1}^m \alpha_i(x) \int_a^b \beta_i(s) f(s) ds = f(x) \quad (30)$$

Обозначим

$$C_i = \int_a^b \beta_i(s) f(s) ds, \quad (31)$$

тогда решение уравнения (30)

$$\varphi(x) = f(x) + \lambda \sum_{i=1}^m C_i \alpha_i(x). \quad (32)$$

Неизвестные коэффициенты  $C_i$  могут быть получены из системы линейных уравнений

$$C_i - \lambda \sum_{j=1}^m a_{ij} C_j = f_i, \quad i = \overline{1, m} \quad (33)$$

различными методами.

$$a_{ij} = \int_a^b \alpha_i(s) \beta_j(s) ds, \quad j, i = \overline{1, m}, \quad (34)$$

$$f_i = \int_a^b \beta_i(s) f(s) ds, \quad i = \overline{1, m}. \quad (35)$$

Если система (33) несовместна, то интегральное уравнение решений не имеет.

**Алгоритм решения неоднородного уравнения с вырожденным ядром**

1. Вычисление интегралов (34) и (35).
2. Составление системы линейных уравнений (33).
3. Решение системы уравнений.
4. Запись решения уравнения в виде (32).

**Пример 4.**

Решить уравнение

$$\varphi(x) - \int_0^{\frac{\pi}{4}} \operatorname{tg} s \varphi(s) ds = \cos^2 x.$$

**Решение.**Ядро  $K(x, s) = \operatorname{tg} s$  - вырожденное. Решение ищем в виде

$$\varphi(x) = f(x) + C_1, \quad i=1, \quad \alpha_1 = 1, \quad \beta_1 = \operatorname{tg} s, \quad f(x) = \cos^2 x, \quad \lambda = 1.$$

Система уравнений содержит в данном случае одно уравнение  $C_1 - a_{11} C_1 = f_1$  (из (33)). По формулам (34) и (35)

$$a_{11} = \int_0^{\frac{\pi}{4}} \operatorname{tg} s ds = \frac{1}{2} \ln 2, \quad f_1 = \int_0^{\frac{\pi}{4}} \operatorname{tg} s \cos^2 s ds = \frac{1}{4}.$$

Тогда из соотношения  $C_1(1 - a_{11}) = f_1$  следует  $C_1 = \frac{1}{2 \cdot (2 - \ln 2)}$ .

$$\text{Решение уравнения } \varphi(x) = \cos^2 x + \frac{1}{2 \cdot (2 - \ln 2)}.$$

**Алгоритм решения однородного уравнения с вырожденным ядром**

$$\varphi(x) - \lambda \sum_{i=1}^m \alpha_i(x) \int_a^b \beta_i(s) \varphi(s) ds = 0. \quad (36)$$

1. Вычисление интегралов (34).
2. Составление определителя

$$D(\lambda) = \begin{vmatrix} 1 - \lambda a_{11} & -\lambda a_{12} & \dots & -\lambda a_{1m} \\ -\lambda a_{21} & 1 - \lambda a_{22} & \dots & -\lambda a_{2m} \\ \dots & \dots & \dots & \dots \\ -\lambda a_{m1} & -\lambda a_{m2} & \dots & 1 - \lambda a_{mm} \end{vmatrix}.$$

Характеристические числа  $\lambda_k$  однородного уравнения (36), являются корнями уравнения  $D(\lambda) = 0$ , количество которых  $p, 1 \leq p \leq m$ ,  $k = \overline{1, p}$ .

3. Решение системы однородных уравнений

$$\begin{cases} (1 - \lambda a_{11})C_1 - \lambda a_{12}C_2 - \dots - \lambda a_{1m}C_m = 0, \\ -\lambda a_{21}C_1 - (1 - \lambda a_{22})C_2 - \dots - \lambda a_{2m}C_m = 0, \\ \dots \\ -\lambda a_{m1}C_1 - \lambda a_{m2}C_2 - \dots - (1 - \lambda a_{mm})C_m = 0 \end{cases} \quad (37)$$

относительно  $C_1^{(k)}, C_2^{(k)}, \dots, C_m^{(k)}, k = \overline{1, p}$ , соответствующему какому-либо  $k$ -у характеристическому числу  $\lambda_k$ .**4. Определение собственных функций**

$$\varphi_k(x) = \lambda \sum_{i=1}^m C_i^{(k)} \alpha_i(x), \quad k = \overline{1, p}. \quad (38)$$

**Пример 5.**

Решить уравнение

$$\varphi(x) - 4 \int_0^1 x s^2 \varphi(s) ds = 0.$$

**Решение.**

$$i=1, \quad \alpha_1 = 4x, \quad \beta_1 = s^2, \quad \lambda = 4.$$

$$\text{Вычислим } a_{11} = \int_0^1 \alpha_1(s) \beta_1(s) ds = \int_0^1 s \cdot s^2 ds = \frac{1}{4}.$$

Из уравнения  $D(\lambda) = 1 - \lambda = 0$  следует  $\lambda = 1$ . Из системы (37) $(1 - \lambda)C_1 = 0$ , при  $\lambda = 1$ ,  $C_1$  произвольная постоянная.Решение исходного интегрального уравнения  $\varphi(x) = Cx$ .**Пример 6 [39]**

Показать, что интегральное уравнение

$$\varphi(x) - \lambda \int_0^1 (3x-2) s \varphi(s) ds = 0$$

не имеет характеристических чисел и собственных функций.

**Решение.**

$$\varphi(x) = \lambda (3x-2) \int_0^1 s \varphi(s) ds.$$

Пусть  $C = \int_0^1 s \varphi(s) ds$ , тогда  $\varphi(x) = C \lambda (3x-2)$ .Подставим в выражение для  $C$  и получим  $C = \int_0^1 s C \lambda (3s-2) ds$ ,

$$C \left( 1 - \lambda \int_0^1 (3s^2 - 2s) ds \right) = 0, \text{ откуда следует } C = 0, \varphi(x) \equiv 0, \text{ значит}$$

однородное уравнение имеет только тривиальное решение при любых  $\lambda$ , то есть не имеет характеристических чисел и собственных функций.**12.1.5. Способы аппроксимации ядер вырожденными**

Рассмотрим кратко два способа такой аппроксимации, более приспособленных к табличному заданию аппроксимируемого ядра.

1) Способ Бетмена [8].

$\tilde{K}(x, s)$  - вырожденное ядро,  $K(x, s)$  - аппроксимирующее ядро,  $x_1, x_2, \dots, x_n; s_1, s_2, \dots, s_n$  — некоторые точки отрезка  $[a, b]$ .

$$\tilde{K}(x, s) = K(x, s) - \frac{\begin{vmatrix} K(x, s) & K(x, s_1) & \dots & K(x, s_n) \\ K(x_1, s) & K(x_1, s_1) & \dots & K(x_1, s_n) \\ \dots & \dots & \dots & \dots \\ K(x_n, s) & K(x_n, s_1) & \dots & K(x_n, s_n) \end{vmatrix}}{\begin{vmatrix} K(x_1, s_1) & K(x_1, s_2) & \dots & K(x_1, s_n) \\ K(x_2, s_1) & K(x_2, s_2) & \dots & K(x_2, s_n) \\ \dots & \dots & \dots & \dots \\ K(x_n, s_1) & K(x_n, s_2) & \dots & K(x_n, s_n) \end{vmatrix}}, \quad (39)$$

Для ядра  $\tilde{K}(x, s)$ , согласно (39), резольвента  $\tilde{R}$  определяется из уравнения

$$\begin{vmatrix} \tilde{R}(x, s, \lambda) & K(x, s_1) & K(x, s_2) & \dots & K(x, s_n) \\ K(x_1, s) & A_{11} & A_{12} & \dots & A_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ K(x_n, s) & A_{n1} & A_{n2} & \dots & A_{nm} \end{vmatrix} = 0, \quad (40)$$

$$A_{jk} = K(x_j, s_k) - \lambda \int_a^b K(x_j, t) K(t, s_k) dt, \quad j, k = \overline{1, n}. \quad (41)$$

Решение уравнения может быть записано в виде

$$\varphi(x) = f(x) + \lambda \int_a^b \tilde{R}(x, s, \lambda) f(s) ds. \quad (42)$$

2) Комбинированный способ аппроксимации ядра  $K(x, s)$  вырожденным  $\tilde{R}(x, s)$ .

Задача квадратичных приближений в случае аппроксимации такова [12]: необходимо подобрать функции  $\alpha_i(x)$  и  $\beta_i(s)$ ,  $i = \overline{1, m}$ , таким образом, чтобы обратить в минимум

$$I = \int_a^b \int_a^b \left[ K(x, s) - \sum_{i=1}^m \alpha_i(x) \beta_i(s) \right]^2 dx ds. \quad (43)$$

**Алгоритм метода аппроксимации**

1. Положить в нулевом приближении  $\beta_i^{(0)}(s) = 1$ ,  $i = \overline{1, m}$ .
2. Для  $r$ -го приближения ( $r = 2, 3, \dots$ ) вычислить

$$\alpha_i^{(r-1)}(x_j) = \frac{\int_a^b K_{i-1}(x_j, s) \beta_i^{(r-1)}(s) ds}{\int_a^b [\beta_i^{(r-1)}(s)]^2 ds}, \quad j = \overline{1, n}, \quad i = \overline{1, m}, \quad (44)$$

$$\beta_i^{(r)}(s_j) = \frac{\int_a^b K_{i-1}(x, s_j) \alpha_i^{(r-1)}(x) dx}{\int_a^b [\alpha_i^{(r-1)}(x)]^2 dx}, \quad (45)$$

$$K_{i-1}(x, s) = \begin{cases} K(x, s), & i = 1, \\ K_{i-2}(x, s) - \alpha_{i-1}(x) \beta_{i-1}(s), & i = \overline{2, m}. \end{cases} \quad (46)$$

3. Вычислить ошибки.

$$\Delta_i^{(r)} = \max |K_i^{(r)}(x, s)|, \quad x, s \in [a, b]. \quad (47)$$

Итерационный процесс продолжается до выполнения условия

$$\frac{|\Delta_i^{(r)} - \Delta_i^{(r-1)}|}{\Delta_i^{(r)}} \leq E, \quad (E - \text{требуемая точность}), \text{ при котором } r = r_0.$$

Далее полагают  $K_i(x, s) = K_i^{(r_0)}(x, s)$ ,  $\alpha_i(x) = \alpha_i^{(r_0-1)}(x)$ ,

$\beta_i(s) = \beta_i^{(r_0)}(s)$  и увеличивают  $i$  на 1.

Увеличение  $i$  продолжается до тех пор, пока  $\Delta_i^{(r)} \geq E$ .

Рассмотренный процесс приближений сходится для широкого круга ядер. Оценка погрешности приведена в [8].

Дальнейшие вычисления проводятся в соответствии с формулами (29) – (35).

Достоинством метода вырожденных ядер является тот факт, что порядок систем алгебраических уравнений, которые приходится решать для определения решения интегрального уравнения, как правило, значительно более низкий, чем при решении задачи методом квадратур. Математическое обоснование правомерности замены ядра интегрального уравнения близким к нему вырожденным приведено в [8]. В качестве вырожденного ядра, близкого к данному  $K(x, s)$ , можно брать отрезок ряда Тейлора для функции  $K(x, s)$ , отрезок ряда Фурье и так далее.

**Пример 7 [33].**

Решить уравнение

$$\varphi(x) = \sin x + \int_0^1 (1-x \cos xs) \varphi(s) ds$$

заменой его ядра на вырожденное.

*Решение.*

Разложение  $K(x, s) = 1 - x \cos(xs)$  в ряд:

$$K(x, s) = 1 - x + \frac{x^3 s^2}{2} - \frac{x^5 s^4}{24} + \dots$$

$$\text{Пусть } \tilde{R}(x, s) = 1 - x + \frac{x^3 s^2}{2}.$$

$$\text{Решая уравнение } \tilde{\varphi}(x) = \sin x + \int_0^1 (1 - x + \frac{x^3 s^2}{2}) \tilde{\varphi}(s) ds,$$

$$\text{получим } \tilde{\varphi}(x) = \sin x + C_1(1-x) + C_2 x^3, \quad C_1 = \int_0^1 \tilde{\varphi}(s) ds, \quad C_2 = \frac{1}{2} \int_0^1 s^2 \tilde{\varphi}(s) ds.$$

$$C_1 = \int_0^1 (\sin t + C_1(1-t) + C_2 t^3) dt = \frac{1}{2} C_1 + \frac{1}{4} C_2 + 1 - \cos 1,$$

$$C_2 = \frac{1}{2} \int_0^1 (t^2 \sin t + C_1(t^2 - t^3) + C_2 t^5) dt = \frac{1}{24} C_1 + \frac{1}{12} C_2 + \sin 1 - 1 + \frac{1}{2} \cos 1,$$

Решая систему

$$\begin{cases} \frac{1}{2} C_1 - \frac{1}{4} C_2 = 1 - \cos 1, \\ -\frac{1}{24} C_1 + \frac{11}{12} C_2 = \sin 1 + \frac{1}{2} \cos 1 - 1. \end{cases}$$

$$\text{найдем } C_1 = 1.0031, \quad C_2 = 0.1674$$

$$\tilde{\varphi}(x) = \sin x + 1.0031(1-x) + 0.1674 x^3.$$

Если воспользоваться указанными выше оценками и учесть, что точное решение данного уравнения есть  $\varphi(x) \equiv 1$ , то получим

$$\|\varphi - \tilde{\varphi}\| < 0.016.$$

Рассматриваемые ниже методы моментов и коллокаций основаны на представлении приближенного решения  $\tilde{\varphi}(x)$  в виде

$$\tilde{\varphi}(x) = \Phi(x, \bar{C}), \quad \bar{C} = (C_1, C_2, \dots, C_n), \quad (48)$$

$C_i$  - неопределенные коэффициенты, определяемые из условия:

$$r = \varphi(x) - \lambda \int_a^b K(x, s) \varphi(s) dx - f(x) = 0. \quad (49)$$

Близость невязки к нулю не обозначает обязательной близости функций  $\varphi(x)$  и  $\tilde{\varphi}(x)$ .

Требует также рассмотрения вопрос о сходимости приближенного решения  $\tilde{\varphi}(x) \equiv \Phi(x, \bar{C})$  к точному, то есть выяснение условий, при которых  $\lim_{n \rightarrow \infty} \Phi(x, \bar{C}) = y(x)$ .

Эти вопросы рассматриваются в [31].

### 12.1.6. Метод моментов

В методе моментов (метод Бубнова-Галеркина) приближенное решение ищется в виде суммы  $f(x)$  и линейной комбинации заранее выбранных линейно-независимых между собой функций  $\psi_i(x)$ ,  $i = \overline{1, n}$ .

Например, система полиномов Лежандра  $1, x, \frac{3x^2-1}{2}, \dots$  или

$\psi_i(x) = x^{i-1}$  с неопределенными коэффициентами  $C_i$ .

Формируют функцию

$$\Phi(x, C_i) = \sum_{i=1}^n C_i \left( \psi_i(x) - \lambda \int_a^b K(x, s) \psi_i(s) ds \right) - \lambda \int_a^b K(x, s) f(s) ds. \quad (50)$$

Функция  $\Phi$  должна быть ортогональна ко всем функциям  $\psi_i(x)$  на отрезке  $[ab]$ .

Получаем систему линейных уравнений

$$\sum_{j=1}^n C_j (\alpha_{ij} - \lambda \beta_{ij}) = \lambda \gamma_i, \quad (51)$$

$$\alpha_{ij} = \int_a^b \psi_i(x) \psi_j(x) dx,$$

$$\beta_{ij} = \int_a^b dx \int_a^b K(x, s) \psi_i(x) \psi_j(s) ds, \quad (52)$$

$$\gamma_i = \int_a^b dx \int_a^b K(x, s) \psi_i(x) f(s) ds.$$

Решение интегрального уравнения

$$\varphi(x) = f(x) + \lambda \int_a^b K(x, s) \varphi(s) ds$$

$$\tilde{\varphi}(x) = f(x) + \sum_{i=1}^n C_i \psi_i(x) \quad (53)$$

### Пример 8.

Решить методом моментов уравнение

$$\varphi(x) = x + \int_{-1}^1 x s \varphi(s) ds.$$

Решение.

Выбираем  $\psi_1(x) = 1$ ,  $\psi_2(x) = x$ . Решение ищем в виде

$$\tilde{\varphi}(x) = x + C_1 + C_2 x = C_1 + x(1 + C_2). \quad (54)$$

Составляем систему уравнений (51)

$$\begin{cases} C_1(\alpha_{11} - \beta_{11}) + C_2(\alpha_{12} - \beta_{12}) = \gamma_1, \\ C_1(\alpha_{21} - \beta_{21}) + C_2(\alpha_{22} - \beta_{22}) = \gamma_2 \end{cases}$$

По формулам (52) находим

$$\beta_{11} = \int_{-1}^1 dx \int_{-1}^1 x s ds = 0, \quad \beta_{12} = \int_{-1}^1 dx \int_{-1}^1 x s^2 ds = 0,$$

$$\beta_{21} = \int_{-1}^1 dx \int_{-1}^1 x^2 s ds = 0, \quad \beta_{22} = \int_{-1}^1 dx \int_{-1}^1 x^2 s^2 ds = \frac{4}{9},$$

$$\gamma_1 = \int_{-1}^1 dx \int_{-1}^1 x s^2 ds = 0, \quad \gamma_2 = \int_{-1}^1 dx \int_{-1}^1 x^2 s^2 ds = \frac{4}{9}.$$

Решение системы

$$\begin{cases} C_1(2-0) + C_2 \cdot 0 = 0, \\ C_2\left(\frac{2}{3} - \frac{4}{9}\right) = \frac{4}{9}, \end{cases} \quad C_1 = 0, \quad C_2 = 2.$$

В соответствии с (53) решение интегрального уравнения  $\varphi(x) = 3x$ .

Заметим, что ядро уравнения - вырожденное, метод моментов в этом случае дает точное решение.

### 12.1.7. Метод коллокаций

Приближенное решение интегрального уравнения методом коллокаций представляется функцией вида  $\tilde{\varphi}(x) \equiv \Phi(x)$ .

Требуется, чтобы невязка (49) обращалась в нуль на отрезке  $[a, b]$  в заданной системе точек  $x_j$ ,  $j = \overline{1, n}$ ,  $a \leq x_1 < x_2 < \dots < x_n \leq b$ , то есть выполнялось условие

$$r(x_j, C_i) = 0. \quad (55)$$

Неопределенные коэффициенты  $C_i$  определяются из системы уравнений ((56), (57))

$$\sum_{i=1}^n C_i y_i(x_j, \lambda) = f(x_j), \quad j = \overline{1, n}, \quad (56)$$

$$y_i(x, \lambda) = \varphi_i(x) - \lambda \int_a^b K(x, s) \varphi_i(s) ds. \quad (57)$$

Если определитель системы (56)  $\det[y_i(x_j, \lambda)] \neq 0$ , то решение интегрального уравнения

$$\tilde{\varphi}(x) = \sum_{i=1}^n C_i \psi_i(x). \quad (58)$$

### Пример 9.

$$\text{Решить уравнение } \varphi(x) = x + \int_{-1}^1 x \cdot s \cdot \varphi(s) ds$$

методом коллокаций.

Решение.

Выберем точки коллокации  $x_1 = -1, x_2 = 1$ , систему ортогональных функций  $\varphi_1 = 1, \varphi_2(x) = x$ .

Решение ищем в виде  $\tilde{\varphi}(x) = C_1 + C_2 x$ .

Определяем  $y_i(x)$  по формулам (57).

$$y_1(x) = 1 - \int_{-1}^1 x s ds = 1, \quad y_2(x) = x - \int_{-1}^1 x s^2 ds = \frac{1}{3} x.$$

Система уравнений (56)

$$\begin{cases} C_1 y_1(x_1) + C_2 y_2(x_1) = f(x_1), \\ C_1 y_1(x_2) + C_2 y_2(x_2) = f(x_2), \end{cases} \quad \begin{cases} C_1 - \frac{1}{3} C_2 = -1, \\ C_1 + \frac{1}{3} C_2 = 1, \end{cases}$$

$$C_1 = 0, \quad C_2 = 3.$$

Получено решение  $\varphi(x) = 3x$ .

### 12.2. Линейные интегральные уравнения Фредгольма 1-го рода

Общий вид уравнения Фредгольма 1-го рода

$$\int_a^b K(x, s) \varphi(s) ds = f(x). \quad (59)$$

$K(x, s)$  - ядро уравнения,  $f(x)$  - свободный член - известные функции;  $\varphi(s)$  - искомая функция;  $a, b$  - конечные или бесконечные пределы интегрирования.

В общем случае решение уравнения (59) - достаточно сложная задача. Уравнения 1-го рода часто встречаются в технических приложениях.

Задача восстановления сигнала описывается уравнением

$$\int_a^b h(x, s) \varphi(s) ds = f(x), \quad c \leq x \leq d,$$

$h(x, s)$  - функция отклика на единичный импульс или весовая функция линейной системы.

Эта задача имеет различные конкретные реализации: восстановление сигнала в теории автоматического управления, задача обработки изображений, задача редукции измерений за характеристику направленности антенны, задача оптимальной линейной фильтрации, амплитудный синтез непрерывной антенны и так далее.

Существенным моментом решения интегральных уравнений 1-го рода является ее некорректность. Одним из условий корректности любой задачи является требование устойчивости решения. Задача решения уравнения (59) является некорректной, в первую очередь, по причине нарушения требования устойчивости: даже очень малые относительные ошибки могут практически обесценить численное решение, которое не будет иметь ничего общего с точным решением.

Основным методом решения (59) является метод регуляризации, суть которого состоит в замене исходной задачи задачей решения уравнения Фредгольма 2-го рода (при  $q=0$ ) или интегродифференциального уравнения (при  $q \neq 0$ ):

$$\alpha \left[ \varphi_\alpha(t) - q \varphi_\alpha^*(t) + \int_a^b R(t, s) \varphi_\alpha(s) ds \right] = F(t), \quad (60)$$

$$R(t, s) = R(s, t) = \int_c^d \tilde{K}(x, t) \tilde{K}(x, s) dx, \quad (61)$$

$$F(t) = \int_c^d \tilde{K}(x, t) \tilde{f}(x) dx, \quad (62)$$

краевые условия имеют вид

$$\varphi'_\alpha(a) = \varphi'_\alpha(b) = 0. \quad (63)$$

Говорят о регуляризации нулевого порядка, если  $q=0$ ; регуляризации первого порядка, если  $q \neq 0$ .

$$\tilde{K}(x, s) \text{ и } \tilde{f}(x) - \text{приближенные значения } K(x, s) \text{ и } f(x), \text{ такие, что} \\ \|\tilde{f}(x) - f(x)\|_{L_2} \leq \delta, \quad (64)$$

$$\|\tilde{K}(x, s) - K(x, s)\| \leq \xi. \quad (65)$$

Параметр  $\alpha$  - оптимальное значение, параметр регуляризации.

Уравнение (60) - условие минимума сглаживающего функционала Тихонова. Подробнее в [17], [56].

Выбор параметра  $\alpha$  - отдельная задача. Один из самых эффективных - способ невязки, изложенный в [30], [12].

Применительно к уравнениям Фредгольма 1-го рода метод обобщенной невязки сводится к решению уравнения

$$\gamma(\alpha) \equiv \beta(\alpha) - z(\alpha) = 0, \quad (66)$$

где обозначено:

$$\beta(\alpha) = \int_a^d \left[ \int_a^b K(x, s) \varphi_\alpha(s) ds - \tilde{f}(x) \right]^2 dx, \quad (67)$$

$$z(\alpha) = \left( \delta + \xi \sqrt{\gamma(\alpha)} \right)^2 + \mu, \quad (68)$$

$$\gamma(\alpha) = \int_a^b \varphi_\alpha^2(s) ds. \quad (69)$$

Численный алгоритм решения уравнения (60) с выбором параметра  $\alpha$  по обобщенной невязке (66) - (69) приведен в [12], причем указаны два способа решения: метод конечных сумм и разностей и метод регуляризации Тихонова.

### 12.3. Интегральные уравнения с разностными ядрами

Некоторые задачи связи приводят к необходимости решения интегральных уравнений, ядра которых зависят от разности или разности и суммы аргументов.

Статистическая теория приема сигналов рассматривает приемное устройство как некоторый фильтр, выделяющий полезную информацию об интересующем нас параметре из принимаемой смеси сигнала и шума. Обозначим оцениваемый параметр сигнала  $S$  через  $l$ . Функция  $v(t, l)$  [33], представляющая собой импульсную характеристику оптимального линейного фильтра, определяется из интегрального уравнения

$$\int_0^T K(t-t_2) v(t, l) dt = S(t_2, l),$$

$K(t-t_2)$  - функция корреляции исходного шума.

Интегральное уравнение Фредгольма 2-го рода

$$\varphi(x) = f(x) + \int_{-\infty}^{+\infty} K(x-s) \varphi(s) ds, \quad (70)$$

$$f(x) \in L_2(-\infty, \infty) \text{ и } K(x) \in L_2(-\infty, \infty)$$

может быть решено с помощью интегрального преобразования Фурье.

Решением (70) является



$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{F(w)}{1 - \sqrt{2\pi} \tilde{K}(w)} e^{ixw} dw, \quad (71)$$

$F(w), K(w)$  - преобразование Фурье функций  $f(x)$  и  $K(x)$ .

Аналогично, для уравнения Фредгольма 1-го рода с ядром, зависящим от разности аргументов

$$\int_{-\infty}^{+\infty} K(x-s) \varphi(s) ds = f(x) \quad (72)$$

решением является

$$\varphi(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{F(w)}{\tilde{K}(w)} e^{ixw} dw. \quad (73)$$

Численный метод решения уравнения

$$\varphi(x) + \int_r^w K(x-s) \varphi(s) ds = f(x) \quad (74)$$

на конечном отрезке  $[r, w]$  приведен в [33].

Доказано, если известны (найжены) приближенные значения решения уравнения

$$\varphi(x) + \int_0^w \varphi(s) K(x-s) ds = f(x) \quad (75)$$

при  $f(x) = K(x)$  и  $f(x) = K(x-s)$ , то можно построить решение (75) для любой правой части. Интеграл вычисляется с помощью квадратурных формул Гаусса, решая систему (10). Пусть эти решения равны соответственно  $\tilde{g}(x_i)$  и  $\tilde{h}(x_i)$ ,  $i = \overline{1, 6}$ .

Интерполяционные формулы

$$\tilde{g}(x) = K(x) + \sum_{j=1}^6 A_j K(x-x_j) \tilde{g}(x_j),$$

$$\tilde{h}(x) = K(x-w) + \sum_{j=1}^6 A_j K(x-x_j) \tilde{h}(x_j),$$

позволяют вычислить значения  $\tilde{g}(x_i)$  и  $\tilde{h}(x_i)$  в других точках,

$A_j$  - узлы формулы Гаусса для отрезка  $[0, w]$ ,  $j = \overline{1, 6}$ .

Используя  $\tilde{g}(x_i)$  и  $\tilde{h}(x_i)$  строится матрица с элементами  $\gamma(x_i, x_j)$ ; затем находят решение для любой правой части  $f(x)$  в виде

$$\varphi(x_i) = f(x_i) + \sum_{j=1}^6 A_j \gamma(x_i, x_j) f(x_j). \quad (76)$$

Функция  $\gamma(x, t)$  строится специальным образом так, чтобы можно было обратить оператор

$$s\varphi = \varphi(x) + \int_0^w \varphi(t) K(x-t) dt,$$

то есть найти решение интегрального уравнения.

Метод решения с разностно-суммарными ядрами изложен в [33].

## 12.4. Методы решения уравнений Вольтерры 2-го рода

Линейное одномерное уравнение Вольтерры 2-го рода

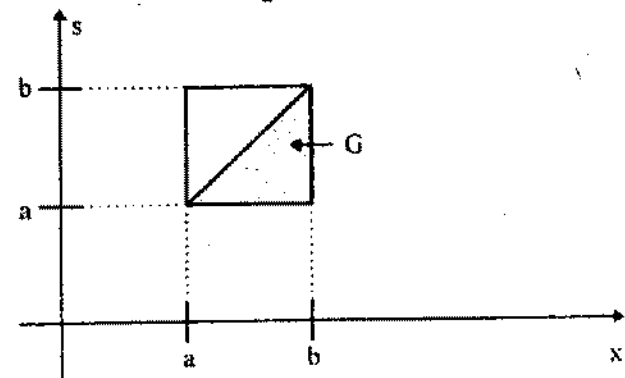
$$\varphi(x) - \int_a^x K(x, s) \varphi(s) ds = f(x), \quad x \in [a; b]. \quad (6)$$

Один из примеров приложения уравнений Вольтерры 2-го рода к задачам связи - задача анализа переходного процесса в электрической цепи с двумя инерционными элементами.

При заданных параметрах цепи  $R, L, C$ , начальном значении тока в катушке индуктивности  $I_0$  и напряжением на конденсаторе  $U_0 = U_C(0)$  в результате построения модели цепи получается уравнение

$$\frac{1}{R} U_C(t) + \int_0^t \left[ C + \frac{1}{L}(t-\tau) \right] U_C(\tau) d\tau = U_0.$$

Существование и единственность решения уравнения (6) связаны с определенными ограничениями на функции  $K(x, s)$  и  $f(x)$ . В частности, решение существует и единственно, если ядро непрерывно внутри и на границе области  $G$ , а  $f(x)$  на промежутке  $[a; b]$  имеет конечное число точек разрыва, причем функция может быть и неограниченной, если  $\int_a^b |f(t)| dt$  имеет конечное значение.



Если доопределить функцию  $K(x, s)$  при  $s > x$ , положив  $K(x, s) = 0$ ,  $x < s \leq b$ , то уравнение (6) можно рассматривать как частный случай уравнения Фредгольма 2-го рода (2) с ядром  $H(x, s)$ , определяемом следующим образом

$$H(x, s) = \begin{cases} K(x, s), & s \leq x, \\ 0, & s > x. \end{cases}$$

Таким образом, для уравнения Вольтерры 2-го рода применимы рассмотренные выше методы решения.

Уравнение (6) содержит интегральный оператор

$$A_B \varphi(x) = \int_a^x K(x, s) \varphi(s) ds, \quad x \leq s, \quad (77)$$

обладающий важным свойством: значение этого оператора при любом  $x$  определяется значениями функции  $\varphi(x)$  только при  $s \leq x$ .

Указанное свойство позволяет применять прием решения уравнения (6), состоящий в том, что решение может быть построено только на части промежутка  $[a, b]$ , например на отрезке  $a \leq x \leq x_1$ , при  $x > x_1$  пользоваться для решения выражением

$$\varphi(x) = \left[ \int_a^x K(x, s) \varphi(s) ds + f(x) \right] + \int_{x_1}^x K(x, s) \varphi(s) ds. \quad (78)$$

Некоторые частные виды уравнения (6):

• Уравнение с вырожденным ядром

$$K(x, s) = \sum_{i=1}^m \alpha_i(x) \beta_i(s), \quad (79)$$

$$\varphi(x) - \sum_{i=1}^m \alpha_i(x) \int_0^x \beta_i(s) \varphi(s) ds = f(x). \quad (80)$$

• Уравнение типа свертки

$$\varphi(x) - \int_0^x K(x-s) \varphi(s) ds = f(x), \quad x \in [0; b] \quad (81)$$

или

$$\varphi(x) - \int_{-\infty}^x K(x-s) \varphi(s) ds = f(x), \quad x \in (-\infty, \infty). \quad (82)$$

Интересно отметить связь задачи Коши для обыкновенных дифференциальных уравнений и интегральных уравнений Вольтерры.

Задача Коши для дифференциального уравнения с постоянными коэффициентами

$$\frac{d^n y(x)}{dx^n} + a_1 \frac{d^{n-1} y(x)}{dx^{n-1}} + \dots + a_n y(x) = f(x), \quad (83)$$

$$y(x_0) = y_0, \quad y'(x_0) = y_1, \quad \dots, \quad y^{(n-1)}(x_0) = y_{n-1}$$

сводится к интегральному уравнению вида (6) с ядром

$$K(x, s) = - \sum_{k=1}^n a_k \frac{(x-s)^{k-1}}{(k-1)!}, \quad \text{зависящим лишь от разности } (x-s) \text{ своих}$$

аргументов.

Задача Коши для дифференциального уравнения с переменными коэффициентами

$$\frac{d^n y(x)}{dx^n} + a_1(x) \frac{d^{n-1} y(x)}{dx^{n-1}} + \dots + a_n(x) y(x) = f(x), \quad x \geq 0,$$

$$y(0) = C_0, \quad y'(0) = C_1, \quad \dots, \quad y^{(n-1)}(0) = C_{n-1}$$

имеет эквивалентное представление в виде интегрального уравнения [39]

$$u(x) + \int_0^x K(x, s) u(s) ds = \varphi(x), \quad (84)$$

$$u(x) = \frac{d^n y(x)}{dx^n}, \quad \int_0^x u(s) ds + C_1 = \frac{d^{n-1} y(x)}{dx^{n-1}},$$

$$\int_a^x dx_1 \int_a^{x_1} dx_2 \dots \int_a^{x_{n-1}} z(x_n) dx_n = \frac{1}{(n-1)!} \int_0^x (x-s)^{n-1} z(s) ds, \quad (85)$$

$$y(x) = C_{n-1} \frac{x^{n-1}}{(n-1)!} + \dots + C_1 x + C_0 + \frac{1}{(n-1)!} \int_0^x (x-s)^{n-1} u(s) ds,$$

$$K(x, s) = \sum_{i=1}^n a_i(x) \frac{(x-s)^{i-1}}{(i-1)!}, \quad (86)$$

$$\begin{aligned} \varphi(x) = & f(x) - C_{n-1} a_1(x) + (C_{n-1} x + C_{n-2}) a_2(x) - \dots \\ & \dots - (C_{n-1} \frac{x^{n-1}}{(n-1)!} + \dots + C_1 x + C_0) a_n(x). \end{aligned} \quad (87)$$

Получение эквивалентных дифференциальных уравнений из интегральных возможно лишь в частных случаях.

Для линейных уравнений с вырожденным ядром (80) эквивалентная система дифференциальных уравнений имеет вид

$$v_i' = \beta_i(x) \left[ \sum_{j=1}^m \alpha_j(x) v_j(x) + f(x) \right], \quad v_i(x) = \int_0^x \beta_i(s) \varphi(s) ds,$$

с начальными условиями  $v_i(a) = 0$ ,  $i = \overline{1, n}$ .

$$\text{Решение} - y(x) = \frac{v_i'(x)}{\beta_i(x)}.$$

**Пример 10.**

Составить интегральное уравнение для решения задачи Коши

$$y'' - 5y' + 6y = 0, \quad y(0) = 0, \quad y'(0) = 1.$$

**Решение.**

Из записи уравнения (83) следует:  $a_1 = -5, \quad a_2 = 6, \quad f(x) \equiv 0.$

Полагаем  $\frac{d^2 y}{dx^2} = \varphi(x).$

Тогда, интегрируя, получаем  $\frac{dy}{dx} = \int_0^x \varphi(t) dt + y'(0) = 1 + \int_0^x \varphi(t) dt.$

С учетом (85)

$$y(x) = \int_0^x \left( \int_0^t \varphi(t) dt + 1 \right) dt = \int_0^x dt \int_0^t \varphi(t) dt + x = \int_0^x (x-t) \varphi(t) dt + x.$$

Подставляя  $\frac{d^2 y}{dx^2}, \quad \frac{dy}{dx}$  и  $y(x)$  в исходное уравнение, получим

$$\varphi(x) - 5 \left( 1 + \int_0^x \varphi(t) dt \right) + 6 \left( \int_0^x (x-t) \varphi(t) dt + x \right) = 0.$$

Окончательный вид интегрального уравнения

$$\varphi(x) = 5 - 6x + \int_0^x [5 - 6(x-t)] \varphi(t) dt.$$

**Пример 11.**

Сведем к задаче Коши решить уравнение

$$y(x) = 1 + \int_0^x \frac{ch(t)}{ch(x)} y(t) dt.$$

**Решение.**

Полагаем  $v(x) = \int_0^x ch(t) y(t) dt$ , тогда  $y(x) = 1 + \frac{1}{ch(x)} v(x).$

$$v'(x) = ch(x) y(x) = ch(x) \left( 1 + \frac{1}{ch(x)} v(x) \right)$$

или

$$v'(x) - v(x) = ch(x) - \text{искомое дифференциальное уравнение.}$$

Начальное условие:  $v(0) = 0.$

Решение уравнения

$$v(x) = \frac{1}{2} \cdot (x \cdot e^x + sh(x)), \text{ откуда } y(x) = 1 + \frac{1}{2} \cdot \frac{x \cdot e^x + sh(x)}{ch(x)}.$$

Получение вырожденного ядра связано с аппроксимацией функции нескольких переменных функциями, каждая из которых зависит от одной переменной.

Приведем несколько примеров такой аппроксимации [80].

- Ряд Тейлора.
- Применяется отрезок ряда Тейлора по той переменной (или нескольким переменным), где возможно разложение в ряд Тейлора ядра  $K(x, s)$

$$\tilde{K}(x, s) = \sum_{i=0}^m \frac{1}{i!} (x - x_0)^i K_{x_i}^{(i)}(x_0, s), \quad x, s \in [a, b], \text{ или}$$

$$\tilde{K}(x, s) = \sum_{i=0}^m \frac{1}{i!} (s - s_0)^i K_{s_i}^{(i)}(x, s_0), \quad x_0, s_0 \in [a, b]. \quad (88)$$

- Двойной ряд Тейлора

$$\tilde{K}(x, s) = \sum_{i=0}^m \sum_{j=0}^m a_{ij} (x - x_0)^i (s - s_0)^j, \quad (89)$$

$$a_{ij} = \frac{\partial^{i+j}}{\partial x^i \partial s^j} K(x, s) \text{ в точке } x_0, s_0 \in [a, b].$$

- Тригонометрический полином.

Полагая период равным  $2T$ , где  $T = b - a$ , получаем

$$\tilde{K}(x, s) = \frac{1}{2} a_0(s) + \sum_{i=1}^m a_i(s) \cos \frac{i\pi x}{T}, \quad (90)$$

$$a_i(s) - \text{коэффициенты Фурье, } i = 0, 1, 2, \dots \quad a_i(s) = \frac{2}{T} \int_a^b K(x, s) \cos \frac{i\pi x}{T} dx.$$

Другие методы замены ядра вырожденным приведены выше (для уравнений Фредгольма).

В уравнениях Вольтерры интеграл с переменным верхним пределом вносит трудности при замене интегральных операторов конечными суммами. Некоторые квадратурные методы решения уравнений Вольтерры основаны на совместном использовании замкнутых формул и формул открытого типа.

Важной особенностью интегралов с переменным верхним пределом является возможность использования каждого значения подынтегральной функции для вычисления многих значений функции, являющейся результатом интегрального преобразования. Это позволяет применять квадратурные формулы, обеспечивающие нужную точность при небольшом количестве вычислений.

Следует также отметить, что при большом (или заранее неизвестном) отрезке интегрирования за счет большого количества шагов интегрирования накапливаются погрешности.

Квадратурные формулы содержат значения подинтегральной функции в ряде точек.

$$\int_a^{x_i} K(x, s) \varphi(s) ds = \sum_{j=1}^i A_j K(x_i, x_j) \varphi(x_j), \quad i = \overline{1, n}.$$

Значение  $\varphi(x_i)$  можно определить из системы уравнений

$$\varphi_i - \sum_{j=1}^i A_j K_{ij} \varphi_j = f_i, \quad i = \overline{1, n}. \quad (91)$$

$$\varphi_i = (1 - A_i K_{ii})^{-1} \left( f_i + \sum_{j=1}^{i-1} A_j K_{ij} \varphi_j \right) \quad (92)$$

при условии  $1 - A_i K_{ii} \neq 0, \quad i = \overline{1, n},$

$$\tilde{\varphi}(x_i) = \varphi_i, \quad f(x_i) = f_i, \quad K(x_i, x_j) = K_{ij}. \quad (93)$$

В случае применения формулы трапеций с непостоянным шагом на сетке узлов по  $x$ , совпадающей с сеткой узлов по  $s (s_i = x_i)$

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

получаем решение уравнения (6) в виде рекуррентных формул [12]

$$\left\{ \begin{array}{l} \varphi_1 = f_1, \quad \varphi_2 = \frac{f_2 + \frac{h_1}{2} K_{21} \varphi_1}{1 - \frac{h_1}{2} K_{22}}, \\ \dots \\ \varphi_i = \frac{f_i + \frac{h_i}{2} K_{ii} \varphi_i + \sum_{j=2}^{i-1} \left( \frac{x_{j+1} - x_{j-1}}{2} \right) K_{ij} \varphi_j}{1 - \frac{h_i}{2} K_{ii}}, \quad i = \overline{3, n}. \end{array} \right. \quad (94)$$

При постоянном шаге  $h = const$  рекуррентные формулы имеют вид

$$\left\{ \begin{array}{l} \varphi_1 = f_1, \quad \varphi_i = \frac{f_i + h \sum_{j=1}^{i-1} A_j K_{ij} \varphi_j}{1 - \frac{h}{2} K_{ii}}, \quad i = \overline{2, n}, \end{array} \right. \quad (95)$$

$$x_i = a + (i-1)h, \quad n = \frac{b-a}{h} + 1, \quad A_j = \begin{cases} 0.5, & j=1, \\ 1, & j>1, \end{cases} \quad h \neq \frac{2}{K_{ii}}$$

Приближенные значения погрешностей решения уравнения (6)

$$\Delta \varphi_i \approx \sum_{j=1}^i A_j K_{ij} \Delta \varphi_j + R_i, \quad i = \overline{3, n},$$

$$\Delta \varphi_1 = \Delta \varphi_2 = 0. \quad (96)$$

Погрешность аппроксимации

$$R_i = -\frac{1}{12} \sum_{j=2}^i h_j^3 \left\{ \frac{\partial^2}{\partial s^2} [K(x_i, s) \varphi(s)] \right\}_{s=x_j} \quad (97)$$

может быть выражена в явном виде через  $h_i, K_{ij}$  [41].

Применение формулы трапеций с постоянным шагом для решения уравнения Вольтерры 2-го рода для случая вырожденного ядра позволяет получить расчетные формулы

$$\left\{ \begin{array}{l} \tilde{\varphi}(0) = f(0) = f_1, \\ \tilde{\varphi}_i = \frac{1}{1 - \frac{h}{2} \sum_{i=1}^m \alpha_{ii} \sum_{j=1}^{i-1} \beta_{ij} A_j \tilde{\varphi}_j} \left( f_i + h \sum_{i=1}^m \alpha_{ii} \sum_{j=1}^{i-1} \beta_{ij} A_j \tilde{\varphi}_j \right) \end{array} \right. \quad (98)$$

$$i = \overline{2, n}, \quad x_i = a + (i-1)h,$$

$$A_j = \begin{cases} 0.5, & j=1, \\ 1, & j>1. \end{cases}$$

### Пример 12.

Решить уравнение

$$\varphi(x) = 1 + \int_0^x \varphi(s) ds, \quad x \in [0; 0.4]$$

методом квадратур.

Решение.

Исходное уравнение - уравнение Вольтерры 2-го рода с вырожденным ядром  $K(x, s) = 1$ .

Воспользуемся формулами (98).

$m=1, \quad \alpha = \beta = 1, \quad f(x) = 1$ . Положим  $h = 0.2$ , тогда  $i = 2, 3$ .

$$\varphi_1 = 1, \quad \varphi_2 = \frac{1}{1 - \frac{h}{2}} (f_2 + h A_1 \varphi_1) = \frac{1}{0.9} (1 + 0.2 \cdot 0.5 \cdot 1) = 1.222,$$

$$\varphi_3 = \frac{1}{1 - \frac{h}{2}} (f_3 + h(A_1 \varphi_1 + A_2 \varphi_2)) = \frac{1}{0.9} (1 + 0.2 \cdot (0.5 \cdot 1 + 1.222)) = 1.4998.$$

Решение сведем в таблицу 1.

Таблица 1

$i$	1	2	3
$x_i$	0	0.2	0.4
$\varphi_i$	1	1.222	1.4998

Плодотворной является идея использования квадратурных формул открытого типа [44].

Итерационные методы решения Вольтерры имеют широкое применение прежде всего из-за простоты алгоритма. Кроме того, для уравнений Вольтерры методы простой итерации при условиях непрерывности функций  $f(x)$  на  $[a, b]$  и  $K(x, s)$  при  $a \leq x \leq b$ ,  $(s \leq x)$  сходятся и имеют факториальную скорость сходимости.

Рекуррентная формула простой итерации

$$\varphi_k(x) = f(x) + \int_a^x K(x, s) \varphi_{k-1}(s) ds, \quad k=1, 2, \dots \quad (99)$$

$$\text{Пусть } y_0(x) = f(x), \quad y_{k+1}(x) = \int_a^x K(x, s) y_k(s) ds. \quad (100)$$

$$\text{Искомое решение } \varphi(x) = \sum_{k=0}^{\infty} y_k(x). \quad (101)$$

Практически для окончания итерационного процесса используется условие

$$\frac{\|\varphi_k - \varphi_{k-1}\|}{\|\varphi_k\|} \leq E. \quad (102)$$

$E$  - заданная погрешность

**Пример 13.**

Решить уравнение

$$\varphi(x) = 1 + \int_0^x \varphi(s) ds$$

методом итерации, взяв  $\varphi_0(x) \equiv 0$ .

**Решение.**

$$K(x, s) \equiv 1, \quad \varphi_1(x) \equiv 1, \quad \text{по формулам (99) - (101)}$$

$$\varphi_2(x) = 1 + \int_0^x 1 \cdot ds = 1 + x, \quad \varphi_3(x) = 1 + \int_0^x (1+s) ds = 1 + x + \frac{x^2}{2},$$

$$\varphi_4(x) = 1 + \int_0^x (1+s+\frac{s^2}{2}) ds = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}.$$

Искомое решение

$$\varphi_n(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^{n-1}}{(n-1)!}, \quad \text{при } n \rightarrow \infty \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x.$$

Заметим, что точным решением уравнения является функция  $\varphi(x) = e^x$ . Приведем для сравнения значения решения по методу итерации и точного в ряде точек.

Таблица 2

$i$	1	2	3	4
$x_i$	0	0.2	0.4	0.6
$\varphi_i$	1	1.2	1.48	1.816
$e^{x_i}$	1	1.2214	1.4918	1.822

Решение с помощью резольвенты также возможно для уравнений Вольтерры.

Рассмотрим семейство уравнений ( $\lambda$  - фиксированный параметр)

$$\varphi(x) = f(x) + \lambda \int_a^x K(x, s) \varphi(s) ds. \quad (103)$$

Решая уравнение (103) методом последовательных приближений, взяв в качестве  $\varphi_0(x) = f(x)$ , получим

$$\varphi_1(x) = f(x) + \lambda \int_a^x K_1(x, s) f(s) ds, \quad K_1(x, s) = K(x, s).$$

$$\varphi_2(x) = f(x) + \lambda \int_a^x K(x, s) \varphi_1(s) ds =$$

$$= f(x) + \lambda \int_a^x K(x, s) \left[ f(s) + \lambda \int_a^s K_1(x, t) f(t) dt \right] ds =$$

$$= f(x) + \lambda \int_a^x K_1(x, s) f(s) ds + \lambda^2 \int_a^x \left( \int_t^x K(x, t) K_1(s, t) dt \right) f(s) ds =$$

$$= f(x) + \lambda \int_a^x K_1(x, s) f(s) ds + \lambda^2 \int_a^x K_2(x, s) f(s) ds,$$

$$K_2(x, s) = \int_s^x K(x, t) \cdot K_1(t, s) dt,$$

$$\varphi_n = f(x) + \sum_{j=1}^n \lambda^j \int_a^x K_j(x, s) f(s) ds =$$

$$= f(x) + \lambda \int_a^x \left( \sum_{j=1}^n \lambda^{j-1} K_j(x, s) \right) f(s) ds, \quad (104)$$

$$K_j(x, s) = \int_s^x K(x, t) K_{j-1}(t, s) dt, \quad j=2, 3, \dots, \quad (105)$$

$$K_1(x, s) = K(x, s).$$

$$\text{Функция } R(x, s, \lambda) = \sum_{j=1}^{\infty} \lambda^{j-1} K_j(x, s) \quad (106)$$

называется резольвентой ядра  $K(x, s)$ .

При любых фиксированных  $\lambda$  ряд (106) сходится равномерно относительно  $x \in [a, b]$  и  $s \in [0, x]$  к функции  $R(x, s, \lambda)$ , если ядро непрерывно.

В пределе, при  $\lambda \rightarrow \infty$  из (104) следует решение уравнения Вольтерры 2-го рода через резольвенту.

$$\varphi(x) = f(x) + \lambda \int_a^x R(x, s, \lambda) f(s) ds. \quad (107)$$

**Пример 14 [66].**

Найти резольвенту ядра  $K(x, s) = x$  и, используя ее, решить интегральное уравнение

$$\varphi(x) = x - \frac{1}{2} \int_0^x x \varphi(s) ds.$$

*Решение.*

Из рекуррентных соотношений (105) получаем

$$K_1(x, s) = x,$$

$$K_2(x, s) = \int_s^x K(x, t) K_1(t, s) dt = \int_s^x x s ds = x \frac{x^2 - s^2}{2},$$

$$K_3(x, s) = \int_s^x K(x, t) K_2(t, s) dt = \int_s^x x t \frac{t^2 - s^2}{2} dt = \frac{1}{2} x \left( \frac{x^2 - s^2}{2} \right)^2,$$

... ..

$$K_j(x, s) = x \frac{1}{(j-1)!} \left( \frac{x^2 - s^2}{2} \right)^{j-1}, \quad j = 1, 2, \dots$$

Находим резольвенту (106)

$$R(x, s, \lambda) = x \sum_{j=1}^{\infty} \frac{\lambda^{j-1}}{(j-1)!} \left( \frac{x^2 - s^2}{2} \right)^{j-1} = x e^{\lambda \frac{x^2 - s^2}{2}}.$$

Решение исходного уравнения при  $\lambda = -\frac{1}{2}$ ,  $f(x) = x$

$$\varphi(x) = x - \frac{1}{2} \int_0^x x e^{-\frac{x^2 - s^2}{4}} s ds = x - x e^{-\frac{x^2}{4}} \int_0^x d \left( e^{\frac{s^2}{4}} \right) = x e^{-\frac{x^2}{4}}.$$

## 12.5. Методы решения уравнений Вольтерры 1-го рода

К уравнениям Вольтерры 1-го рода сводится ряд задач, и, в первую очередь, часто встречающаяся на практике задача восстановления сигналов, поступающих на входы измерительных приборов и систем наблюдения, которые подвержены искажениям. Пусть задана некоторая линейная динамическая система,  $x(t)$  - ее входной,  $y(t)$  - выходной сигналы. Известно, что зависимость  $y(t)$  от  $x(t)$  может быть описана уравнением

$$y(t) = \int_{-\infty}^{+\infty} g(t, s) x(s) ds,$$

$g(t, s)$  - весовая функция, определяемая свойствами системы.

Если, в частности, выполнено условие физической реализуемости ( $g(t, s) = 0$  при  $s > t$ ) и система до момента  $t_0$  находилась в покое ( $x(t) = 0$  при  $t < t_0$ ), то получаем уравнение Вольтерры 1-го рода

$$y(t) = \int_{t_0}^t g(t, s) x(s) ds$$

Общий вид линейного интегрального уравнения Вольтерры 1-го рода

$$A\varphi \equiv \int_a^x K(x, s) \varphi(s) ds = f(x), \quad x \in [a, b]. \quad (108)$$

Важными и широко распространенными разновидностями интегрального уравнения Вольтерры 1-го рода являются уравнения

$$\int_{-\infty}^x K(x-s) \varphi(s) ds = f(x), \quad -\infty < x < \infty, \quad (109)$$

$$\int_0^x K(x-s) \varphi(s) ds = f(x), \quad x \in [0, b], \quad (110)$$

$$\int_0^x K(x, s) F(s, \varphi(s)) ds = f(x), \quad x \in [a, b], \quad (111)$$

$$\int_{-\infty}^{+\infty} K(x-s) F(s, \varphi(s)) ds = f(x), \quad -\infty < x < \infty \quad (112)$$

и другие.

Отличительной особенностью решений этих уравнений является то, что с одной стороны, уравнения (108) - (112) - частные случаи уравнений Фредгольма 1-го рода, то есть их можно решать методами регуляризации. С другой стороны, при определенных условиях, налагаемых на функции  $K(x, s)$  и  $f(x)$ , уравнения имеют

решения, основанные на дискретизации исходного уравнения. В первом случае решение методами регуляризации значительно сложнее соответствующих решений уравнений Фредгольма за счет того, что матрица алгебраических уравнений уже является полной.

Применение прямых методов дискретизации часто вызывает неустойчивость приближенных решений (некорректность задачи).

Разработаны методы решения, объединяющие достоинства указанных двух подходов к решению [1], [2].

Если непрерывное решение уравнения (108) существует, то функция  $f(x)$  должна иметь непрерывные производные до  $n$ -го порядка включительно, причем

$$f(0) = f'(0) = \dots = f^{(n-1)}(0) = 0, \quad (113)$$

при этом решение имеет вид

$$\varphi(x) = f^{(n)}(x). \quad (114)$$

Предполагается, что ядро  $K(x, s)$  - непрерывная функция.

Рассмотрим типовые методы решения уравнения (108).

• **Приведение к уравнению 2-го рода.**  
Это преобразование может быть проведено различными способами.

1) Посредством дифференцирования (однократного, или  $p$ -кратного), при условии дифференцируемости функций  $K(x, s)$  и  $f(x)$  нужное число раз.

**Пример 15.**

Решить уравнение

$$\int_0^x \sin(x-s) \varphi(s) ds = 1 - \cos x,$$

сведем его к уравнению Вольтерры 2-го рода.

**Решение.**

Дифференцируем уравнения по переменной  $x$ .

$$\sin(x-x) \varphi(x) + \int_0^x \cos(x-s) \varphi(s) ds = \sin x.$$

Дифференцирование уравнения 1-го рода

$\int_0^x \cos(x-s) \varphi(s) ds = \sin x$  позволяет получить эквивалентное исход-

ному уравнению 2-го рода  $\varphi(x) - \int_0^x \sin(x-s) \varphi(s) ds = \cos x$ .

Решение уравнения -  $\varphi(x) \equiv 1$ .

2) Интегрирование по частям.

Пусть  $\int_a^x \varphi(s) ds = Y(x)$ ,  $x \in [a, b]$ ,  $u = K(x, s)$ ,  $dv = \varphi(s) ds$ .

Тогда из (108) получим уравнение Вольтерры 2-го рода:

$$K(x, x) Y(x) - \int_a^x K'_s(x, s) Y(s) ds = f(x), \quad x \in [a, b]. \quad K(x, x) \neq 0,$$

$$Y(x) - \int_0^x \frac{K'_s(x, s)}{K(x, x)} Y(s) ds = \frac{f(x)}{K(x, x)}. \quad (115)$$

Решение уравнения

$$\varphi(s) = \frac{dy(s)}{ds}, \quad s \in [a, b].$$

Интегрирование по частям

$u = \varphi(s)$ ,  $dv = K(x, s) ds$ , тогда

$$v(x, x) \varphi(x) - \int_a^x v(x, s) \frac{d\varphi(s)}{ds} ds = f(x), \quad x \in [a, b],$$

$$v(x, s) = \int_a^s K(x, t) dt. \quad (116)$$

$$\text{Если } v(x, x) \neq 0, \text{ то } \varphi(x) - \int_a^x \frac{v(x, s)}{v(x, x)} \frac{d\varphi(s)}{ds} ds = \frac{f(x)}{v(x, x)}, \quad x \in [a, b]. \quad (117)$$

**Пример 16.**

Привести уравнение  $\int_0^x e^{x-s} \varphi(s) ds = \sin x$

к уравнению Вольтерры 2-го рода.

**Решение.**

Интегрируем по частям  $u = e^{x-s}$ ,  $dv = \varphi(s) ds$ .

$$K(x, s) = e^{x-s}, \quad K(x, x) \equiv 1, \quad v(x, s) = \int_0^s e^{x-t} dt = e^x - e^{x-s}.$$

На основании (115)  $Y(x) + \int_0^x e^{x-s} Y(s) ds = \sin x$ . Это уравнение

Вольтерры 2-го рода имеет решение  $\varphi(x) = \cos(x) - \sin(x)$ .

• Уравнение (108) можно решать методом квадратур.  
Если отрезок  $[a, b]$  разбит на  $(n-1)$  частей, выбраны узлы дискретизации  $x = x_i$ ,  $i = 1, n$ ,  $x_1 = a$ ,  $x_n = b$ , с учетом условий

$$\int_a^{x_i} K(x_i, s) \varphi(s) ds = f(x_i) \quad (118)$$

получается система уравнений  $\sum_{j=1}^i A_j K_{ij} \varphi_j = f_i, \quad i = \overline{1, n}$ . (119)

Алгоритм на основе формулы трапеции описывается [12] формулами:

$$\left\{ \begin{array}{l} \varphi_1 = \frac{f'(a)}{K_{11}}, \quad \varphi_2 = \frac{f_2 - \frac{h_1}{2} K_{21} \varphi_1}{\frac{h_1}{2} K_{22}}, \\ \dots \\ \varphi_i = \frac{f_i - \frac{h_i}{2} K_{i1} \varphi_1 - \sum_{j=2}^{i-1} \left( \frac{x_{j+1} - x_{j-1}}{2} \right) K_{ij} \varphi_j}{\frac{h_i}{2} K_{ii}}, \quad i = \overline{3, n}. \end{array} \right. \quad (120)$$

$$K_{ii} \neq 0. \quad (121)$$

$$f'(a) = -\frac{x_2 - x_1 + x_3 - x_1}{(x_2 - x_1) \cdot (x_3 - x_1)} f_1 + \frac{x_3 - x_1}{(x_2 - x_1) \cdot (x_3 - x_2)} f_2 - \frac{x_2 - x_1}{(x_3 - x_1) \cdot (x_3 - x_2)} f_3. \quad (122)$$

Погрешность вычисления  $\Delta \varphi_i$  определяется по формулам

$$\Delta \varphi_1 = \Delta \varphi_2 = 0, \quad \Delta \varphi_i = -\frac{R_i + \sum_{j=2}^{i-1} \left( \frac{x_{j+1} - x_{j-1}}{2} \right) K_{ij} \varphi_j}{\frac{h_i}{2} K_{ii}}, \quad (123)$$

$R_i$  вычисляется по формуле (97).

При постоянном шаге  $h_i = h = \text{const}$  имеем

$$\left\{ \begin{array}{l} \varphi_1 = \frac{f'(a)}{K_{11}}, \\ \varphi_i = \frac{2}{K_{ii}} \left( \frac{f_i}{h} - \sum_{j=1}^{i-1} a_j K_{ij} \varphi_j \right), \quad i = \overline{2, n}, \\ x_i = a + (i-1)h, \quad a_j = \begin{cases} 0.5, & j=1, \\ 1, & j>1, \end{cases} \end{array} \right. \quad (124)$$

$$f'(a) = \frac{-3f_1 + 4f_2 - f_3}{2h}.$$

Если ядро  $K(x, s) = \sum_{l=1}^m \alpha_l(x) \beta_l(s), \quad l = \overline{1, m}$

разделяющееся, то для решения уравнения получаем рекуррентные выражения

$$\varphi(0) = \frac{f'(0)}{\sum_{l=1}^m \alpha_l(0) \beta_l(0)}, \quad (125)$$

$$\varphi(x_i) = \frac{1}{\sum_{l=1}^m A_l \alpha_l(x_i) \beta_l(x_i)} \left[ f(x_i) - \sum_{l=1}^m \alpha_l(x_i) \sum_{j=1}^{i-1} A_j \beta_l(x_j) \varphi(x_j) \right].$$

• Метод регуляризации Тихонова для уравнения Вольтерры 1-го рода.

Для обеспечения устойчивости решения уравнения вводится условие минимума сглаживающего функционала.

$$\int_a^b [A\varphi - f(x)]^2 dx + \alpha \Omega(\varphi) = \min_{\varphi}, \quad (126)$$

$$\Omega(\varphi) = \int_a^b [\varphi^2(s) + q(\varphi'(s))^2] ds, \quad (127)$$

$\alpha > 0$  - параметр регуляризации (определяется способом невязки);  $q = 0$ , если порядок регуляризации нулевой;  $q > 0$ , если порядок регуляризации первый.

При выполнении условий (126), (127) уравнение сводится к уравнению второго рода.

Для решения практически всех видов уравнений, рассмотренных выше, предназначен пакет программ [12].

В частности предложены процедуры (модули), позволяющие решать уравнения Фредгольма 2-го рода методом квадратур, уравнения Фредгольма 1-го рода методом Тихонова с выбором параметра регуляризации методом обобщенной невязки и при известном точном значении параметра.

Также имеются модули для решения уравнений Вольтерры 1-го и 2-го рода методом квадратур и ряд других.

Ядра уравнений предполагается задавать в двух вариантах: в виде процедур-функций или в виде массивов.

Возможно обращение к ядрам и в смешанном варианте. Допускается табличное задание ядер уравнений. Все программы пакета снабжены достаточным комментарием, описываются структуры и вид параметров процедур. Приведены полные тексты программ, включая программы модулей низших уровней, тесты пакета, результаты решения тестовых задач.



## СПИСОК ЛИТЕРАТУРЫ

1. Абрамов Л.Н. Капустин В.Ф. Математическое программирование. - Л.:ЛГУ, 1981. - 326 с.
2. Акулич И.Л. Математическое программирование в примерах и задачах. - М.:Высшая школа, 1993. - 336 с.
3. Ахмед Н., Рао К.Р. Ортогональные преобразования при обработке цифровых сигналов. - М.:Связь, 1980.-248 с.
4. Баңди Б. Методы оптимизации. - М.:Радио и связь, 1988. - 128 с.
5. Баңди Б. Основы линейного программирования. М.:Радио и связь, 1989. - 176 с.
6. Бахвалов Н.С. Численные методы, т. 1. - М.:Наука, 1973. - 630 с.
7. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы. - М.:Наука, 1987. - 600 с.
8. Березин И.С., Жидков Н.П. Методы вычислений, т. 1, 2. - М.:Физматгиздат, 1982. - 640 с.
9. Бут Э.Д. Численные методы. - М.:Физматгиздат, 1959. - 239 с.
10. Васильев Ф.П. Лекции по методам решения экстремальных задач. - М.:МГУ, 1974. - 374 с.
11. Васильев Ф.П. Численные методы решения экстремальных задач. - М.:Наука, 1981. - 400 с.
12. Верлань А.Ф., Сизиков В.С. Интегральные уравнения: методы, алгоритмы, программы. - К.:Наукова думка, 1986. - 291 с.
13. Волков А.Е. Численные методы. - М.:Наука, 1982. - 248 с.
14. Воробьев Н.Н. Числа Фибоначчи. - М.:Наука, 1987. - 248 с.
15. Воробьева Г.Н., Данилова А.Н. Практикум по вычислительной математике. - М.:Высшая школа, 1990. - 308 с.
16. Воробьев А.Ф., Даниелян Т.Я., Данилина Н.И. Вычислительная математика. - М.:Статистика, 1966. - 164 с.
17. Воскобойников Ю.Е. О сегментном методе построения численного решения интегрального уравнения 1 рода с разностным ядром. Алгоритмы обработки теплофизического эксперимента. - Новосибирск: Ин-т теплофизики Сиб. отделение АН СССР, 1975.
18. Вышеславцев А.П. Сборник задач по теории оптимизации. Методы поиска. - М.:МЭИ, 1983. - 57 с.
19. Габасов Р., Кириллова Ф.М. Методы оптимизации. - Минск, БГУ, 1975. - 272 с.
20. Горинштейн А.М. Численное решение задач радиотехники и техники связи на ЭЦВМ. - М.:Связь, 1972. - 200
21. Горинштейн А.М. Практика решения инженерных задач на ЭВМ. - М.:Радио и связь, 1984. - 232 с.
22. Гутер Р.С., Овчинский Б.В. Элементы численного анализа и математической обработки результатов опыта. - М.:Наука, 1970. - 432 с.
23. Демидович Б.П., Марон И.А. Основы вычислительной математики. - М.:Наука, 1966. - 665 с.
24. Демидович Б.П., Марон И.А., Шувалова Э.З. Численные методы анализа. - М.:Наука, 1968. - 400 с.

25. Дымарский Я.С., Лозинский Н.Н., Макушкин А.Т., Розенбарг В.Я., Эрглис В.Р. Справочник программиста, том I. - П., 1969.

26. Дьяконов В.П. Справочник по алгоритмам и программам на языке БЕЙСИК для персональных ЭВМ. - М.:Наука, 1987. - 240 с.

27. Де Борк. Практическое руководство по сплайнам. - М.:Радио и связь, 1985 - 304 с.

28. Зельдович Я.Б., Мышкис А.Д. Элементы прикладной математики. - М.:Наука, 1972. - 592 с.

29. Ищук В.А. О двусторонних методах вида Рунге-Кутты решения нелинейного интегрального уравнения типа Вольтерра. - Сборник Института кибернетики, 1974.

30. Карманов В.Г. Математическое программирование. - М.:Наука, 1975. - 272 с.

31. Канторович Л.В., Горстков А.Б. Математическое оптимальное планирование в экономике. - М.:Знание, 1968. - 96 с.

32. Калиткин Н.Н. Численные методы. - М.:Наука, 1978. - 500 с.

33. Кальмушевски И.И. Методы интегральных уравнений в задачах связи. - Одесса, 1986. - 46 с.

34. Каплан И.А. Практические занятия по высшей математике, - Харьков, 1972. - 972 с.

35. Карандаев И.С. Начала линейного, нелинейного и динамического программирования. - М.:Знание, 1967. - 158 с.

36. Карпелевич Ф.И., Садовский Л.Е. Элементы линейной алгебры и линейного программирования. - М.:Наука, 1965. - 275 с.

37. Кожевников Н.И., Краснощекова Т.И., Шишкин Н.Е. Ряды и интеграл Фурье. Теория Поля. Аналитические и специальные функции. Преобразование Папласа. - М.:Наука, 1964. - 184 с.

38. Копченкова Н.В., Марон И.А. Вычислительная математика в примерах и задачах. - М.:Наука, 1972. - 368 с.

39. Краснов М.Л., Киселев А.И., Макаренко Г.И. Интегральные уравнения. - М.:Наука, 1976. - 192 с.

40. Краснов М.Л. Интегральные уравнения. - М.:Наука, 1975. - 298 с.

41. Крылов В.И., Бобков В.В., Монастырный П.И. Вычислительные методы. В 2-х т. - М.:Наука, 1977. - 339 с.

42. Крылов В.И., Бобков В.В., Монастырный П.И. Начала теории вычислительных методов. Уравнения в частных производных. - Минск, Наука и техника, 1986. - 308 с.

43. Крылов В.И., Бобков В.В., Монастырный П.И. Вычислительные методы высшей математики, том I. - М.:Наука, 1976. - 303 с.

44. Крылов В.И., Шульгина Л.Т. Справочная книга по численному интегрированию. - М.:Наука, 1966. - 370 с.

45. Кузнецов Ю. А., Кузубов В. И., Вольщенко А. Б. Математическое программирование. - М.:Высшая школа, 1980 - 300 с.

46. Левшиц В.М., Литвин Б.Ф. Приближенные вычисления и программирование на ЭВМ "Наири-2". - Л.:Машиностроение, 1977. - 240 с.

47. Марчук Г.И. Методы вычислительной математики. М.:Наука, 1989. - 455 с.

48. Моисеев Н.И., Иванилов Ю.П., Столярова Е.М. Методы оптимизации. - М.:Наука, 1978. - 352 с.

49. Мудров А.Е. Численные методы для ПЭВМ на языках БЕЙСИК, ФОРТРАН и ПАСКАЛЬ. - Томск, Раско, 1991. - 272 с.

50. Пашкев С.Д., Минязов Р.И., Могилевский В.Д. Машинные методы оптимизации в технике связи. - М.: Связь, 1976. - 272 с.
51. Петровский И.Г. Лекции по теории обыкновенных дифференциальных уравнений. - М.: Наука, 1970. - 280 с.
52. Писенничный Б.Н., Данилин Ю.Н. Численные методы в экстремальных задачах. - М.: Наука, 1974. - 320 с.
53. Романовский П.И. Ряды Фурье. Теория Поля. Аналитические и специальные функции. Преобразование Лапласа. - М.: Наука, 1973. - 336 с.
54. Соболев И.М. Численные методы Монте-Карло. - М.: Наука, 1979. - 312 с.
55. Стечкин С.Б., Субботин Ю.Н. Сплайны в вычислительной математике. - М.: Наука, 1976. - 248 с.
56. Сухарев А.Г., Тимохов А.В., Фёдоров В. В. Курс методов оптимизации. - М.: Наука, - 1986 - 324 с.
57. Турецкий А.Х. Теория интерполирования в задачах. - Минск. - М.: Высшая школа, 1968. - 518 с.
58. Турчак Л.И. Основы численных методов. - М.: Мир, 1987. - 350 с.
59. Уайлд Д. Методы поиска экстремума: Пер. с англ./Ред. А.А. Фельдбаум. - М.: Наука, 1967. - 267 с.
60. Харди Г.Х., Рогозинский В.В. Ряды Фурье. - М.: Физматиздат, 1959. - 156 с.
61. Шуп Т.Е. Прикладные численные методы в физике и технике. - М.: Высшая школа, 1990. - 254 с.
62. Шуп Т.Е. Решение инженерных задач на ЭВМ. - М.: Мир, 1990. - 235 с.
63. Юдин Д.Б., Гольштейн Е.Г. Линейное программирование. - М.: Наука, 1990. - 425 с.
64. Алгоритмы и программы восстановления зависимостей. Под редакцией Вапника В.Н. - М.: Наука, 1984. - 815 с.
65. Инженерные расчеты на ЭВМ. Под редакцией Троицкого В.А. - Л.: Машиностроение, 1979. - 288 с.
66. Сборник задач по математике, ч. 3,4. Под редакцией Ефимова А.В. - М.: Наука, 1984. - 303 с.
67. Справочник по специальным функциям. Пер. с англ. - Под редакцией Абрамова М. и Стиган И., - М.: Наука, 1979. - 832 с.
68. Дж.-Н. Ланс Численные методы для быстродействующих вычислительных машин. Пер. с англ. - М.: ИЛ, 1962. - 208 с.
69. Дж. Форсайт, М. Малькольм, К. Моулер Машинные методы математических вычислений. - М.: Мир, 1980. - 279 с.
70. Д. Мак-Кракен, У. Дори Численные методы и программирование на ФОРТРАНЕ. - М.: Мир., 1977. - 584 с.
71. Харкевич А.А. Спектры и анализ. - М.: Физматиздат, 1957. - 236 с.
72. Численные методы условной оптимизации. Под редакцией Ф. Гилл, У. Мюррей. - М.: Мир., 1977. - 293 с.
73. Чуа А.О., Лин Пен-Мин Машинный анализ электронных схем. Алгоритмы и вычислительные методы. Пер. с англ. - М.: Энергия, 1980. - 640 с.
74. Э. Полак Численные методы и оптимизация. - М.: Мир, 1972. - 374 с.
75. Шура-Бура М.Р. Аппроксимация функций многих переменных. В.М., вып. 2. - М., 1967.

## СОДЕРЖАНИЕ

Предисловие.....	3
1. СИСТЕМЫ ЛИНЕЙНЫХ УРАВНЕНИЙ.....	4
1.1. Метод Гаусса.....	4
1.2. Метод Гаусса – Жордана.....	10
1.3. Метод Холесского решения систем линейных уравнений.....	14
1.4. Метод прогонки.....	15
1.5. Методы итерации.....	24
1.6. Исследование и решение систем линейных уравнений методом Жордана.....	27
2. НЕЛИНЕЙНЫЕ УРАВНЕНИЯ.....	31
2.1. Метод половинного деления.....	32
2.2. Метод итерации.....	35
2.3. Усовершенствованный метод итерации.....	37
2.4. Метод Ньютона – Рафсона.....	39
2.5. Методы решения алгебраических уравнений.....	43
2.6. Определение комплексных корней алгебраических уравнений.....	45
2.7. Метод парабол.....	55
3. СИСТЕМЫ НЕЛИНЕЙНЫХ УРАВНЕНИЙ.....	63
3.1. Метод Ньютона.....	63
3.2. Метод итерации.....	66
3.2.1. Приведение системы $\vec{f}(x) = 0$ к виду, удобному для итерации.....	73
4. АППРОКСИМАЦИЯ ФУНКЦИЙ.....	79
4.1. Постановка задачи интерполяции.....	79
4.2. Интерполяционный полином Лагранжа.....	80
4.3. Интерполяционные формулы Ньютона.....	86
4.4. Обратная интерполяция.....	90
4.5. Интерполяционная формула Ньютона для произвольных значений аргумента.....	91
4.6. Сплайн – интерполяция.....	93
4.7. Метод наименьших квадратов.....	99
4.8. Ортогональные функции. Ряд Фурье.....	109
4.9. Ряд Фурье для периодической функции.....	111
4.10. Ряд Фурье для произвольного периода.....	113
4.11. Интеграл Фурье.....	114
4.12. Комплексная форма записи ряда Фурье.....	115
4.13. Интеграл Фурье в комплексной форме.....	116
4.14. Многомерная интерполяция. Интерполяционные формулы Эрмита.....	122

5. ЧИСЛЕННЫЕ МЕТОДЫ ИНТЕГРИРОВАНИЯ.....	123
5.1. Методы Ньютона - Котеса.....	124
5.1.1. Формулы прямоугольников.....	124
5.1.2. Формула трапеций.....	125
5.1.3. Формула Симпсона.....	126
5.2. Сплайн-квадратура.....	134
5.3. Квадратуры Гаусса.....	137
5.4. Интегрирование быстро осциллирующих функций.....	146
5.5. Интегрирование периодических функций.....	149
5.6. Несобственные интегралы.....	152
5.7. Кратные интегралы.....	156
5.8. Кубатурная формула Коробова.....	157
5.9. Кубатурная формула Симпсона.....	158
5.10. Метод Монте-Карло вычисления двойных интегралов.....	161
5.11. Вычисление неопределенного интеграла.....	162
6. ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ.....	168
6.1. Методы Рунге-Кутты.....	168
6.1.1. Геометрическая иллюстрация методов Рунге-Кутты.....	171
6.2. Метод Рунге - Кутта - Мерсона.....	176
6.3. Метод А. Н. Крылова последовательных сближений.....	177
6.4. Метод Адамса.....	179
6.5. Метод Гира.....	182
6.6. Системы обыкновенных дифференциальных уравнений.....	183
6.7. Уравнения высших порядков.....	190
6.8. Дифференциальные уравнения второго порядка.....	190
6.9. Краевые (граничные) задачи.....	191
6.10. Численные методы решения краевой (граничной) задачи.....	193
6.10.1. Метод стрельбы.....	193
6.10.2. Метод конечных разностей.....	195
7. ОПТИМИЗАЦИЯ.....	200
7.1. Метод прямого перебора.....	202
7.2. Методы последовательного перебора.....	204
7.2.1. Метод дихотомии.....	204
7.2.2. Метод золотого сечения.....	208
7.2.3. Оптимизация функций методом квадратичной интерполяции.....	212
7.2.4. Метод Фибоначчи.....	215
7.3. Метод ломаных.....	219
7.4. Метод касательных.....	224
7.5. Метод Ньютона.....	227

8. ОПТИМИЗАЦИЯ ФУНКЦИИ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ.....	228
8.1. Основные понятия и определения.....	228
8.2. Классический метод решения задачи оптимизации.....	232
8.3. Поисковые методы оптимизации.....	233
8.4. Метод конфигураций.....	234
8.5. Градиентные методы оптимизации.....	234
8.6. Метод градиентного поиска с дроблением шага (GR).....	238
8.7. Методы наискорейшего спуска (NS).....	242
8.8. Метод сопряженных направлений.....	247
8.9. Метод Флетчера - Ривза.....	250
8.10. Метод Ньютона.....	253
8.11. Метод Давидона - Флетчера - Пауэлла (ДФ).....	257
9. ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ.....	260
9.1. Постановка задачи.....	260
9.2. Геометрическая иллюстрация задачи линейного программирования с ограничениями - неравенствами.....	262
9.3. Симплексный метод решения ЗЛП.....	266
9.4. Определение начального базисного решения.....	270
9.5. Алгоритм получения начального допустимого базиса.....	272
9.6. Двойственность в линейном программировании.....	288
9.6.1. Геометрическая иллюстрация двойственных задач.....	289
9.6.2. Симплексный метод решения двойственной задачи.....	291
9.6.3. Двойственный симплексный метод.....	293
10. НЕЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ.....	301
10.1. Метод множителей Лагранжа.....	302
10.1.1. Геометрическая интерпретация метода Лагранжа.....	303
10.2. Выпуклое программирование.....	307
10.3. Квадратичное программирование.....	310
10.4. Численные методы решения задачи нелинейного программирования.....	317
10.4.1. Прямые методы.....	317
10.4.2. Методы спуска.....	319
10.4.3. Методы преобразования.....	323
10.4.3.1. Барьерные функции.....	323
10.4.3.2. Штрафные функции.....	328
10.4.4. Метод Фиакко-Макормика решения задачи оптимизации с ограничениями.....	331
11. СОБСТВЕННЫЕ ЗНАЧЕНИЯ И СОБСТВЕННЫЕ ВЕКТОРЫ.....	333
11.1. Метод итерации.....	334
11.2. Метод Леверрье-Фаддеева.....	344
11.1. Метод Данилевского.....	354

11.2. Метод Якоби.....	357
11.3. QR-метод.....	364
12. ИНТЕГРАЛЬНЫЕ УРАВНЕНИЯ.....	367
12.1. Линейные интегральные уравнения Фредгольма 2-го рода.....	368
12.1.1. Метод квадратур для уравнений Фредгольма 2-го рода.....	369
12.1.2. Решение интегральных уравнений с помощью резольвенты.....	371
12.1.3. Метод последовательных приближений.....	373
12.1.4. Метод вырожденных ядер.....	375
12.1.5. Способы аппроксимации ядер вырожденными.....	377
12.1.6. Метод моментов.....	381
12.1.7. Метод коллокаций.....	382
12.2. Линейные интегральные уравнения Фредгольма 1-го рода.....	383
12.3. Интегральные уравнения с разностными ядрами.....	385
12.4. Методы решения уравнений Вольтерры 2-го рода.....	387
12.5. Методы решения уравнений Вольтерры 1-го рода.....	397
Список литературы.....	402
Содержание.....	405