

Министерство образования и науки Российской Федерации

Федеральное государственное бюджетное образовательное учреждение  
высшего профессионального образования  
«Национальный исследовательский университет „МЭИ“»

Институт электроэнергетики (ИЭЭ)

---

Направление подготовки: 140400 Электроэнергетика и электротехника

Специальность: 140202 Нетрадиционные и возобновляемые источники энергии

Квалификация (степень) выпускника: бакалавр

Форма обучения: очная

#### 4. Конспект лекций

«Математические задачи энергетики возобновляемых источников»

Литература

- 1 **Приближение функций**
  - 1.1 Функции одной переменной
    - 1.1.1 **Понятие функции одной переменной**
    - 1.1.2 **Способы задания функции одной переменной**
  - 1.2 Интерполирование функций
    - 1.2.1 **Постановка задачи интерполяции**
    - 1.2.2 **Методы интерполяции**
  - 1.3 Аппроксимация функций
    - 1.3.1 **Постановка задачи аппроксимации функции одной переменной**
    - 1.3.2 **Порядок определения вида аппроксимирующей функции**
    - 1.3.3 **Постановка задачи аппроксимации функции одной переменной**
- 2 **Численное интегрирование**
  - 2.1 Методы численного интегрирования
    - 2.1.1 **Метод прямоугольников**
    - 2.1.2 **Метод трапеций**
    - 2.1.3 **Метод парабол (метод Симпсона)**
- 3 **Введение в теорию систем и принятия решений**
  - 3.1 Теория систем
    - 3.1.1 **Основные понятия**
    - 3.1.2 **Понятие системы на примере электрической и структурной схем**
    - 3.1.3 **Окружающие системы, внешняя среда**
    - 3.1.4 **Понятие «наилучшего» решения**
    - 3.1.5 **Процесс принятия решений при проектировании и эксплуатации систем**
- 4 **Элементы теории вероятностей и математической статистики**
  - 4.1 Основные понятия
    - 4.1.1 **Вероятность, частота событий**
    - 4.1.2 **Случайные величины**
  - 4.2 Закон распределения
    - 4.2.1 **Ряд распределения, многоугольник распределения**
    - 4.2.2 **Функция распределения**
    - 4.2.3 **Свойства функции распределения**
    - 4.2.4 **Вероятность попадания случайной величины в заданный интервал**
    - 4.2.5 **Плотность распределения**
    - 4.2.6 **Основные свойства плотности распределения**
  - 4.3 Числовые характеристики случайных величин

- 4.3.1 Математическое ожидание
- 4.3.2 Дисперсия и среднеквадратическое отклонение
- 4.4 Частные случаи законов распределения
- 4.4.1 Закон равномерной плотности
- 4.4.2 [Ошибка! Источник ссылки не найден.](#)
- 4.5 [Ошибка! Источник ссылки не найден.](#)
- 4.5.1 Простой статистический ряд. Статистическая функция распределения
- 4.5.2 Статистический ряд. Гистограмма
- 4.5.3 Числовые характеристики статистического распределения
- 4.5.4 Выравнивание статистических рядов
- 4.5.5 Критерии согласия. Критерий Пирсона
- 5 Методы оптимизации
- 6 Линейное программирование
  - 6.1 Транспортная задача
  - 6.2 Геометрическая интерпретация задачи линейного программирования
  - 6.3 Симплекс-метод
- 7 Нелинейное программирование
  - 7.1 Численные методы оптимизации функций одной переменной
  - 7.1.1 Методы первого порядка
  - 7.1.2 Методы второго порядка
  - 7.2 Численные методы оптимизации
  - 7.2.1 Методы нулевого порядка
  - 7.2.2 Методы первого порядка. Градиентные методы
  - 7.2.3 Методы второго порядка

## Литература

- Кузин Л.Т., «Основы кибернетики, том 1: Математические основы кибернетики» Учебное пособие для студентов ВУЗов. М.: Энергия, 1977
- Саркисян С.А. и другие, «Теория прогнозирования и принятия решения» Учебное пособие для студентов ВУЗов. М.: Высшая школа, 1977
- Веников В.А. и другие, «Оптимизация режимов электростанций и энергосистем» Учебник для ВУЗов. М.: Энергоиздат, 1981
- Чистяков В.П., «Курс теории вероятностей» М.: Наука, 1982
- Коршунов Ю.М., «Математические основы кибернетики» М.: Энергия, 1972
- Вентцель Е.С., «Теория вероятностей» М.: Наука, 1983
- Малинин Н.К., «Оптимизация режимов ГЭС и ГАЭС методом динамического программирования» М.: издательство МЭИ, 1991

- Васильев Ф.П., «Численные методы решения экстремальных задач»  
Учебное пособие для студентов ВУЗов. М.: Наука, 1988

## Раздел 1

# Приближение функций

Изучение разнообразных явлений в окружающем нас мире приводит к понятию функции. Так, например, ясно, что каждому моменту времени в данной местности соответствует определенная температура воздуха; атмосферное давление меняется в зависимости от высоты местности; продуктивность водоема зависит от продолжительности солнечного освещения; морские приливы и отливы периодически повторяются в зависимости от фазы Луны, и так далее. Во всех этих случаях значению одной величины (время, высота над уровнем моря, продолжительность солнечного освещения, положение Луны относительно Земли и Солнца) ставится в соответствие определенное значений другой величины по определенному закону. Используя математический аппарат, можно исследовать природные закономерности, проводить прогнозирование событий, анализировать прошедшие, и так далее. Для этого необходимо владеть приемами перевода языка природы на язык математики. И одной из первых задач исследователя при обработке экспериментальных данных является задача нахождения имеющейся функциональной зависимости между измеренными величинами. Об этом и пойдет речь в настоящей главе.

### 1.1. Функции одной переменной

#### **1.1.1. Понятие функции одной переменной**

Рассмотрим два числовых множества  $X$  и  $Y$ . Правило  $f$ , по которому каждому числу  $x$ , принадлежащему множеству  $X$ , ставится в соответствие единственное число  $y$ , принадлежащее множеству  $Y$ , называется **числовой функцией**, заданной на множестве  $X$  и принимающей значения во множестве  $Y$ .

$$(y = f(x)) \equiv \left( \forall x \in X, \exists! y \in Y: \{x \xrightarrow{f} y\} \right) \quad (1.1)$$

*выражение в левой части (1.1), — « $y$  функционально зависит от  $x$ », — означает, что для*

каждого  $x$ , входящего в множество  $X$ , найдется единственный  $y$ , входящий в множество  $Y$ , такой, что  $y$  соответствует этому  $x$  по определенному правилу  $f$

Таким образом, задать функцию — значит задать три объекта:

- множество  $X$  (область определения функции);
- множество  $Y$  (область значений функции);
- правило соответствия  $f$  (сама функция).

Процессы, происходящие в системе, необходимо научиться описывать функциональной зависимостью, тогда, зная все значения  $x$  и правило  $f$ , можно найти все значения  $y$ .

### 1.1.2. Способы задания функции одной переменной

#### 1. Аналитический способ.

Функция  $f$  задается в виде формулы  $y = f(x)$ , например,  $y = 3\cos x + 2x^2$ .

Во многих случаях формула бывает неизвестна, и тогда для выражения функциональной зависимости используются другие способы.

#### 2. Графический способ.

В пунктах учета электроэнергии (выработанной или потребленной) можно наблюдать работу приборов-самописцев, регистрирующих величину активной или реактивной мощности в любой момент времени суток. По полученному графику можно определить значения указанных величин в любой момент времени. Графиком функции  $y = f(x)$  называется множество всех точек плоскости с координатами  $(x, f(x))$ . График содержит всю информацию о функции. Имея перед собой график, мы, как бы, «видим функцию».

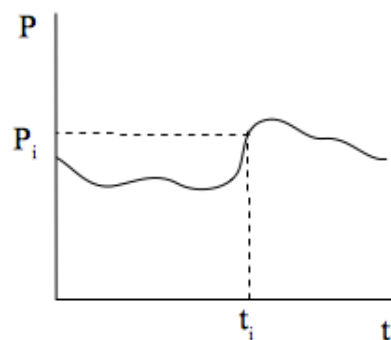


Рис. 1

#### 3. Табличный способ.

Этот способ является наиболее простым. В одной строке таблицы записываются все значения аргумента (числа), а в другой — значения  $f(x)$ , соответствующие каждому  $x$ . Например, зависимость температуры воздуха  $T$  от времени суток  $t$  в определенный день можно представить такой таблицей:

Таблица 1.1

$t$ , час	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
$T$ , °C	8	7	6	5	5	5	7	9	12	17	19	21	22	22	21	22	20	20	20	20	18	15	12	9

Табличные зависимости получаются в результате регистрации результатов опытов, лабораторных экспериментов, периодических замеров каких-либо параметров. Однако, по таблице можно найти значения функции только для тех аргументов, которые в таблицу внесены.

Часто возникают задачи, требующие нахождения значения функции для значения аргумента, не входящего в таблицу. Тогда важно знать аналитическое задание функции.

Для этого функцию, заданную таблично или графически (график всегда можно представить в табличном виде при необходимости), заменяют на некотором отрезке  $[a; b]$  другой, более простой, функцией, близкой к данной и имеющей аналитическое выражение. Существует два основных приема такой замены — интерполирование и аппроксимация

функции-таблицы.

## 1.2. Интерполирование функций

### 1.2.1. Постановка задачи интерполяции

Пусть известные значения некоторой функции  $F$  образуют следующую таблицу:

Таблица 1.2

$x_i$	$x_0$	$x_1$	...	$x_n$
$F(x_i)$	$y_0$	$y_1$	...	$y_n$

При этом требуется получить значение функции  $F$  для такого значения аргумента  $x$ , которое входит в отрезок  $[x_0; x_n]$ , но не совпадает ни с одним из значений  $x_i$  ( $i=0, 1, \dots, n$ ). Геометрически это означает, что нужно найти кривую  $y = F(x_i)$  некоторого определенного

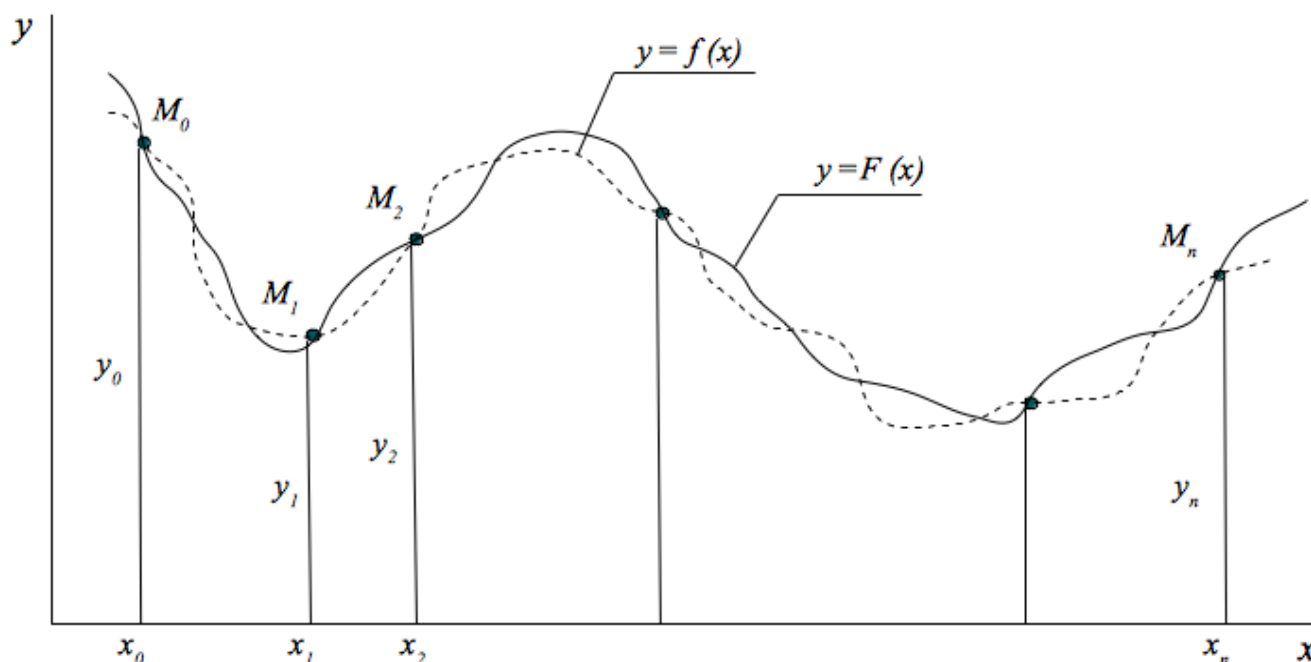


Рис. 2

типа, проходящую через заданную систему точек  $M_i(x_i, y_i)$  ( $i=0, 1, \dots, n$ ) (рис. 2).

Задача интерполирования может иметь в общей постановке бесчисленное множество решений или совсем их не иметь, но задача становится однозначной, если вместо произвольной функции  $F(x)$  искать некоторую функцию конкретного вида, удовлетворяющую требованию строгого совпадения значений  $f(x)$  и  $F(x_i)$  в точках  $x_i$  ( $i=0, 1, \dots, n$ ), то есть  $f(x_0) = F(x_0) = y_0$ ;  $f(x_1) = F(x_1) = y_1$ ; ...;  $f(x_n) = F(x_n) = y_n$ .

В этом случае нахождение приближенной функции называют **интерполяцией** или **интерполированием**, а точки  $x_0, x_1, \dots, x_n$  — **узлами интерполяции**.

Наиболее удобной функцией приближения является алгебраический многочлен степени  $n$ :

$$P_n(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \dots + a_{n-1}x + a_n \quad (1.2)$$

## 1.2.2. Методы интерполяции

Рассмотрим линейную и квадратичную интерполяции.

### 1. Порядок выполнения линейной интерполяции.

Найдем в исходной таблице два соседних значения аргумента, между которыми лежит заданное значение  $x$ , и обозначим их  $x_k$  и  $x_{k+1}$  (так, что  $x_k < x < x_{k+1}$ ).

Соответствующие им значения функции обозначим как  $y_k = f(x_k)$ ,  $y_{k+1} = f(x_{k+1})$ . Будем считать, что в промежутке  $(x_k, x_{k+1})$  данную функцию с достаточной степенью точности можно заменить линейной функцией, то есть дугу графика функции можно заменить стягивающей ее хордой (рис. 3). Такая замена называется линейной интерполяцией.

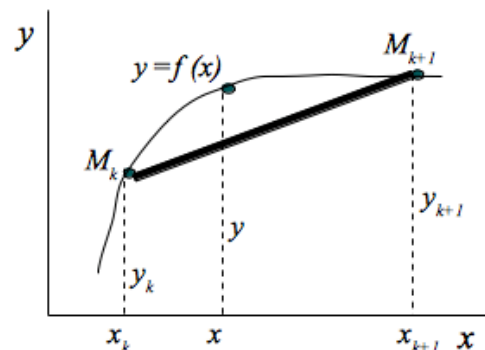


Рис. 3

Уравнение прямой, проходящей через точки  $M_k(x_k, y_k)$  и  $M_{k+1}(x_{k+1}, y_{k+1})$ , имеет следующий вид:

$$\frac{y-y_k}{y_{k+1}-y_k} = \frac{x-x_k}{x_{k+1}-x_k}, \quad (1.3)$$

или в форме уравнения с угловым коэффициентом  $k$ :

$$y = y_k + \frac{y_{k+1}-y_k}{x_{k+1}-x_k}(x-x_k) = y_k + k(x-x_k) \quad (1.4)$$

Применение линейной интерполяции для приближенного вычисления значений функций обосновано в том случае, когда возникающая при этом погрешность невелика, то есть, заметно меньше погрешности измерений натуральных данных.

Пример: интерполируем функцию, заданную таблично, на интервале  $[1; 3]$  методом линейной интерполяции.

Таблица 1.3

$x$	0	1	3	4
$y$	1	3	23	12

Подставим в выражение (1.4) следующие значения:  $x_k = 1$ ,  $x_{k+1} = 3$ ,  $y_k = 3$ ,  $y_{k+1} = 23$ .

$$y = 3 + \frac{23-3}{3-1}(x-1) = 3 + 10(x-1) = 10x - 7 \quad (1.5)$$

Тогда для  $x = 2$ ,  $y = 13$ ;  $x = 1,25$ ,  $y = 5,5$  и так далее.

### 2. Квадратичная интерполяция.

Пусть снова дана функция  $f(x)$ , заданная таблично. Считая, что на промежутке  $(x_k, x_{k+2})$  данную функцию можно с достаточной степенью точности заменить квадратичной функцией, то есть, часть графика функции заменить параболой (рис. 4), необходимо найти значение функции  $f(x)$  в некоторой точке  $x$ , принадлежащей интервалу  $(x_k, x_{k+2})$ .

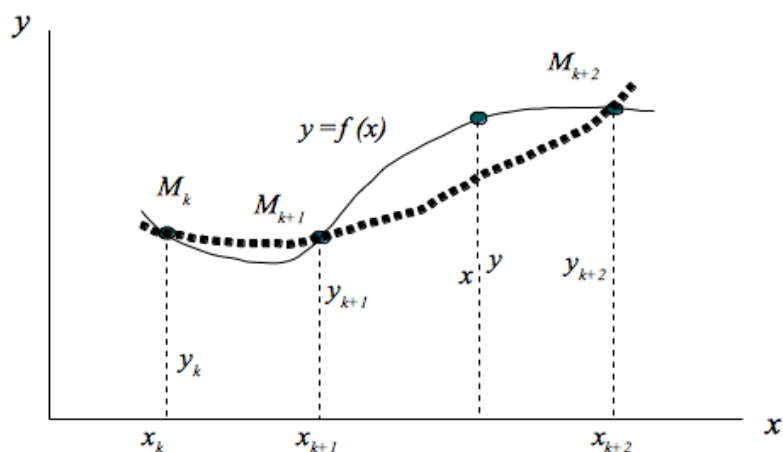


Рис. 4

Будем искать квадратичную функцию в следующем виде:  $y = ax^2 + bx + c$ . Исходя из условия совпадения значений искомой квадратичной функции с табличными значениями функции в трех заданных точках, составим следующую систему уравнений:

$$\begin{cases} y_k = ax_k^2 + bx_k + c \\ y_{k+1} = ax_{k+1}^2 + bx_{k+1} + c \\ y_{k+2} = ax_{k+2}^2 + bx_{k+2} + c \end{cases} \quad (1.6)$$

Это система из трех уравнений с тремя неизвестными, из решения которой получаются коэффициенты искомой квадратичной функции  $a$ ,  $b$  и  $c$ .

Пример: интерполируем функцию, заданную таблично, на интервале  $[1; 3]$  методом квадратичной интерполяции.

Таблица 1.4

$x$	0	1	2	3	4
$y$	1	5	12	23	12

Составим систему уравнений в виде  $y = ax^2 + bx + c$  для точек  $(1; 5)$ ,  $(2; 12)$  и  $(3; 23)$ .

$$\begin{cases} 5 = a \cdot 1^2 + b \cdot 1 + c \\ 12 = a \cdot 2^2 + b \cdot 2 + c \\ 23 = a \cdot 3^2 + b \cdot 3 + c \end{cases} \quad \begin{cases} 5 = a + b + c \\ 12 = 4a + 2b + c \\ 23 = 9a + 3b + c \end{cases} \quad \begin{cases} 5 = a + b + c \\ 7 = 3a + b \\ 18 = 8a + 2b \end{cases} \quad \begin{cases} a = 2 \\ b = 1 \\ c = 2 \end{cases} \quad (1.7)$$

Получаем функцию  $y = 2x^2 + x + 2$ . Взяв ее значение в точке, скажем,  $x = 1,5$ , получим  $y = 8$ .

### 1.3. Аппроксимация функций

#### **1.3.1. Постановка задачи аппроксимации функции одной переменной**



Задача **аппроксимации** (приближения) функции заключается в следующем: необходимо для заданной таблично функции  $F(x_i)$  найти заданную аналитически функцию  $f(x)$  заданного типа, которая в точках  $x_i$  принимает значения, как можно более близкие к табличным  $y_i$ . Задача аппроксимации функции обязательно должна учитывать характер поведения исходной функции на всем интервале наблюдений; при этом полное совпадение значений приближающей функции с табличными обязательным требованием не является, так как исходная функция сама по себе является в некоторой степени приближением из-за наличия погрешности при измерениях.

### 1.3.2. Порядок определения вида аппроксимирующей функции

1. По исходной таблице строится точечный график;
2. Проводится плавная кривая, отражающая характер расположения точек;
3. Полученная кривая сравнивается с известными видами функций;
4. Устанавливается вид функции;
5. Рассчитываются ее параметры.

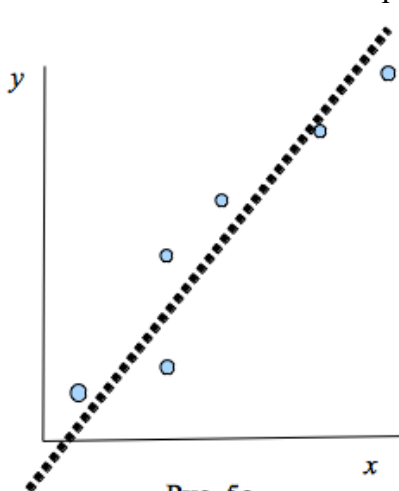


Рис. 5а

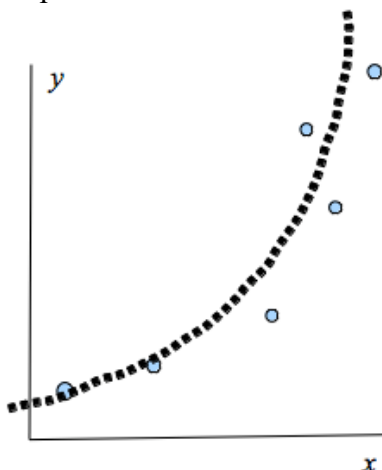


Рис. 5б

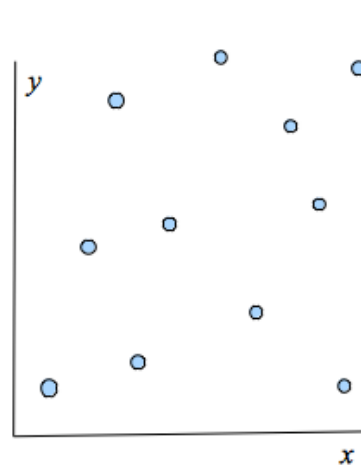


Рис. 5в

Рассмотрим три случая, приведенных на рисунке 5.

- а) отклонения точек от прямой небольшие и случайные;
- б) зависимость между  $x$  и  $y$  нелинейная;
- в) связь между  $x$  и  $y$ , если и существует, то описанию с удовлетворительной точностью не поддается.

В экспериментальной практике в качестве приближающих функций в зависимости от характера точечного графика  $F$  используются следующие приближающие функции:

1.  $y = ax + b$  — линейная
2.  $y = ax^2 + bx + c$  — квадратичная
3.  $y = ax^m$  — степенная
4.  $y = ae^{mx}$  — показательная
5.  $y = \frac{1}{ax+b}$  — дробно-линейная
6.  $y = a \ln x + b$  — логарифмическая
7.  $y = \frac{a}{x} + b$  — гиперболическая
8.  $y = \frac{x}{ax+b}$  — дробно-рациональная

### 1.3.3. Определение параметров аппроксимирующей функции методом

## наименьших квадратов

Метод оценки параметров приближающей функции, минимизирующий сумму квадратов отклонений наблюдений зависимой переменной от значения искомой функции, называется методом наименьших квадратов и заключается в следующем: для функции  $F$ , заданной таблично, находится  $f$  определенного вида такая, чтобы

$$\sigma = \sum_{i=1}^n (y_i - f(x_i))^2 \rightarrow \min \quad (1.8)$$

Рассмотрим в общем виде метод нахождения приближающей функции с тремя параметрами,  $y_i = F(x_i)$ ,  $y = f(x, a, b, c)$ .

$$\sum_{i=1}^n (y_i - f(x_i, a, b, c))^2 \rightarrow \min \quad (1.9)$$

Таким образом, задача сводится к отысканию минимума суммы квадратов отклонений. Из высшей математики известно, что экстремум функция имеет в той точке, где ее производная равна нулю. Применив необходимое условие экстремума,

$$\frac{\partial f}{\partial a} = \frac{\partial f}{\partial b} = \frac{\partial f}{\partial c} = 0 \quad (1.10)$$

получим систему уравнений для определения неизвестных параметров  $a$ ,  $b$  и  $c$ .

$$\begin{cases} \sum_{i=1}^n ((y_i - f(x_i, a, b, c)) \cdot f'_a(x_i, a, b, c)) = 0 \\ \sum_{i=1}^n ((y_i - f(x_i, a, b, c)) \cdot f'_b(x_i, a, b, c)) = 0 \\ \sum_{i=1}^n ((y_i - f(x_i, a, b, c)) \cdot f'_c(x_i, a, b, c)) = 0 \end{cases} \quad (1.11)$$

Решив систему относительно трех неизвестных, определим вид искомой функции  $f$ .

Для найденной эмпирической формулы в соответствии с исходной таблицей можно найти сумму квадратов отклонений:

$$\sigma = \sum_{i=1}^n (y_i - f(x_i, a, b, c))^2 \rightarrow \min \quad (1.12)$$

Если сложно сказать о виде приближающей функции, рассматривают несколько функций, а окончательный выбор делается в пользу той, у которой наименьшая  $\sigma$ .

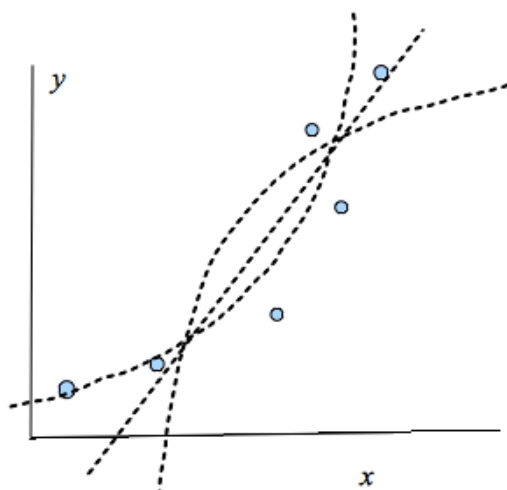


Рис. 6

$$\sigma_k = \sum_{i=1}^n (y_i - f_k(x_i, a, b, c))^2, \quad (1.13)$$

$$\sigma = \min(\sigma_1, \sigma_2, \dots, \sigma_k, \dots, \sigma_m) \quad (1.14)$$

Пример: применение метода в случае линейной функциональной зависимости.

Таблично заданы точки:

$x_1$	$x_2$	...	$x_n$
$y_1$	$y_2$	...	$y_n$

Определить приближающую функцию вида  $f = ax + b$ .

Оценим коэффициенты по методу наименьших квадратов:

$$\sum_{i=1}^n (y_i - (ax_i + b))^2 \rightarrow \min$$

$$\frac{\partial f}{\partial a} = x, \frac{\partial f}{\partial b} = 1 \quad (1.15)$$

Используя (1.11) и (1.15), запишем систему:

$$\begin{cases} \sum_{i=1}^n ((y_i - ax_i - b) \cdot x_i) = 0 \\ \sum_{i=1}^n ((y_i - ax_i - b) \cdot 1) = 0 \end{cases} \quad \begin{cases} \sum_{i=1}^n y_i x_i - a \sum_{i=1}^n x_i^2 - b \sum_{i=1}^n x_i = 0 \\ \sum_{i=1}^n y_i - a \sum_{i=1}^n x_i - nb = 0 \end{cases} \quad (1.16)$$

Разделив каждую часть на  $n$ , получим:

$$\begin{aligned} \left(\frac{1}{n} \sum_{i=1}^n x_i^2\right) a + \left(\frac{1}{n} \sum_{i=1}^n x_i\right) b &= \left(\frac{1}{n} \sum_{i=1}^n y_i x_i\right) \\ \left(\frac{1}{n} \sum_{i=1}^n x_i\right) a + b &= \left(\frac{1}{n} \sum_{i=1}^n y_i\right) \end{aligned} \quad (1.17a)$$

Введем обозначения:

$$\frac{1}{n} \sum_{i=1}^n x_i = M_x \quad (1.18a)$$

$$\frac{1}{n} \sum_{i=1}^n y_i = M_y \quad (1.18б)$$

$$\frac{1}{n} \sum_{i=1}^n x_i y_i = M_{xy} \quad (1.18в)$$

$$\frac{1}{n} \sum_{i=1}^n x_i^2 = M_{x2} \quad (1.18г)$$

Объединим систему (1.17a) и обозначения (1.18), получим систему:

$$\left. \begin{aligned} aM_{x2} + bM_x &= M_{xy} \\ aM_x + b &= M_y \end{aligned} \right\} \quad (1.17б)$$

Решив ее, найдем:

$$a = \frac{M_{xy} - M_x M_y}{M_{x2} - M_x^2} \quad (1.19)$$

$$b = M_y - aM_x \quad (1.20)$$

Пример: применение метода в случае упрощенной квадратичной формы функциональной зависимости.

Таблично заданы точки:

$x_1$	$x_2$	...	$x_n$
$y_1$	$y_2$	...	$y_n$

Определить приближающую функцию вида  $f = ax^2 + b$ .

Оценим коэффициенты по методу наименьших квадратов:

$$\sum_{i=1}^n (y_i - (ax_i^2 + b))^2 \rightarrow \min$$

$$\frac{\partial f}{\partial a} = x^2, \frac{\partial f}{\partial b} = 1 \quad (1.21)$$

$$\begin{cases} \sum_{i=1}^n ((y_i - ax_i^2 - b) \cdot x_i^2) = 0 \\ \sum_{i=1}^n ((y_i - ax_i^2 - b) \cdot 1) = 0 \end{cases} \quad (1.22)$$

$$\begin{cases} \sum_{i=1}^n y_i x_i^2 - a \sum_{i=1}^n x_i^4 - b \sum_{i=1}^n x_i^2 = 0 \\ \sum_{i=1}^n y_i - a \sum_{i=1}^n x_i^2 - nb = 0 \end{cases} \quad (1.23)$$

$$\begin{cases} \left( \frac{1}{n} \sum_{i=1}^n x_i^4 \right) a + \left( \frac{1}{n} \sum_{i=1}^n x_i^2 \right) b = \left( \frac{1}{n} \sum_{i=1}^n y_i x_i^2 \right) \\ \left( \frac{1}{n} \sum_{i=1}^n x_i^2 \right) a + b = \left( \frac{1}{n} \sum_{i=1}^n y_i \right) \end{cases} \quad (1.24a)$$

$$\frac{1}{n} \sum_{i=1}^n x_i^2 = M_{x2} \quad (1.18г)$$

$$\frac{1}{n} \sum_{i=1}^n y_i = M_y \quad (1.18б)$$

$$\frac{1}{n} \sum_{i=1}^n x_i^2 y_i = M_{x2y} \quad (1.18д)$$

$$\frac{1}{n} \sum_{i=1}^n x_i^4 = M_{x4} \quad (1.18е)$$

$$\begin{cases} aM_{x4} + bM_{x2} = M_{x2y} \\ aM_{x2} + b = M_y \end{cases} \quad (1.24б)$$

$$a = \frac{M_{x2y} - M_{x2}M_y}{M_{x4} - M_{x2}^2} \quad (1.25)$$

$$b = M_y - aM_{x2} \quad (1.26)$$

## Раздел 2

# Численное интегрирование

Из курса физики известен принцип сохранения энергии в замкнутой системе. Любые изменения энергии системы почти всегда приводят к необходимости вычисления различного рода интегралов от соответствующих функций. Часто функция  $f(x)$  имеет такой вид, что взять определенный интеграл по формуле Ньютона-Лейбница, известной из курса высшей математики, оказывается невозможно. В этих случаях прибегают к методам приближенного интегрирования, то есть к методам, позволяющим найти численное значение определенного интеграла приближенно с любой степенью точности.

## 2.1. Методы численного интегрирования

### 2.1.1. Метод прямоугольников

Идея численного интегрирования предельно проста и вытекает из геометрического смысла определенного интеграла — значение определенного интеграла численно равно площади криволинейной трапеции, ограниченной графиком функции  $y = f(x)$ , осью абсцисс и прямыми  $x = a$ ,  $x = b$ . Находя приближенно площадь криволинейной трапеции, мы получаем значение **интеграла**.

Считаем, что функция  $y = f(x)$  интегрируема на сегменте  $[a, b]$ . Требуется вычислить ее интеграл  $\int_a^b f(x)dx$ . Составим интегральную сумму для  $f(x)$  на сегменте, для чего разобьем его на  $n$  равных между собой частей с помощью точек  $x_1, x_2, \dots, x_n$ . Если длину каждой части мы при этом обозначим, как  $\Delta x = \frac{b-a}{n}$ , то для каждой точки будем иметь  $x_k = a + k\Delta x$ , где  $k = 0, 1, 2, \dots, n$ , а функция тогда примет вид  $y_k = f(a + k\Delta x)$ .

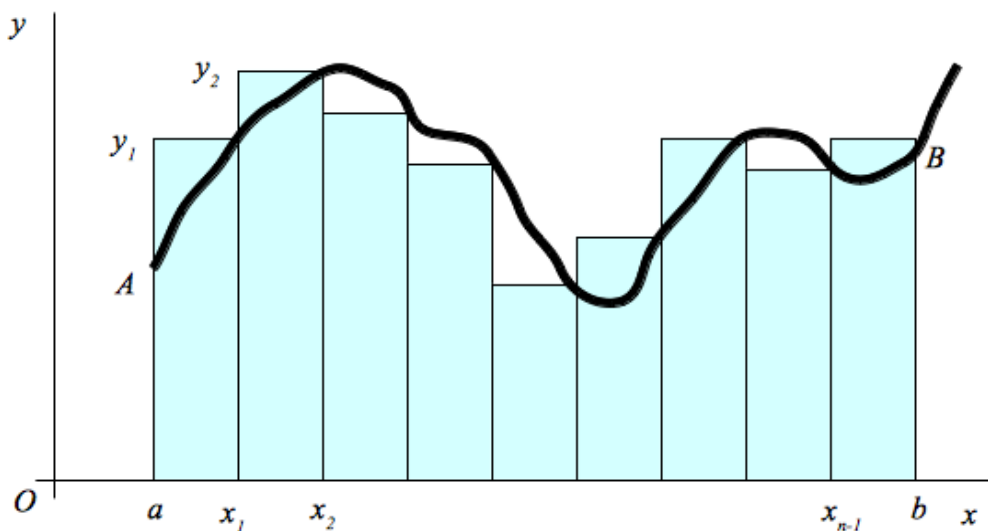


Рис. 7

Тогда сумма  $\sum_{k=1}^n y_k \Delta x$  будет интегральной для функции  $f(x)$  на отрезке  $[a, b]$ .

По определению интеграла имеем:  $\int_a^b f(x)dx = \lim_{\Delta x \rightarrow 0} \sum_{k=1}^n y_k \Delta x$ , тогда в качестве приближенного значения  $\int_a^b f(x)dx$  естественно взять  $\sum_{k=1}^n y_k \Delta x$ , то есть положить:

$\int_a^b f(x)dx \approx \sum_{k=1}^n y_k \Delta x$ , что с учетом  $\Delta x = \frac{b-a}{n}$  выглядит как

$$\int_a^b f(x)dx \approx \frac{b-a}{n} (y_1 + y_2 + \dots + y_n) \quad (2.1)$$

Это приближенное равенство называется **формулой прямоугольников**.

В том случае, когда  $f(x) \geq 0$ , формула (2.1) с геометрической точки зрения означает,

что площадь криволинейной трапеции  $aABb$ , ограниченной дугой кривой  $y=f(x)$ , осью  $Ox$  и прямыми  $x=a$ ,  $x=b$ , принимается приближенно равной площади ступенчатой фигуры, образованной из  $n$  прямоугольников с основаниями  $\Delta x = \frac{b-a}{n}$  и высотами  $y_1, y_2, \dots, y_n$ .

Способ приближенного вычисления определенного интеграла по этой формуле принято называть **методом прямоугольников**.

Всякое приближенное вычисление имеет определенную ценность лишь тогда, когда оно сопровождается оценкой допущенной при этом погрешности. Абсолютная величина погрешности  $R_n$ , которую мы допускаем, вычисляя интеграл  $\int_a^b f(x)dx$  по формуле прямоугольников, может быть оценена по формуле:

$$|R_n| \leq \frac{M(b-a)^2}{2n}, \quad (2.2)$$

где  $M$  — некоторое положительное число, такое, что для всех  $x \in [a, b]$  выполняется неравенство  $|f(x)| \leq M$ .

При неограниченном возрастании  $n$ , то есть при увеличении числа равных частей, на которые мы делим отрезок интегрирования, величина погрешности  $R_n$ , как видно из формулы (2), будет неограниченно уменьшаться, то есть, чем тоньше мы поделим на части отрезок интегрирования, тем точнее вычислим интеграл.

Хотя в инженерной практике чаще всего применяют именно метод прямоугольников, в математике существуют иные методы, обеспечивающие более точное вычисление интеграла.

### 2.1.2. Метод трапеций

$$\text{Площадь трапеции: } \frac{b-a}{n} \cdot \frac{y_k + y_{k+1}}{2}$$

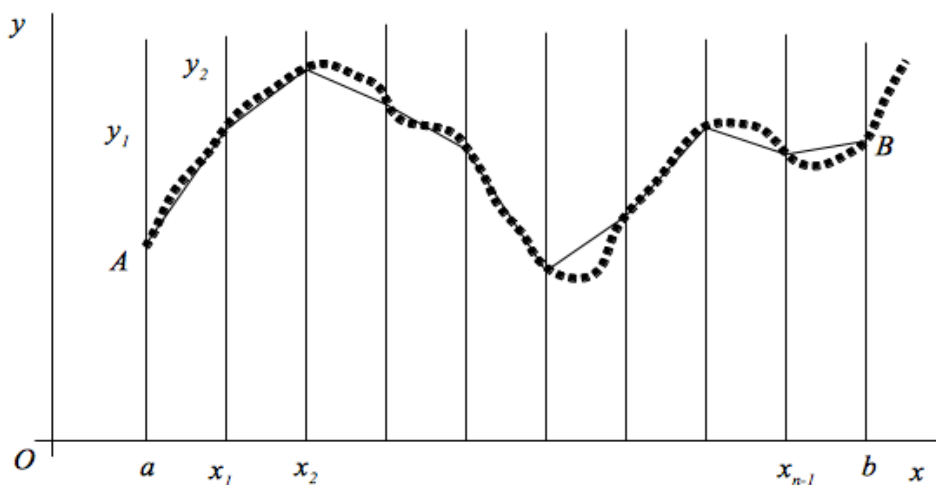


Рис. 8

Тогда приближенное значение интеграла, как суммы площадей этих трапеций, равно:

$$\int_a^b f(x)dx \approx \frac{b-a}{n} \left( \frac{y_0}{2} + y_1 + y_2 + \dots + y_{n-1} + \frac{y_n}{2} \right) \quad (2.3)$$

### 2.1.3. Метод парабол (метод Симпсона)

В этом методе хорды, стягивающие дуги, заменяют некоторыми параболой.

Сущность метода: отрезок интегрирования  $[a, b]$  делят на  $2n$  равных частей, для каждого  $x$ :  $x_0 = a, x_1, x_2, \dots, x_{2n-1}, x_{2n} = b$  определяются значения подынтегральной функции  $y$ :  $y_0, y_1, y_2, \dots, y_{2n}$ . Производят квадратичную интерполяцию на каждом из отрезков разбиения, то есть заменяют каждую дугу графика исходной функции дугой параболы с вертикальной осью. Формула парабол в окончательном виде примет вид:

$$\int_a^b f(x)dx \approx \frac{b-a}{6n} ((y_0 + y_{2n}) + 2(y_2 + y_4 + \dots + y_{2n-2}) + 4(y_1 + y_3 + \dots + y_{2n-1})) \quad (2.4)$$

## Раздел 3

# Введение в теорию систем и принятия решений

«Кибернетика» — греческое слово, происходящее от слова «кибер», означающего приставку «над-», и слова «наутик», означающего «моряк», таким образом «кибернаутик» — это кормчий. В научном плане к этому слову впервые обратился древнегреческий философ Платон, который назвал так науку управлять обществом, затем на две тысячи лет это слово забыли и вновь обрели лишь в середине XX века в более широком смысле.

Кибернетика — наука об общих законах автоматизированного управления, применимая к любой системе вне зависимости от ее физической природы. Кибернетика учит нас, как эффективно автоматизировать управление, но она не может без содействия других наук указать разумные цели управления, то есть, ради чего это управление производится. Для решения этих вопросов должен применяться системный подход с привлечением таких дисциплин, как системотехника, теория систем, синергетика, теория принятия решений и других. Поэтому в современном мире кибернетику чаще называют «прикладной математикой».

Исторически, при образовании кустарных производств, управление осуществлялось одним человеком, сейчас это невозможно. Производство разбито на отделы, подразделения, каждый из которых решает свои конкретные задачи.

Случилось так, что различные по природе явления, изучаемые различными науками, стали описываться одинаковыми математическими моделями: например, движение ракеты, работа промышленного предприятия и процесс размножения клеток могут быть описаны системами дифференциальных уравнений. Также существуют единые подходы к решению задач, принадлежащих к процессу управления в различных науках, но имеющих одинаковые математические или кибернетические модели.

## 3.1. Теория систем

### 3.1.1. Основные понятия

**Система** — множество предметов (объектов) в совокупности с их признаками и

связями между ними.

**Предметы (объекты)** — части или компоненты системы.

**Признаки** — свойства предметов (объектов).

$S = \{S_1, \dots, S_n; R_{12}, R_{13}, \dots\}$  —  
система,  
 $S_i$  — предметы,  
 $R_{ij}$  — связи.

Система производства,  
передачи и распределения  
электроэнергии:

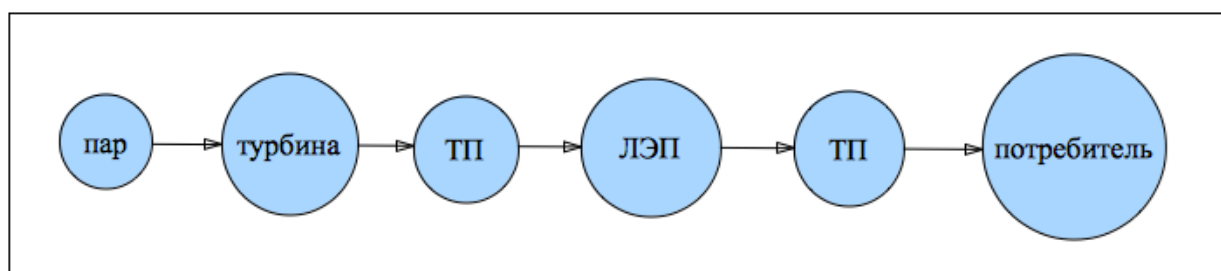
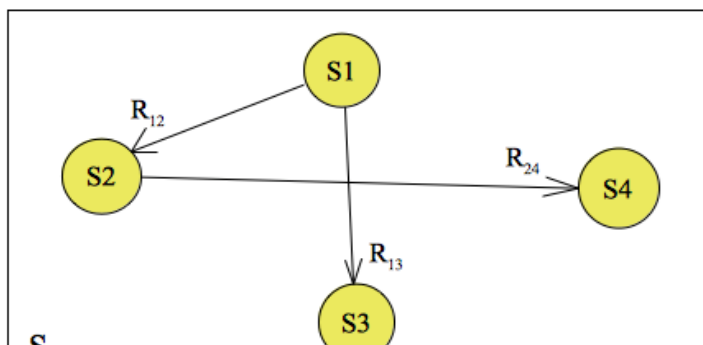


Рис. 10

Каждый из объектов может быть рассмотрен как отдельная система со своими внутренними предметами, связями и свойствами. Границы системы рассматриваются в зависимости от целей исследований.

### 3.1.2. Понятие системы на примере электрической и структурной схем

В энергетике система может быть представлена электрической схемой (рис. 11) и схемой замещения (рис. 12). Электрическая схема несет на себе функциональную нагрузку, схема замещения лишь показывает связь между элементами, процессы, происходящие в которых, описываются передаточным числом  $W = X_{вых} / X_{вх}$ , где  $X_{вых}$  и  $X_{вх}$  — выходной и входной сигналы соответственно.

Парадокс науки состоит в том, что, если мы хотим исследовать сложное явление, мы должны представить его упрощенно, то есть перейти от реальности к модели (например, от электрической схемы к структурной).



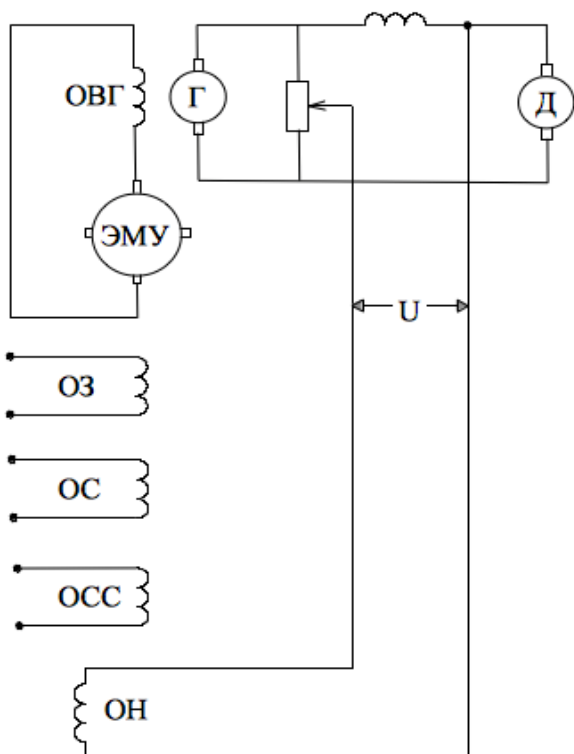


Рис. 11

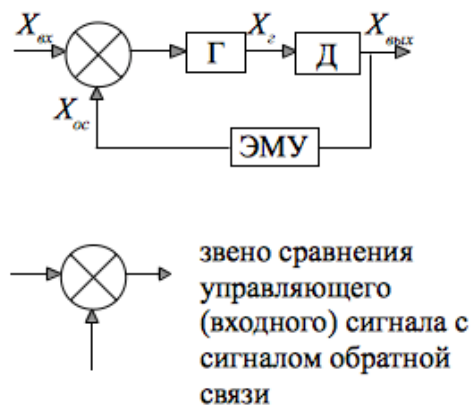


Рис. 12

$$P \uparrow, n \downarrow, U \uparrow$$

$$\uparrow U_{\text{двг}} = \overset{\text{const}}{U}_{\text{оз}} - \overset{U_{\text{ар}}}{U}_{\text{он}} \downarrow$$

### 3.1.3. Окружающие системы, внешняя среда

Любая рассматриваемая система находится в окружении множества явлений и факторов, влияющих на нее. В свою очередь, и рассматриваемая система влияет на эти внешние, окружающие ее явления и факторы.

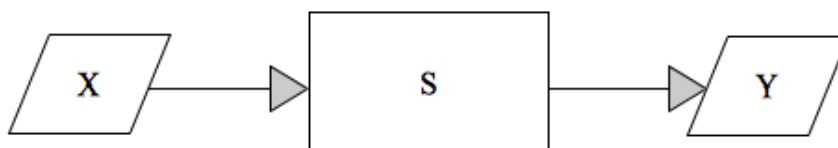


Рис. 13

Структурно это можно представить следующим образом:

$X = \{x_1, x_2, \dots, x_n\}$  — множество внешних воздействий на систему (входные параметры);

$Y = \{y_1, y_2, \dots, y_n\}$  — множество выходов системы, характеризующих состояние, поведение и функционирование системы (выходные параметры).

**Функционирование** системы — это целенаправленное изменение ее выходных параметров, изменение ее состояния.

На функционирование системы влияют целенаправленное изменение входных параметров, при этом  $X = \{E, U\}$ , где  $E$  — **неуправляемые входные воздействия** (среда, климат и тому подобные), а  $U$  — **управляемые входные воздействия**.

С учетом этого структурная схема системы примет следующий вид:

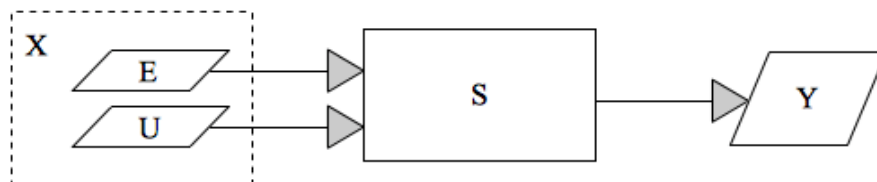


Рис. 14

Целенаправленное изменение выходных параметров  $Y$  посредством управляемого входа  $U$  при заданной среде  $E$  предполагает наличие целей системы.

**Цели системы** — это то, чего мы хотим достигнуть в процессе функционирования системы  $S$ .

$$C = \{c_1, c_2, \dots, c_n\}$$

Пример: цель — поддержание величины напряжения на выводах приемника.

Пути достижения:

- регулирование под нагрузкой (без отключения приемника)
- регулирование без нагрузки
- учет потерь напряжения в элементах цепи при проектировании

Пример: цель — регулирование скорости системы Г-Д-ЭМУ поперечного поля.

Достигается:

- посредством обратной связи по скорости (через тахометр)
- посредством обратной связи по напряжению (через делитель напряжения)

Цель может быть достигнута разными путями, то есть, например,  $C$  может быть достигнута как при  $U_1$ , так и при  $U_2$  и  $U_3$ .

Наличие нескольких способов достижения целей определяет задачу принятия решений — среди нескольких способов достижения целей выбрать и реализовать «наилучший».

Процесс организации «наилучшего» функционирования системы и есть **управление системой**.

### 3.1.4. Понятие «наилучшего» решения

Для того, чтобы выбрать «наилучшее» решение из множества возможных, необходимо иметь возможность сравнивать решения между собой, для чего следует ввести понятие качества решения, выражаемое показателем качества или их множеством:

$$P = \{p_1, p_2, \dots, p_m\} \quad (3.1)$$

Значения  $P$  характеризуют степень достижения поставленных целей.

Пример: цель функционирования ГЭС — обеспечение потребителей энергией  $\mathcal{E}$  за время  $T$ .

Цель достигается управлением:

$$\begin{aligned} U_1 &\rightarrow \mathcal{E}_1 \\ U_2 &\rightarrow \mathcal{E}_2 \end{aligned} \quad (3.2)$$

$$\begin{aligned} &\dots \\ U_n &\rightarrow \mathcal{E}_n \end{aligned}$$

Допустим, «наилучшее» решение — максимум  $\mathcal{E}$ , то есть  $\max_u P$ .

С помощью показателей качества можно сравнивать решения между собой, но нам необходимо выбрать одно «наилучшее», то есть должен существовать критерий выбора «наилучшего» решения.

$$K = \text{extr}_u P = \text{opt}_u P \quad (3.3)$$

«Наилучшее» — оптимальное, эффективное, рациональное (в отношении принятого критерия выбора) решение.

Задача принятия решений возникает, когда существует множество способов достижения поставленных целей и необходимо из них выбрать один «наилучший». В принятой нами терминологии это означает, что требуется найти такое управление, чтобы множество показателей качества было оптимальным (экстремальным).

$$U = \{U_1, U_2, \dots, U_n\} \quad (3.4)$$

$$P = \{p_1, p_2, \dots, p_m\} \quad (3.1)$$

$$K = \text{extr}_u P \quad (3.3)$$

Тогда найденное решение  $U_i$  является «наилучшим» — оптимальным, рациональным, эффективным (относительно выбранных критериев).

В общем случае:

$$P = \{p_1, p_2, \dots, p_m\} \quad (3.1)$$

$$K_1 = \text{extr}_u p_1, K_2 = \text{extr}_u p_2, \dots, K_m = \text{extr}_u p_m \quad (3.5)$$

$$K = \{K_1, K_2, \dots, K_m\} \quad (3.6)$$

### 3.1.5. Процесс принятия решений при проектировании и эксплуатации систем

Схема принятия решений содержит несколько этапов:

1. Вербальное описание:
  - описание целей функционирования системы;
  - описание работы системы;
  - описание показателей качества функционирования системы;

- формирование критериев выбора «наилучших» решений.
- 1. Формализованное описание:
  - построение функциональной модели системы:
$$Y = F(E, U) \quad (3.7)$$
  - построение модели расчета показателей качества функционирования системы:
$$P = G(E, U, Y) \quad (3.8)$$
  - формализация критериев выбора «наилучших» решений:
$$K = \underset{U}{extr} G(E, U, Y) \quad (3.9)$$
  - формализация описания ограничений:
$$Y \subseteq Y^{доп}, U \subseteq U^{доп} \quad (3.10)$$

$$\Psi(E, U, Y) = 0 \quad (3.11)$$
- 1. Сбор исходной информации (информационное обеспечение):

$$E, U^{доп}, Y^{доп}$$
- 2. Поиск эффективных решений:

$$\text{ЭВМ}$$
- 3. Окончательный выбор решения.
- 4. Исполнение решения, коррекция.

## Раздел 4

# Элементы теории вероятностей и математической статистики

### 4.1. Основные понятия

Задачей теории вероятностей является изучение вероятностных закономерностей массовых однородных случайных событий (например, опытов при  $S = const$ ).

Наблюдаемые события и явления можно разделить на три вида:

- **достоверные**,  $\Omega$  — к ним относится любое событие, которое обязательно произойдет, если будет осуществлена определенная совокупность условий  $S$ .
- **невозможные**,  $\emptyset$  — к ним относится любое событие, которое заведомо не произойдет, если будет осуществлена определенная совокупность условий  $S$ .
- **случайные** — к ним относится любое событие, которое при осуществлении определенной совокупности условий  $S$  может как произойти, так и не произойти.

Следует подчеркнуть, что теория вероятностей имеет дело не с явлениями окружающего мира непосредственно, а с их математическими моделями.

Теория вероятностей — это математическая модель явлений природы, технических или социологических объектов, систем и процессов, обладающих статистической устойчивостью.

Задачей же математической статистики является получение оценок параметров вероятностных закономерностей массовых однородных случайных событий по ограниченному числу экспериментальных данных.

### 4.1.1. Вероятность, частота событий

**Вероятность события** — числовая функция  $P$ , характеризующая возможность появления события  $A$ .

Вероятность можно оценить, как относительную долю благоприятных исходов  $m$  от общего числа событий  $n$ :

$$P(A) = \frac{m}{n} \quad (4.1)$$

$$0 \leq P(A) \leq 1 \quad (4.2)$$

$$P(\Omega) = 1, P(\emptyset) = 0 \quad (4.3)$$

Для математической статистики: если произведена серия из  $n$  опытов, в результате которых величина  $A$  появляется  $m$  раз, то частотой события или статистической вероятностью  $A$  в данной серии опытов называется отношение  $m$  к  $n$ .

$$P^*(A) = \frac{m}{n} \quad (4.4)$$

### 4.1.2. Случайные величины

Понятие **случайной величины** используется в математической статистике — это такая величина, конкретное значение которой невозможно однозначно предсказать при реализации определенной совокупности условий  $S$ .  $X$  — случайная величина,  $x$  — ее конкретное значение.

Случайные величины подразделяются на:

- **дискретные** — например, количество студентов, пропустивших текущее занятие по МЗЭВИЭ:

$$X = \{0, 1, 2, 3, \dots, n\}, \text{ где } n \text{ — количество студентов в группе;}$$

- **непрерывные** — например, количество воды, притекающей к ГЭС за год, в кубометрах:  $x \in [0; \infty)$ .

## 4.2. Закон распределения

### 4.2.1. Ряд распределения, многоугольник распределения

Рассмотрим прерывную случайную величину  $X$ , возможные значения которой —  $x_1, x_2, x_3, \dots, x_n$ . Каждое из этих значений возможно, но не достоверно, и величина  $X$  может принять каждое из них с некоторой вероятностью. В результате опыта величина  $X$  примет одно из этих значений, то есть произойдет одно из полной группы несовместных событий:

$$\begin{cases} X = x_1 \\ X = x_2 \\ X = x_3 \\ \dots \\ X = x_n \end{cases} \quad (4.5)$$

Обозначим вероятности этих событий:

$$P(X = x_1) = p_1, P(X = x_2) = p_2, \dots, P(X = x_n) = p_n. \quad (4.6)$$

Так как эти несовместные события образуют полную группу, то

$$\sum_{i=1}^n p_i = 1, \quad (4.7)$$

то есть сумма вероятностей всех возможных значений случайной величины равна единице. Эта суммарная вероятность каким-то образом распределена между отдельными значениями. Случайная величина будет полностью описана с вероятностной точки зрения, если мы зададим это распределение, то есть в точности укажем, какой вероятностью обладает каждое из событий. Этим мы установим так называемый закон распределения случайной величины.

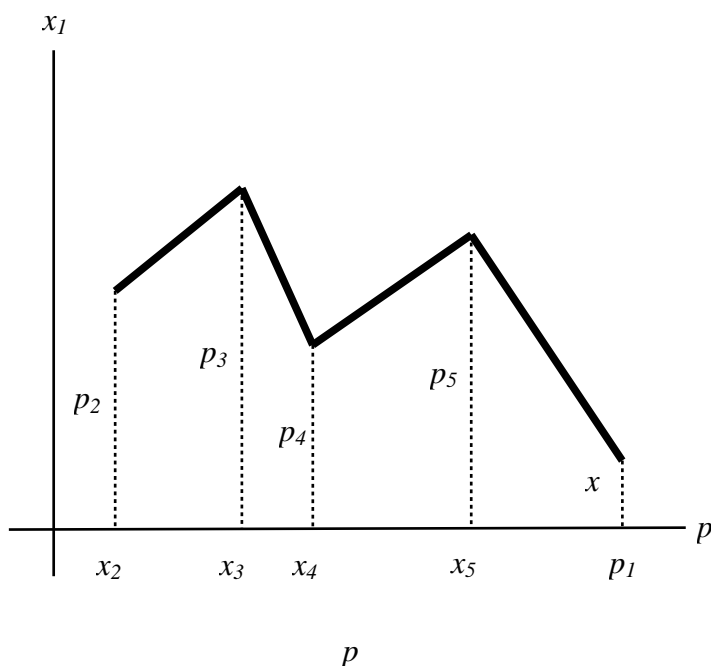
**Законом распределения случайной величины** называется всякое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями. Про случайную величину говорят, что она подчинена данному закону распределения.

Простейшей формой задания этого закона является таблица, в которой перечислены возможные значения случайной величины и соответствующие вероятности:

$x_i$	$x_1$	$x_2$	$x_3$	...	$x_n$
$p_i$	$p_1$	$p_2$	$p_3$	...	$p_n$

Такую таблицу называют **рядом распределения случайной величины X**.

Чтобы придать ряду распределения более наглядный вид, часто прибегают к его графическому изображению: по оси абсцисс откладываются возможные значения случайной величины, а по оси ординат — вероятности этих значений. Для наглядности полученные точки соединяются отрезками. Такая фигура называется **многоугольником распределения** (рис. 15). Многоугольник распределения, так же, как и ряд распределения, полностью характеризует прерывную случайную величину; он является одной из форм закона распределения.



Ряд и многоугольник распределения применимы только к дискретной случайной величине, но никак не к непрерывной.

### 4.2.2. Функция распределения

Для количественной характеристики распределения вероятностей случайной величины удобно пользоваться не вероятностью события  $X = x$ , а вероятностью события  $X < x$ , где  $x$  — некоторая текущая переменная. Вероятность этого события называется функцией распределения случайной величины называется **функцией распределения случайной величины  $X$**  и обозначается  $F(x)$ :

$$F(x) = P(X < x). \quad (4.8)$$

Функция распределения  $F(x)$  есть вероятность того, что случайная величина  $X$  всегда меньше своих возможных значений  $x$ .

Функцию распределения называют также интегральной функцией или интегральным законом распределения. Это самая универсальная форма закона распределения — она существует как для дискретных, так и для непрерывных случайных величин.

### 4.2.3. Свойства функции распределения

Рассмотрим случайную величину  $X$  как случайную точку на оси  $Ox$ , которая в результате опыта может занять то или иное положение. Тогда функция распределения  $F(x)$  есть вероятность того, что случайная точка  $X$  в результате опыта попадет левее точки  $x$ .

Будем увеличивать  $x$ , то есть перемещать точку  $x$  вправо по оси абсцисс. Очевидно, при этом вероятность того, что точка  $X$  попадет левее нее, не может уменьшаться, следовательно, функция распределения  $F(x)$  с возрастанием  $x$  убывать не может:

$$x_2 > x_1 \rightarrow F(x_2) \geq F(x_1) \quad (4.9)$$

При перемещении точки  $x$  вправо попадание точки  $X$  левее нее в пределе становится неизбежным, то есть достоверным:

$$F(+\infty) = 1. \quad (4.10)$$

Теперь станем передвигать точку  $x$  влево — в пределе попадание точки  $X$  левее нее становится невозможным:

$$F(-\infty) = 0. \quad (4.11)$$

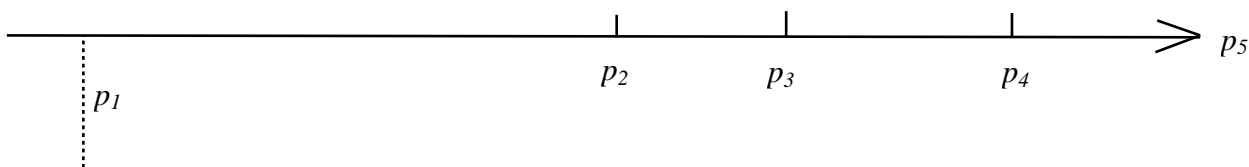


Рис. 16

График функции распределения  $F(x)$  в общем случае представляет собой график неубывающей функции (рис. 17), значения которой начинаются от 0 и доходят до 1 за промежуток, ограниченный наименьшим и наибольшим возможными значениями случайной величины.

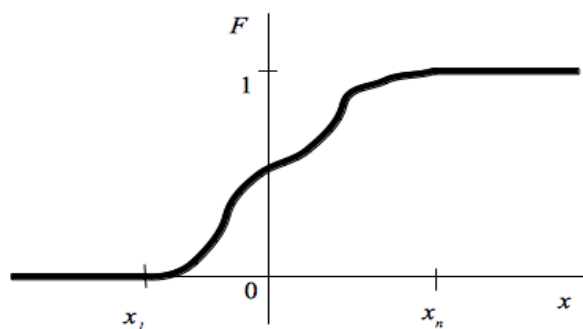


Рис. 17

Функция распределения любой прерывной случайной величины всегда есть разрывная ступенчатая функция, скачки которой происходят в точках, соответствующих возможным значениям случайной величины, и равны вероятностям этих значений. Сумма всех скачков  $F(x)$  равна единице.

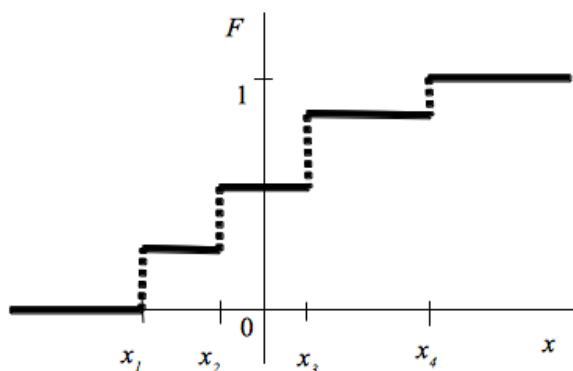


Рис. 18

#### 4.2.4. Вероятность попадания случайной величины в заданный интервал



Для определения вероятности попадания случайной величины в интервал от  $\alpha$  до  $\beta$  условимся левый конец  $\alpha$  включать в интервал, а правый  $\beta$  — не включать. Тогда

$$\alpha \leq X < \beta.$$

Выразим вероятность этого события через функцию распределения величины  $X$ . Для этого рассмотрим три события:

событие  $A$ , состоящее в том, что  $X < \beta$ ;

событие  $B$ , состоящее в том, что  $X < \alpha$ ;

событие  $C$ , состоящее в том, что  $\alpha \leq X < \beta$ .

Учитывая, что  $A = B + C$ , согласно теореме сложения вероятностей имеем:

$$P(X < \beta) = P(X < \alpha) + P(\alpha \leq X < \beta) \quad (4.8)$$

$$F(\beta) = F(\alpha) + P(\alpha \leq X < \beta), \quad (4.9)$$

$$\text{откуда } P(\alpha \leq X < \beta) = F(\beta) - F(\alpha), \quad (4.10)$$

то есть вероятность попадания случайной величины на заданный участок равна приращению функции распределения на этом участке.

#### 4.2.5. Плотность распределения

Пусть имеется некоторая непрерывная случайная величина  $X$  с функцией распределения  $F(x)$ , которую мы предположим непрерывной и дифференцируемой. Вычислим вероятность попадания этой случайной величины на участок от  $x$  до  $x + \Delta x$ :

$$P(x \leq X < x + \Delta x) = F(x + \Delta x) - F(x), \quad (4.11)$$

то есть приращение функции распределения на этом участке. Рассмотрим отношение этой вероятности к длине участка, то есть среднюю вероятность, приходящуюся на единицу длины на этом участке, и будем приближать  $\Delta x$  к нулю. Что получим в предельном случае? Производную от функции распределения:

$$\lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = F'(x).$$

Введем обозначение:

$$f(x) = F'(x), \quad (4.18)$$

и назовем эту функцию **плотностью распределения** (также называется «**плотностью вероятности**») непрерывной случайной величины  $X$ .

Кривая, изображающая плотность распределения случайной величины, называется **кривой распределения** (рис. 19).

Плотность распределения так же, как и функция распределения, есть одна из форм закона распределения, но существует она только для непрерывных случайных величин.

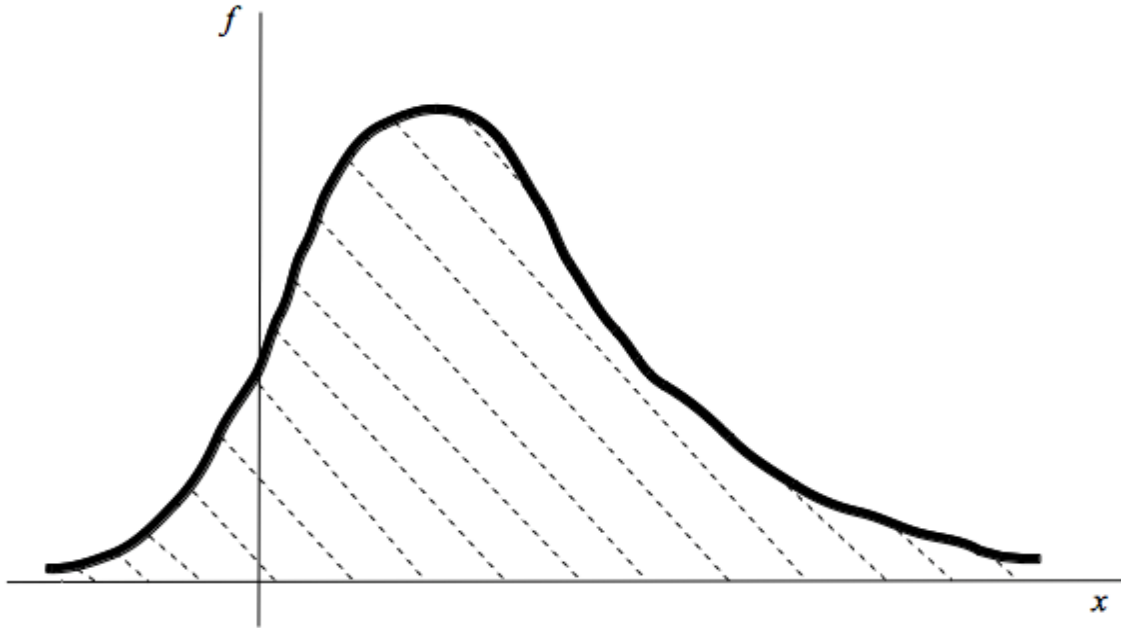


Рис. 19

#### 4.2.6. Основные свойства плотности распределения

1. Плотность распределения есть неотрицательная функция:

$$f(x) \geq 0 \quad (4.19)$$

Это свойство непосредственно вытекает из того, что функция распределения  $F(x)$  есть неубывающая функция.

2. Интеграл в бесконечных пределах от плотности распределения равен единице:

$$\int_{-\infty}^{\infty} f(x) dx = 1$$

Это следует из того, что  $F(+\infty) = 1$ .

Геометрически основные свойства плотности распределения означают, что:

- всякая кривая распределения лежит не ниже оси абсцисс;
- полная площадь, ограниченная кривой распределения и осью абсцисс, равна единице.

Функция распределения  $F(x)$ , как и всякая вероятность, есть величина безразмерная.

Размерность плотности распределения  $f(x)$ , как видно из формулы  $f(x) = F'(x) = \frac{dF}{dx}$ , обратна размерности случайной величины.

Пример: случайная величина  $X$  подчинена закону распределения с плотностью:

$$\begin{cases} f(x) = a \cos x & \text{при } -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} \\ f(x) = 0 & \text{при } x < -\frac{\pi}{2} \text{ или } x > \frac{\pi}{2} \end{cases}$$

- найти коэффициент  $a$ ;
- построить график плотности распределения  $f(x)$ ;
- найти функцию распределения  $F(x)$  и построить ее график;

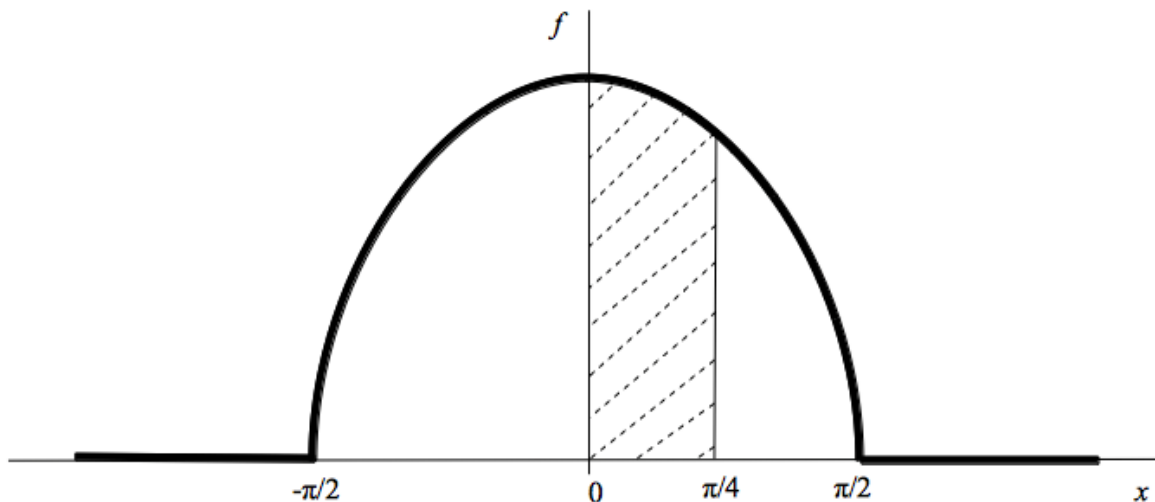


Рис. 20

- найти вероятность попадания величины  $X$  на участок от 0 до  $\frac{\pi}{4}$ .

Для определения коэффициента  $a$  воспользуемся свойством плотности распределения:

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} a \cos x dx = 2a = 1,$$

откуда  $a = \frac{1}{2}$ , тогда  $f(x) = \frac{\cos x}{2}$ .

График плотности распределения  $f(x)$  представлен на рисунке 20.

Функция распределения находится, исходя из того, что плотность распределения — это ее производная:

$$F(x) = \int_{-\infty}^x f(x) dx$$

Таким образом,

$$\left\{ \begin{array}{ll} F(x) = \int_{-\infty}^x 0 dx = 0 & \text{при } x < -\frac{\pi}{2} \\ F(x) = 0 + \int_{-\frac{\pi}{2}}^x \frac{\cos x}{2} dx = \frac{\sin x + 1}{2} & \text{при } -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} \\ F(x) = 0 + \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\cos x}{2} dx + \int_{\frac{\pi}{2}}^{\infty} 0 dx = 1 & \text{при } x > \frac{\pi}{2} \end{array} \right.$$

График функции  $F(x)$  изображен на рисунке 21.

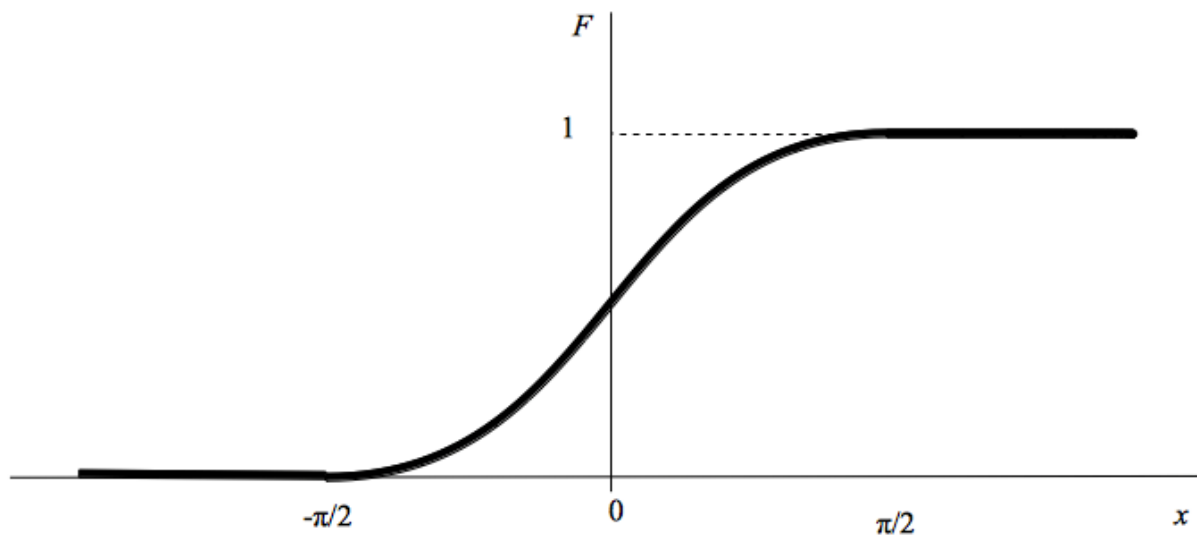


Рис. 21

Вероятность попадания случайной величины на участок находим, как приращение функции распределения на этом участке:

$$P(0 < X < \frac{\pi}{4}) = \frac{\sin \frac{\pi}{4} + 1}{2} - \frac{\sin 0 + 1}{2} = \frac{\sqrt{2}}{4}.$$

Пример для самостоятельного решения: плотность распределения задана формулой:

$$f(x) = \frac{1}{\pi(1+x^2)}.$$

- построить график плотности;
- найти вероятность того, что величина  $X$  попадет на участок от  $-1$  до  $+1$ .

$$P(-1 < X < 1) = \int_{-1}^1 \frac{dx}{\pi(1+x^2)} = \frac{1}{\pi} \arctg x \Big|_{-1}^1 = \frac{1}{2}$$

### 4.3. Числовые характеристики случайных величин

Мы изучили законы распределения.

Для дискретной случайной величины это:

- функция распределения;
- ряд (графически — многоугольник) распределения.

Для непрерывной:

- функция распределения;
- плотность (графически — кривая) распределения.

Каждый закон распределения представляет собой некоторую функцию, и указание этой функции полностью описывает случайную величину с вероятностной точки зрения.

Однако, во многих практических вопросах достаточно бывает указать только отдельные числовые параметры, характеризующие существенные черты распределения случайной величины.

Характеристики, назначение которых — выразить в сжатой форме наиболее существенные особенности распределения, называются числовыми характеристиками случайной величины. С их помощью удастся решить задачу, оставляя в стороне законы распределения и оперируя одними числовыми характеристиками.

В теории вероятностей и математической статистике применяется большое количество различных числовых характеристик, имеющих различное назначение и различные области применения. Из них рассмотрим:

- математическое ожидание;
- дисперсию;
- среднеквадратическое отклонение случайной величины.

#### 4.3.1. Математическое ожидание

Рассмотрим дискретную случайную величину  $X$ , имеющую возможные значения  $x_1, x_2, \dots, x_n$  с вероятностями  $p_1, p_2, \dots, p_n$ . Нам требуется охарактеризовать каким-то числом положение значений случайной величины на оси абсцисс с учетом того, что эти значения имеют различные вероятности. Для этой цели естественно воспользоваться так называемым «средним взвешенным» из значений  $x_i$ , причем каждое значение  $x_i$  при осреднении должно учитываться с «весом», пропорциональным вероятности этого значения. Таким образом, мы вычислим среднее значение случайной величины  $X$ , которое мы обозначим  $M[X]$ :

$$m_X = M[X] = \frac{x_1 p_1 + x_2 p_2 + \dots + x_n p_n}{p_1 + p_2 + \dots + p_n} = \frac{\sum_{i=1}^n x_i p_i}{\sum_{i=1}^n p_i}, \quad (4.20)$$

или, учитывая, что  $\sum_{i=1}^n p_i = 1$ ,

$$m_X = M[X] = \sum_{i=1}^n x_i p_i \quad (4.21)$$

Это среднее взвешенное и называется **математическим ожиданием случайной величины**.

Математическим ожиданием случайной величины называется сумма произведений всех возможных значений случайной величины на вероятности этих значений.

На графике математическое ожидание совпадает с «модой случайной величины» — наиболее вероятным значением.

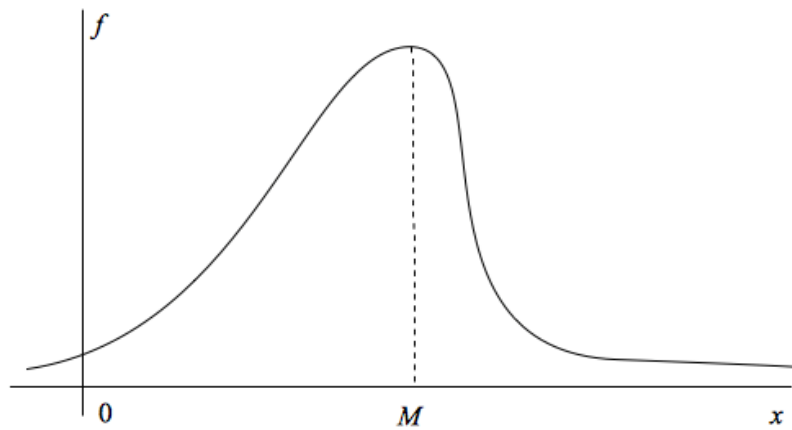


Рис. 22

Измеряется математическое ожидание в тех же единицах, что и случайная величина.

### 4.3.2. Дисперсия и среднееквадратическое отклонение

**Дисперсия случайной величины** есть характеристика рассеивания, разбросанности значений случайной величины около ее математического ожидания.

Для непосредственного вычисления дисперсии служат формулы:

$$D_X = D[X] = \sum_{i=1}^n (x_i - m_X)^2 p_i \text{ — для дискретной величины,} \quad (4.22)$$

$$D_X = D[X] = \int_{-\infty}^{\infty} (x - m_X)^2 f(x) dx \text{ — для непрерывной.} \quad (4.23)$$

Дисперсия случайной величины имеет размерность квадрата самой величины; для наглядной характеристики рассеивания удобнее пользоваться величиной с совпадающей размерностью. Для этого из дисперсии извлекают квадратный корень. Полученная величина называется средним квадратическим отклонением (иначе — «стандартом») случайной величины  $X$ . Среднее квадратическое отклонение будем обозначать буквой «сигма»:

$$\sigma_X = \sigma[X] = \sqrt{D_X}.$$

Пример: определить характеристики величины  $X$  — математическое ожидание, дисперсию и среднееквадратическое отклонение.

Ряд распределения величины  $X$  имеет вид:

$x_i, B$	0	1	2	3
$p_i$	0,216	0,432	0,288	0,064

Вычисляем числовые характеристики величины  $X$ :

$$m_X = M[X] = \sum_{i=1}^n x_i p_i = 0 \cdot 0,216 + 1 \cdot 0,432 + 2 \cdot 0,288 + 3 \cdot 0,064 = 1,2B,$$

$$D_X = D[X] = \sum_{i=1}^n (x_i - m_X)^2 p_i = (0 - 1,2)^2 \cdot 0,216 + (1 - 1,2)^2 \cdot 0,432 + \\ + (2 - 1,2)^2 \cdot 0,288 + (3 - 1,2)^2 \cdot 0,064 = 0,72B^2$$

$$\sigma_X = \sigma[X] = \sqrt{D_X} = \sqrt{0,72} = 0,848В.$$

## 4.4. Частные случаи законов распределения

### **4.4.1. Закон равномерной плотности**

В некоторых задачах практики встречаются непрерывные случайные величины, о которых заранее известно, что их возможные значения лежат в интервале, в пределах которого все значения случайной величины одинаково вероятны, то есть обладают одной и той же плотностью вероятности. О таких величинах говорят, что они распределяются по закону равномерной плотности.

Например: поезда метрополитена идут с интервалом в две минуты. Пассажир выходит на платформу в некоторый момент времени. Время  $T$ , в течение которого ему придется ждать следующего поезда, представляет собой случайную величину, распределенную с равномерной плотностью на интервале  $(0, 2)$  минут.

Как найти эту плотность вероятности? Для этого рассмотрим случайную величину  $X$ , подчиненную закону равномерной плотности на участке от  $\alpha$  до  $\beta$ , и напишем для нее выражение плотности распределения  $f(x)$ , которая постоянная и равна  $c$  на отрезке  $(\alpha, \beta)$ ; вне этого отрезка она равна нулю.

$$f(x) = \begin{cases} c & \text{при } \alpha < x < \beta \\ 0 & \text{при } x < \alpha \text{ и } x > \beta \end{cases}$$

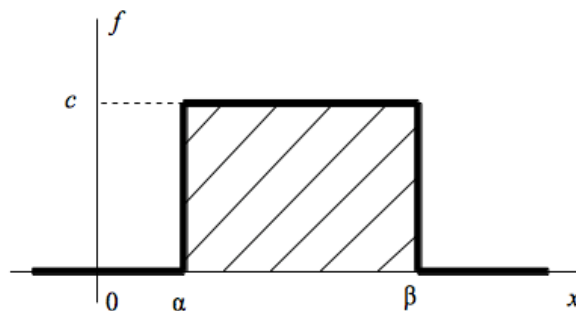


Рис. 23

Так как площадь, ограниченная кривой распределения, равна единице:

$$c(\beta - \alpha) = 1,$$

то

$$c = \frac{1}{\beta - \alpha}.$$

Таким образом, закон равномерной плотности имеет следующий вид:

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{при } \alpha < x < \beta \\ 0 & \text{при } x < \alpha \text{ и } x > \beta \end{cases}$$

Напишем выражение для функции распределения  $F(x)$ . Функция распределения выражается площадью кривой распределения, лежащей левее точки  $x$ . Следовательно,

$$F(x) = \begin{cases} 0 & \text{при } x < \alpha \\ \frac{x - \alpha}{\beta - \alpha} & \text{при } \alpha < x < \beta \\ 1 & \text{при } x > \beta \end{cases}$$

График функции  $F(x)$  приведен на рисунке 24

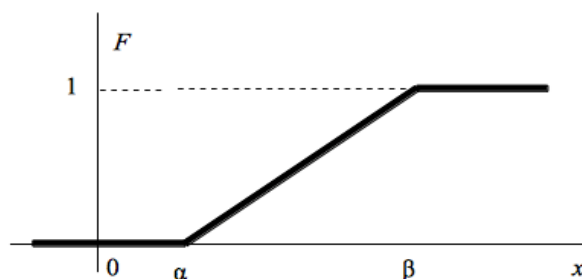


Рис. 24

#### 4.4.2. Нормальный закон распределения

Нормальный закон распределения (или закон Гаусса) — наиболее часто встречающийся на практике закон распределения. Он является предельным законом, к которому приближаются другие законы распределения при весьма часто встречающихся типичных условиях.

Нормальный закон распределения характеризуется плотностью распределения вида:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

Параметр  $m$  — математическое ожидание величины  $X$ . Этот параметр, особенно в задачах стрельбы, часто называют также центром рассеивания.

Параметр  $\sigma$  — среднее квадратическое отклонение величины  $X$ .

Кривая распределения по нормальному закону имеет симметричный холмообразный вид (рисунок 25). Максимальная ордината кривой, равная  $\frac{1}{\sigma\sqrt{2\pi}}$ , соответствует точке  $x = m$ ; по мере удаления от точки  $m$  плотность распределения падает, и при  $x \rightarrow \pm\infty$  кривая

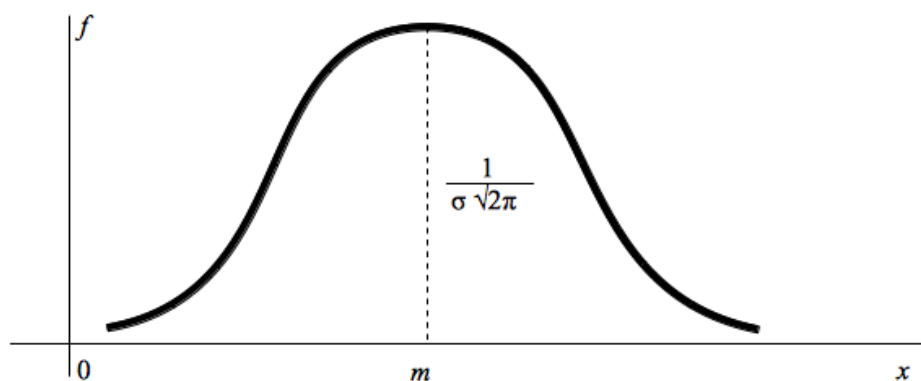


Рис. 25

асимптотически приближается к оси абсцисс.

Из выражения для плотности нормального распределения видно, что центром



симметрии распределения является центр рассеивания  $m$ . Это ясно из того, что при изменении знака разности  $(x - m)$  на обратный выражение не меняется. Если изменять центр рассеивания  $m$ , кривая распределения будет смещаться по оси абсцисс, не меняя своей формы (рисунок 26). Центр распределения характеризует положение распределения на оси абсцисс.

Размерность центра рассеивания — та же, что размерность случайной величины  $X$ .

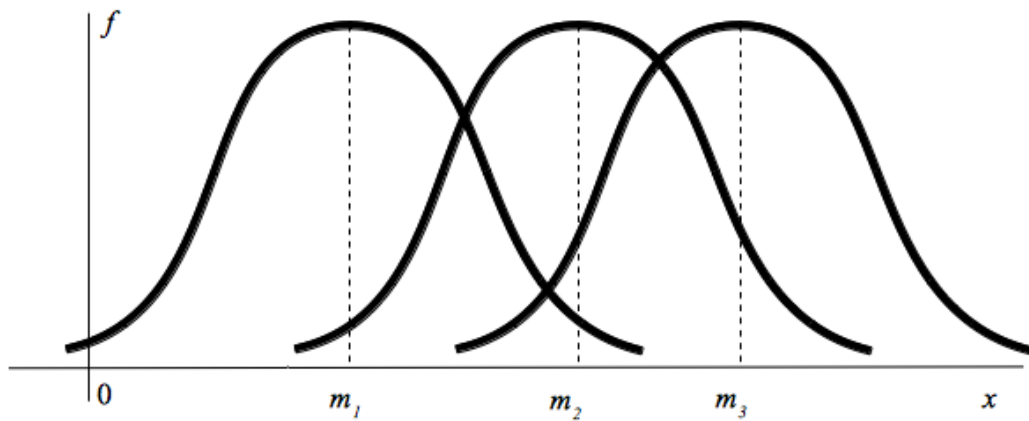


Рис. 26

Параметр  $\sigma$  характеризует не положение, а самую форму кривой распределения. Это есть характеристика рассеивания. Наибольшая ордината кривой распределения обратно пропорциональна  $\sigma$ ; при увеличении  $\sigma$  максимальная ордината уменьшается, так как площадь под кривой распределения всегда равна единице.

На рисунке 27 показаны три нормальные кривые при  $m = 0$ ; из них кривая I соответствует наименьшему значению  $\sigma$ , а кривая III — наибольшему. Изменение параметра  $\sigma$  равносильно изменению масштаба кривой распределения — увеличению масштаба по одной оси вместе с уменьшением по другой.

Размерность параметра  $\sigma$  также совпадает с размерностью случайной величины  $X$ .

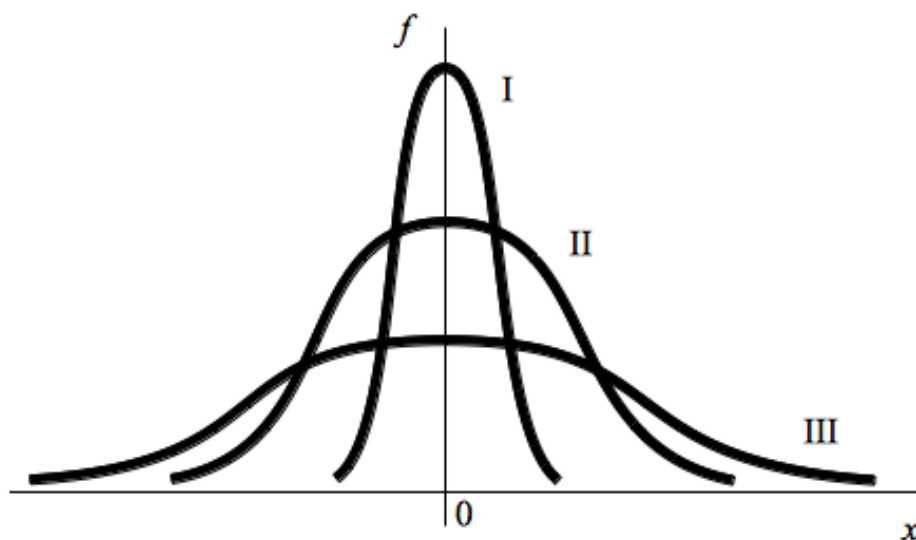


Рис. 27

## **4.5. Основные задачи математической статистики**

Разработка методов регистрации, описания и анализа статистических экспериментальных данных, получаемых в результате наблюдения массовых случайных явлений, составляет предмет специальной науки — математической статистики.

Все задачи математической статистики касаются вопросов обработки наблюдений над массовыми случайными явлениями.

Рассмотрим основные из них:

### **1. Задача определения закона распределения случайной величины (или системы случайных величин) по статистическим данным.**

При обработке обширных статистических данных закон распределения случайной величины можно определить наиболее точно. На практике нам всегда приходится иметь дело с ограниченным количеством экспериментальных данных; в связи с этим результаты наших наблюдений и их обработки всегда содержат больший или меньший элемент случайности. Возникает вопрос о том, какие черты наблюдаемого явления относятся к постоянным, устойчивым и действительно присущи ему.

### **2. Задача проверки правдоподобия гипотез.**

Статистический материал может с большим или меньшим правдоподобием подтверждать или не подтверждать справедливость той или иной гипотезы. Например: согласуются ли результаты эксперимента с гипотезой о том, что данная случайная величина подчинена закону распределения  $F(x)$ ?

Для решения подобных вопросов математическая статистика выработала ряд специальных приемов.

### **3. Задача нахождения неизвестных параметров распределения.**

Часто при обработке статистического материала вовсе не возникает вопрос об определении законов распределения исследуемых случайных величин. Обыкновенно это бывает связано с крайне недостаточным объемом экспериментального материала. Иногда же характер закона распределения качественно известен до опыта, из теоретических

соображений; тогда возникает более узкая задача обработки наблюдений — определить только числовые характеристики случайной величины.

С задачей отыскания «подходящих значений» числовых характеристик тесно связана задача оценки их точности и надежности при ограниченном статистическом наборе.

### 4.5.1. Простой статистический ряд. Статистическая функция распределения

Предположим, что изучается некоторая случайная величина  $X$ , закон распределения которой в точности неизвестен, и требуется определить этот закон из опыта. С этой целью над случайной величиной  $X$  производится ряд независимых опытов (наблюдений). В каждом из этих опытов случайная величина принимает определенное значение.

Таблица, в которой записаны случайные величины (результаты опытов) в порядке их поступления называется простым статистическим рядом. Например:

$i$ , номер опыта	1	2	3	4	5	6	7	8	9	10
$x_i$ , результат опыта	1	5	15	11	1	15	8	11	1	11

Одним из способов обработки статистического материала является построение статистической функции распределения случайной величины.

**Статистической функцией распределения случайной величины  $X$**  называется частота события  $X < x$  в данном статистическом материале:

$$F^*(x) = P^*(X < x).$$

Для того, чтобы найти значение статистической функции распределения при данном  $x$ , достаточно подсчитать число опытов, в которых величина  $X$  приняла значение, меньшее  $x$ , и разделить на общее число  $n$  произведенных опытов.

Статистическая функция распределения любой случайной величины, — прерывной или непрерывной, — представляет собой прерывную ступенчатую функцию, скачки которой соответствуют наблюдаемым значениям случайной величины и по величине равны частотам этих значений. Если каждое отдельное значение случайной величины  $X$  было наблюждено только один раз, скачок статистической функции распределения в каждом значении будет равен  $\frac{1}{n}$ , где  $n$  — число наблюдений.

Пример: построить статистическую функцию распределения для случайной величины, заданной таблицей выше.

Так как наименьшее наблюдаемое значение величины равно 1, то функция распределения для  $x < 1$  равна нулю:

$$F^*(x < 1) = 0.$$

$x = 1$  встречается три раза, следовательно, его частота равна  $\frac{3}{10}$  и именно такой скачок будет иметь  $F^*(x)$  в точке 1. В промежутке от 1 до 5 функция будет иметь значение  $\frac{3}{10}$ , которое затем в точке 5 поднимется до  $\frac{4}{10}$ , так как значение случайной величины, равное пяти, наблюдалось один раз. Таким же образом строится весь остальной график статистической функции распределения, изображенный на рисунке 28.

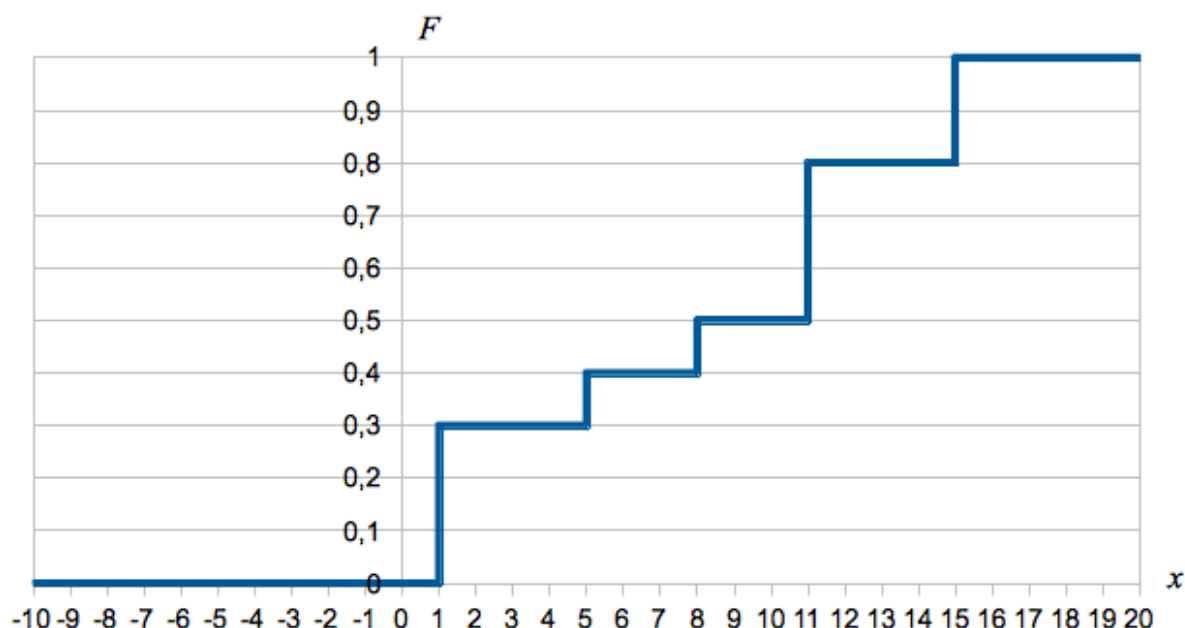


Рис. 28

При увеличении числа опытов  $n$ , согласно теореме Бернулли, при любом  $x$  частота события  $X < x$  приближается (сходится по вероятности) к вероятности этого события. Следовательно, при увеличении  $n$  статистическая функция  $F^*(x)$  приближается (сходится по вероятности) к подлинной функции распределения  $F(x)$  случайной величины  $X$ .

Если  $X$  — непрерывная случайная величина, то при увеличении числа наблюдений  $n$  число скачков функции  $F^*(x)$  увеличивается, сами скачки уменьшаются, и график функции  $F^*(x)$  неограниченно приближается к плавной кривой  $F(x)$  — функции распределения величины  $X$ .

#### 4.5.2. Статистический ряд. Гистограмма

При большом числе опытов (порядка сотен и более) использование простого статистического ряда и функции распределения становится громоздким. В этом случае простой ряд преобразовывается в статистический следующим образом: выстраивают значения случайной величины в порядке возрастания, затем делят весь диапазон наблюдений на интервалы или «разряды», подсчитывают количество попаданий  $m_i$  случайной величины в каждый  $i$ -ый разряд, и, наконец, находят частоту, соответствующую каждому разряду, поделив  $m_i$  на общее число наблюдений  $n$ :

$$p_i^* = \frac{m_i}{n}.$$

Сумма частот всех разрядов, очевидно, должна быть равна единице.

Таблица, в которой приведены разряды в порядке их возрастания и соответствующие частоты, называется статистическим рядом.

$I_i$	$x_1 \dots x_2$	$x_2 \dots x_3$	...	$x_i \dots x_{i+1}$	...	$x_k \dots x_{k+1}$
$p_i^*$	$p_1^*$	$p_2^*$	...	$p_i^*$	...	$p_k^*$

Здесь  $I_i$  — обозначение  $i$ -го разряда;  $x_i \dots x_{i+1}$  — его границы;  $p_i^*$  — соответствующая частота;  $k$  — число разрядов.

Случайная величина, попавшая на границу интервалов, может быть условно отнесена или в предыдущий интервал, или в следующий, или в долях распределена между ними; однако, выбранная такая условность должна быть единой для всех таких пограничных величин в данном ряду.

Практика показывает, что в большинстве случаев рационально выбирать число разрядов порядка 10–20. Длины разрядов могут быть как одинаковыми, так и различными. Проще, разумеется, брать их одинаковыми, но в частных случаях их делят неравномерно (это зависит от плотности распределения).

Пример: преобразовать простой ряд из параграфа 4.5.1 в статистический.

$I_i$	1 ... 5	5 ... 8	8 ... 13	13 ... 18
$m_i$	4	1	3	2
$p_i^*$	$\frac{4}{10}$	$\frac{1}{10}$	$\frac{3}{10}$	$\frac{2}{10}$

$$n = 10,$$

$$\sum p_i^* = 1.$$

**Пример:** произведено 500 измерений скорости ветра в условной точке. Результаты измерений сведены в статистический ряд:

$I_i, \frac{м}{с}$	1 ... 4	4 ... 8	8 ... 12	12 ... 16	16 ... 20	20 ... 24	24 ... 28	28 ... 32
$m_i$	6	25	72	133	120	88	46	10
$p_i^*$	0,012	0,050	0,144	0,266	0,240	0,176	0,092	0,020

Графическое изображение статистического ряда называется гистограммой.

Гистограмма строится следующим образом:

- по оси абсцисс откладываются разряды
- на каждом из разрядов, как на основании, строится прямоугольник, площадь которого равна частоте этого разряда; соответственно, высота этого прямоугольника равна частоте разряда, разделенной на его ширину

Из способа построения диаграммы следует, что площадь ее равна единице.

Построим гистограмму для статистического ряда из последнего примера.

$I_i, \frac{м}{с}$	1 ... 4	4 ... 8	8 ... 12	12 ... 16	16 ... 20	20 ... 24	24 ... 28	28 ... 32
$p_i^*$	0,012	0,050	0,144	0,266	0,240	0,176	0,092	0,020
$\frac{p_i^*}{I_i}, \left(\frac{м}{с}\right)^{-1}$	0,0040	0,0125	0,0360	0,0665	0,0600	0,0440	0,0230	0,0050

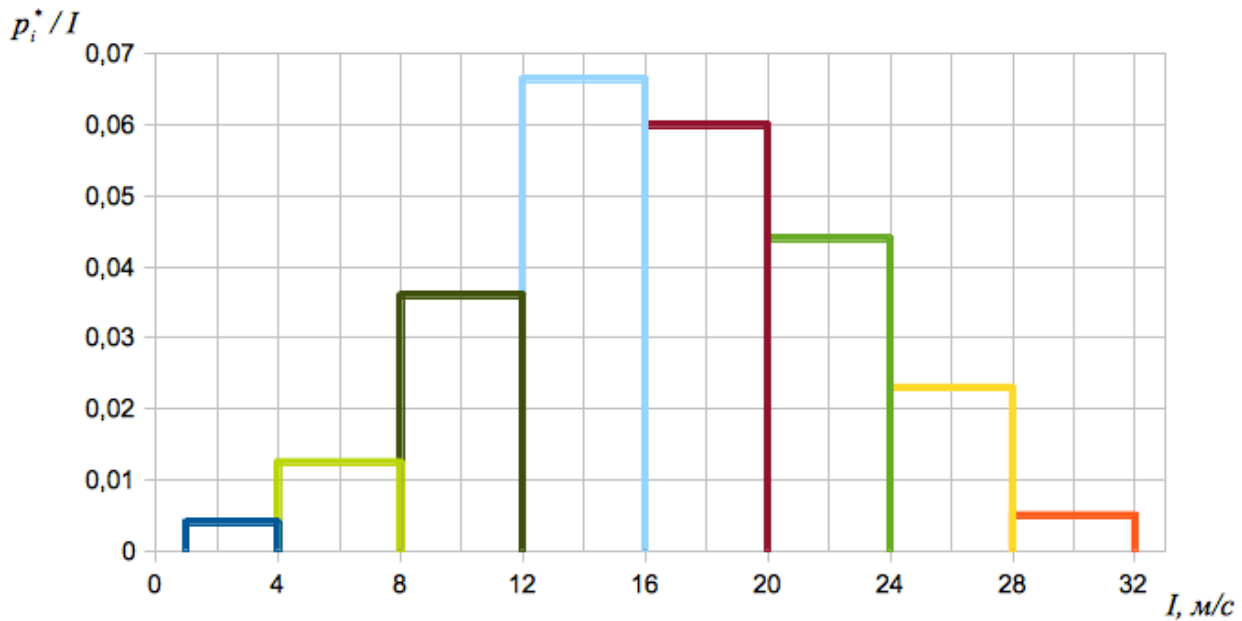


Рис. 29

Очевидно, при увеличении числа опытов можно выбирать все более и более мелкие разряды; при этом гистограмма будет приближаться к кривой, ограничивающей площадь, равную единице, а это не что иное, как график плотности распределения величины  $X$ .

Пользуясь данными статистического ряда, можно приближенно построить и статистическую функцию распределения величины  $X$ . Построение точной статистической функции распределения с несколькими сотнями скачков во всех наблюдаемых значениях  $X$  слишком трудоемко и себя не оправдывает. Для практики достаточно построить статистическую функцию распределения по нескольким точкам. В качестве этих точек удобно взять границы разрядов —  $x_1, x_2, \dots, x_{k+1}$ . Тогда, очевидно,

$$\begin{aligned}
 F^*(x_1) &= 0; \\
 F^*(x_2) &= p_1^*; \\
 &\dots \\
 F^*(x_k) &= \sum_{i=1}^{k-1} p_i^*; \\
 F^*(x_{k+1}) &= \sum_{i=1}^k p_i^* = 1.
 \end{aligned}$$

Соединяя полученные точки ломаной линией или плавной кривой, получим приближенный график статистической функции распределения.

Пример: построить приближенную статистическую функцию распределения по данным статистического ряда из последнего примера.

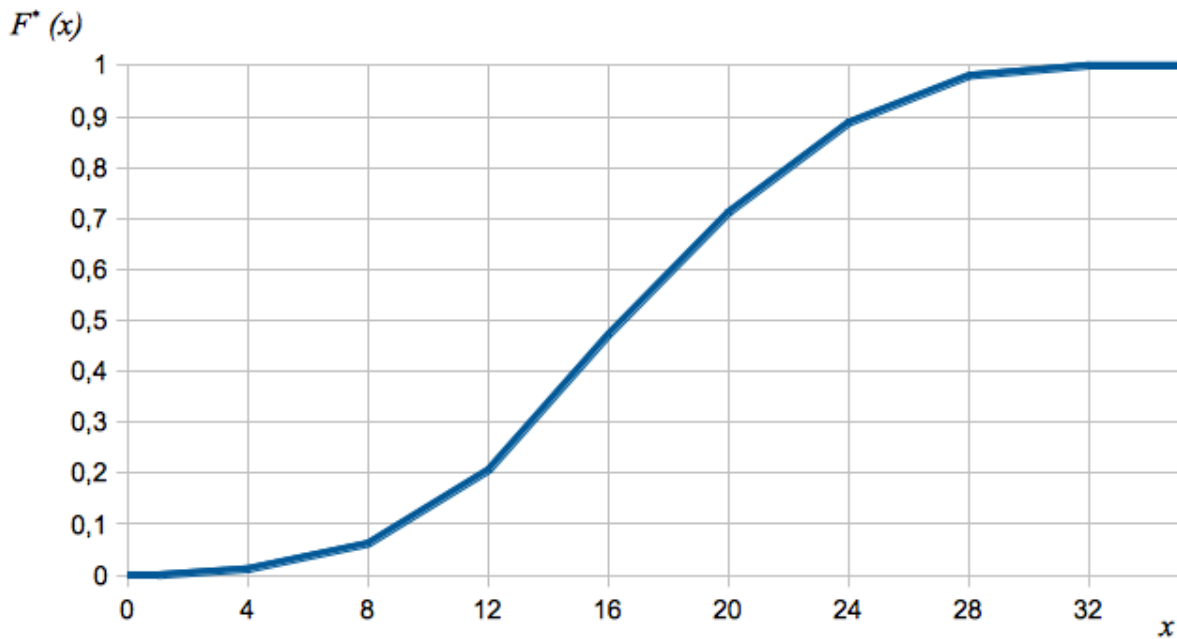


Рис. 30

#### 4.5.3. Числовые характеристики статистического распределения

Каждой числовой характеристике случайной величины  $X$  соответствует ее статистическая аналогия. Для математического ожидания случайной величины такой аналогией является среднее арифметическое наблюдаемых значений случайной величины:

$$M^*[X] = \frac{\sum_{i=1}^n x_i}{n},$$

где  $x_i$  — значение случайной величины, наблюдаемое в  $i$ -ом опыте, а  $n$  — число опытов.

Эту характеристику называют **статистическим средним случайной величины**.

При очень большом количестве опытов статистические числовые характеристики будут выражаться приближенными формулами:

$$m_X^* = M^*[X] = \sum_{i=1}^k \tilde{x}_i p_i^*,$$

$$D_X^* = D^*[X] = \sum_{i=1}^k (\tilde{x}_i - m_X^*)^2 p_i^*,$$

$$\sigma_X^* = \sqrt{D_X^*},$$

где  $\tilde{x}_i$  — «представитель»  $i$ -го разряда,  $p_i^*$  — частота  $i$ -го разряда,  $k$  — число разрядов.

Как видно, данные формулы полностью аналогичны формулам числовых характеристик (математического ожидания, дисперсии и среднеквадратического отклонения) случайной величины  $X$ , с той только разницей, что вместо вероятностей  $p_i$  здесь стоят частоты разрядов  $p_i^*$ , вместо математического ожидания  $m_X$  — статистическое среднее  $m_X^*$ , вместо числа возможных значений величины — число разрядов.

При увеличении числа наблюдений все статистические характеристики будут сходиться по вероятности к соответствующим математическим характеристикам и при

достаточном  $n$  могут быть приняты приближенно равными им.

#### 4.5.4. Выравнивание статистических рядов

Задача **выравнивания ряда** (выдвижения гипотезы) заключается в том, чтобы подобрать теоретическую плавную кривую распределения, наилучшим образом описывающую данное статистическое распределение, которое для удобства мы представим в виде гистограммы.

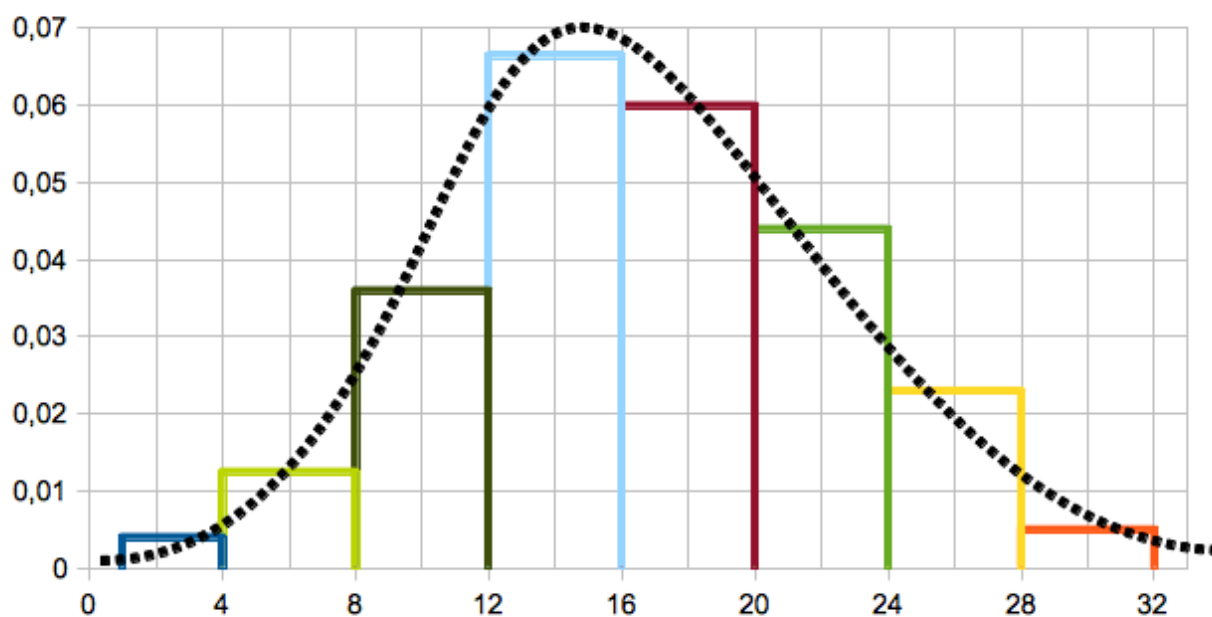


Рис. 31

Принципиальный вид теоретической кривой выбирается заранее из соображений, связанных с существом задачи, или просто сообразно с внешним видом статистического распределения. Затем задача выравнивания статистического ряда переходит в задачу рационального выбора тех значений параметров, при которых соответствие между статистическим и теоретическим распределениями оказывается наилучшим.

Если считать, что величина  $X$  подчиняется нормальному закону распределения:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-m)^2}{2\sigma^2}},$$

то задача выравнивания переходит в задачу о рациональном выборе параметров  $m$  и  $\sigma$  в выражении.

**Пример:** требуется выровнять данное распределение с помощью нормального закона.

$I_i$	-4 ... -3	-3 ... -2	-2 ... -1	-1 ... 0	0 ... 1	1 ... 2	2 ... 3	3 ... 4
$x_{icp}$	-3,5	-2,5	-1,5	-0,5	0,5	1,5	2,5	3,5
$p_i^*$	0,012	0,050	0,144	0,266	0,240	0,176	0,092	0,020



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-m)^2}{2\sigma^2}}$$

Нормальный закон зависит от двух параметров —  $m$  и  $\sigma$ . Подберем их так, чтобы сохранить математическое ожидание и дисперсию статистического распределения.

Вычислим статистическое среднее ряда:

$$\begin{aligned} m_X^* &= -3,5 \cdot 0,012 - 2,5 \cdot 0,050 - 1,5 \cdot 0,144 - 0,5 \cdot 0,266 + 0,5 \cdot 0,240 + 1,5 \cdot 0,176 + 2,5 \cdot \\ &\quad \cdot 0,092 + 3,5 \cdot 0,020 = \\ &= -0,042 - 0,125 - 0,216 - 0,133 + 0,12 + 0,264 + 0,23 + 0,07 = 0,168 \end{aligned}$$

Вычислим дисперсию:

$$\begin{aligned} D_X^* &= (-3,5 - 0,168)^2 \cdot 0,012 + (-2,5 - 0,168)^2 \cdot 0,050 + (-1,5 - 0,168)^2 \cdot 0,144 \\ &\quad + (-0,5 - 0,168)^2 \cdot 0,266 + \\ &+ (0,5 - 0,168)^2 \cdot 0,240 + (1,5 - 0,168)^2 \cdot 0,176 + (2,5 - 0,168)^2 \cdot 0,092 + (3,5 - 0,168)^2 \\ &\quad \cdot 0,020 = \\ &= 0,161 + 0,356 + 0,401 + 0,119 + 0,026 + 0,312 + 0,500 + 0,222 = 2,097 \end{aligned}$$

Выберем параметры распределения:

$$m = m_X^* = 0,168$$

$$\sigma = \sqrt{D_X^*} = 1,448$$

Конечный вид плотности распределения:

$$f(x) = \frac{1}{1,448 \cdot \sqrt{2\pi}} e^{\frac{-(x-0,168)^2}{2 \cdot 1,448^2}} = 0,2755 \cdot e^{-0,2385 \cdot (x-0,168)^2}$$

$x$	-4	-3	-2	-1	0	1	2	3	4
$f(x)$	0,004	0,025	0,090	0,199	0,274	0,234	0,124	0,041	0,008

Гистограмма заданного ряда и график полученной плотности распределения изображены на рисунке 32.

Из графика видно, что теоретическая кривая распределения  $f(x)$ , сохраняя, в основном, существенные особенности статистического распределения, свободна от случайных неправильностей хода гистограммы, которые могут быть отнесены за счет случайных причин.

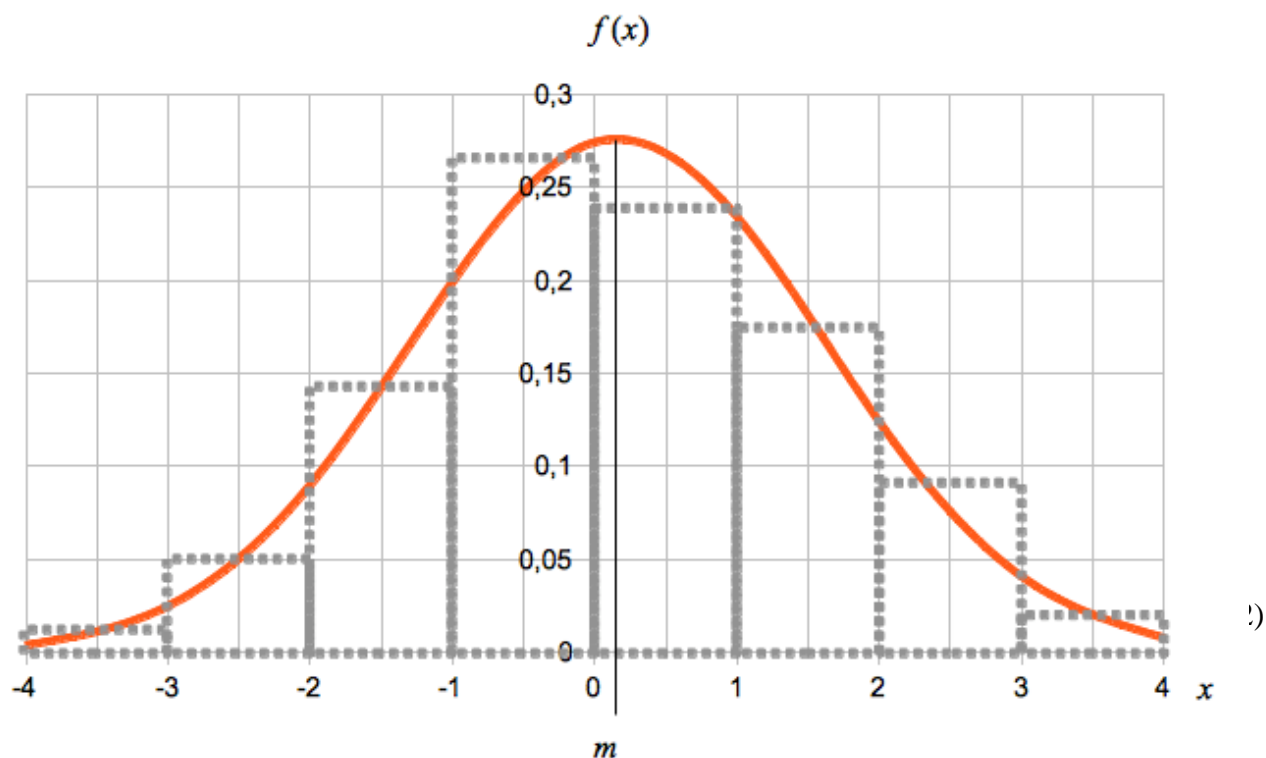


Рис. 32

#### 4.5.5. Критерии согласия. Критерий Пирсона

На основании данного статистического материала нам предстоит проверить гипотезу  $H$ , состоящую в том, что случайная величина  $X$  подчиняется некоторому определенному закону распределения. Это выполняется с помощью критериев согласия.

Рассмотрим один из наиболее часто используемых критериев согласия — так называемый «критерий  $\chi^2$ » (хи квадрат) Пирсона.

Предположим, что произведено  $n$  независимых опытов, в каждом из которых случайная величина  $X$  приняла определенное значение. Результаты опытов сведены в  $k$  разрядов и оформлены в виде статистического ряда:

$I_i$	$x_1 \dots x_2$	$x_2 \dots x_3$	...	$x_k \dots x_{k+1}$
$p_i^*$	$p_1^*$	$p_2^*$	...	$p_k^*$

Требуется проверить, согласуются ли экспериментальные данные с гипотезой о том, что случайная величина  $X$  имеет данный закон распределения (заданный функцией распределения  $F(x)$  или плотностью распределения  $f(x)$ ). Назовем этот закон распределения «теоретическим».

Зная теоретический закон распределения, можно найти теоретические вероятности попадания случайной величины в каждый из разрядов:

$$p_1, p_2, \dots, p_k.$$

Проверяя согласованность теоретического и статистического распределений, мы будем исходить из расхождений между теоретическими вероятностями  $p_i$  и статистическими

частотами  $p_i^*$ . Естественно выбрать в качестве меры расхождения сумму квадратов отклонений, взятых с некоторыми весами  $c_i$ :

$$U = \sum_{i=1}^k c_i (p_i^* - p_i)^2.$$

Коэффициенты  $c_i$  («веса» разрядов) вводятся потому, что в общем случае отклонения, относящиеся к различным разрядам, нельзя считать равноправными по значимости. Действительно, одно и то же по абсолютной величине отклонение  $p_i^* - p_i$  может быть малозначимым, если сама вероятность  $p_i$  велика, и очень заметным, если она мала. Поэтому естественно «веса»  $c_i$  взять обратно пропорциональными вероятностям разрядов  $p_i$ .

Далее возникает вопрос о том, как выбрать коэффициент пропорциональности.

К. Пирсон показал, что если положить

$$c_i = \frac{n}{p_i},$$

то при больших  $n$  закон распределения величины  $U$  приближается к так называемому «распределению  $\chi^2$ »:

$$\chi^2 = n \sum_{i=1}^k \frac{(p_i^* - p_i)^2}{p_i}.$$

Введем  $n$  под знак суммы и, учитывая, что  $p_i^* = \frac{m_i}{n}$ , где  $m_i$  — число значений в  $i$ -ом разряде, приведем формулу к виду:

$$U = \chi^2 = \sum_{i=1}^k \frac{(m_i - np_i)^2}{np_i}.$$

Распределение  $\chi^2$  зависит от параметра  $r$ , называемого числом «степеней свободы» распределения. Число «степеней свободы»  $r$  равно числу разрядов  $k$  минус число независимых условий («связей»), наложенных на частоты  $p_i^*$ . Примерами таких условий могут быть:

- $\sum_{i=1}^k p_i^* = 1$ , если мы требуем только того, чтоб сумма частот была равна единице (требование неизбежное и накладывается во всех случаях);
- $\sum_{i=1}^k \tilde{x}_i p_i^* = m_X$ , если мы подбираем теоретическое значение с тем условием, чтобы его ожидание совпадало со статистическим средним ряда;
- $\sum_{i=1}^k (\tilde{x}_i - m_X^*)^2 p_i^* = D_X$ , если мы требуем, кроме того, совпадения теоретической и статистической дисперсий;

и так далее.

Таким образом, схема применения критерия  $\chi^2$  к оценке согласованности теоретического и статистического распределений сводится к следующему:

- определяется мера расхождения  $\chi^2$  по соответствующей формуле;
- определяется число степеней свободы  $r$  как число разрядов  $k$  минус число положенных связей  $s$ :

$$r = k - s;$$

- по  $r$  и  $\chi^2$  с помощью таблицы определяется вероятность того, что величина, имеющая распределение  $\chi^2$  с  $r$  степенями свободы, превзойдет данное значение  $\chi^2$ . Если эта вероятность весьма мала ( $< 0,1$ ), гипотеза отбрасывается как неправдоподобная. Если эта вероятность достаточно велика, гипотезу можно признать не противоречащей опытными данным.

**Пример:** проверить согласованность теоретического и статистического распределений.

$I_i$	-4 ... -3	-3 ... -2	-2 ... -1	-1 ... 0	0 ... 1	1 ... 2	2 ... 3	3 ... 4
$p_i^*$	0,012	0,050	0,144	0,266	0,240	0,176	0,092	0,020

Пользуясь теоретическим нормальным законом распределения с параметрами

$$m = 0,168, \quad \sigma = 1,448,$$

находим вероятности попадания в разряды по формуле

$$p_i = \Phi^* \left( \frac{x_{i+1} - m}{\sigma} \right) - \Phi^* \left( \frac{x_i - m}{\sigma} \right),$$

где  $x_i, x_{i+1}$  — границы  $i$ -го разряда.

Затем составляем сравнительную таблицу чисел попаданий в разряды  $m_i$  и соответствующих значений  $n p_i$  ( $n = 500$ ).

$I_i$	-4 ... -3	-3 ... -2	-2 ... -1	-1 ... 0	0 ... 1	1 ... 2	2 ... 3	3 ... 4
$p_i^*$	0,012	0,050	0,144	0,266	0,240	0,176	0,092	0,020
$m_i$	6	25	72	133	120	88	46	10
$np_i$	6,2	26,2	71,2	122,2	131,8	90,5	38,2	10,5

Находим значение меры расхождения:

$$\chi^2 = \sum_{i=1}^k \frac{(m_i - np_i)^2}{np_i} = 3,94.$$

Определяем число степеней свободы (в данном случае  $s = 3$ ):

$$r = k - s = 8 - 3 = 5.$$

По таблице находим для  $r = 5$ :

$$\begin{aligned} \text{при } \chi^2 = 3,00 \quad p &= 0,70; \\ \text{при } \chi^2 = 4,35 \quad p &= 0,50. \end{aligned}$$

Следовательно, искомая вероятность  $p$  при  $\chi^2 = 3,94$  приближенно равна 0,56. Эта

вероятность малой не является, поэтому гипотезу о том, что рассматриваемая величина  $X$  распределена по нормальному закону, можно считать правдоподобной.

## Раздел 5

# Методы оптимизации

**Оптимизация** (поиск наиболее эффективного решения) заключается в определении максимального (минимального) значения целевой функции (например, критерия оценки эффективности)  $f(x_1, x_2, \dots, x_n)$  при заданных ограничениях:

$$g_i(x_1, x_2, \dots, x_n) \leq b_i,$$

где  $f, g_i$  — заданные функции,  $b_i$  — действительные числа.

Решением данных задач занимается математическое программирование — совокупность математических методов, предназначенных для исследования и решения оптимизационных задач.

Оптимизационные задачи:

- математическое программирование (определенные, детерминированные методы)
  - линейное программирование
  - нелинейное программирование
  - выпуклое программирование
  - квадратичное программирование
  - многоэкстремальные задачи
  - параметрическое программирование
  - дробно-линейное программирование
  - целочисленное программирование
  - динамическое программирование
- стохастическое программирование
  - вероятностный характер
  - неопределенный характер

Стохастическое программирование — задачи, в которых исходная информация содержит элементы неопределенности, либо некоторые параметры задачи носят случайный характер с известными вероятностными характеристиками. Неопределенность присуща многим задачам возобновляемой энергетики из-за стохастического характера метеорологических процессов. Так, в гидроэнергетике это испарение, осадки, в ветроэнергетике — скорость ветра и его напор, в солнечной энергетике — солнечная радиация.

Основная трудность стохастического программирования состоит в постановке самой задачи из-за сложности исходной информации.

Существует множество способов учета неопределенности, которые зависят от ее характера и от того, как и что будет в первую очередь влиять на функционирование системы. Один из способов — это замена неопределенных величин либо ожидаемыми значениями, либо наихудшими критическими. Тогда становится возможным далее применять детерминированный подход.

Детерминированные задачи делятся на группы в зависимости от свойств целевой функции  $f$  и ограничений  $g_i$ .

Линейное программирование — целевая функция и функции ограничений

(образующие множество допустимых решений) задаются системой линейных равенств и неравенств. В линейном программировании существуют классы задач, структура которых позволяет создать специальные методы их решения, например, транспортные задачи, задачи использования ресурсов, «о диете» и ряд других.

Нелинейное программирование применяется, когда задача описывается нелинейными целевой функцией и функциями ограничений. Нелинейное программирование принято подразделять на:

- выпуклое программирование
- квадратичное программирование
- многоэкстремальные задачи

Выпуклое программирование (рис. 33) — целевая функция и область допустимых значений выпуклы, то есть, такие, что, если соединить любые два значения  $x_1$  и  $x_2$  из этой области, то отрезок будет лежать внутри нее.

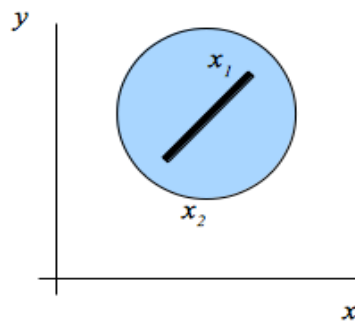


Рис. 33

Область невыпукла (рис. 34) — в области существуют такие пары точек, отрезок между которыми выходит за пределы области.

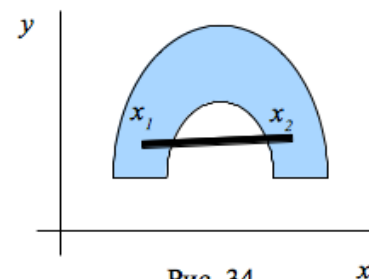


Рис. 34

Задачи выпуклого программирования всегда имеют один экстремум (одно решение), имеют эффективные методы решения.

Задачи невыпуклого программирования имеют несколько решений, они относятся к многоэкстремальным и эффективных решений не имеют.

Квадратичное программирование — отличительной особенностью этих задач является линейность ограничений и квадратичная зависимость целевой функции от переменных:

$$f = c_1x_1^2 + c_2x_2^2 + \dots + c_nx_n^2.$$

Отдельными классами задач математического программирования являются задачи целочисленного, параметрического и дробно-линейного программирования.

В задачах целочисленного программирования неизвестные могут принимать только целочисленные значения. Например, определение числа и состава работающих агрегатов на ГЭС и распределение нагрузки между ними (например, на ГЭС три агрегата мощностью по 50 МВт, еще два — по 100. Нагрузка — 200 МВт, вопрос: сколько агрегатов будут работать и какие именно? Ответов два — или два по 100, или один 100 и два по 50).

В задачах параметрического программирования целевая функция или функция, определяющая область возможных изменений переменных, либо и то, и другое, зависят от некоторых параметров реальных процессов.

В задачах дробно-линейного программирования целевая функция представляет собой отношение двух линейных функций, а функции, определяющие область возможных изменений переменных, являются линейными.

Динамическое программирование не выделяется по математической записи (линейное или нелинейное), а дает новый подход к решению задач. Термин относится скорее к

вычислительному методу, он основан на целенаправленном переборе возможных вариантов решений.

## Раздел 6

# Линейное программирование

Линейное программирование — это раздел математического программирования, предназначенный для решения и исследования оптимальных задач, в которых целевая функция и ограничения заданы линейными функциями. Если функция нелинейна, ее линеаризируют и решают, как задачу линейного программирования.

Задача линейного программирования формулируется следующим образом: требуется найти  $n$  переменных  $x_i$ , которые минимизируют (максимизируют) функцию

$$f = c_1x_1 + c_2x_2 + \dots + c_nx_n$$

при  $m$  ограничениях

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m \end{cases}$$

и условиях неотрицательности переменных

$$x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0.$$

Подлежащая минимизации (максимизации) функция  $f$  в математическом программировании называется **целевой функцией**, а отдельно в линейном программировании — еще и линейной формой.

### 6.1. Транспортная задача

Рассмотрим материал на примере транспортной задачи:

Уголь, добываемый в нескольких месторождениях, отправляется ряду потребителей. Известно, сколько в месяц добывается угля на каждом из месторождений, а также сколько вместе с этим потребляется на каждом из заводов. Известны расстояния и условия сообщения между месторождениями и потребителями, из чего можно подсчитать стоимость перевозки каждой тонны угля из любого месторождения в любой пункт потребления. Требуется при этих условиях организовать перевозку угля таким образом, чтобы суммарные затраты на нее были минимальны.

Сформулируем задачу математически:

Пусть для простоты заданы всего три месторождения:  $M_1$ ,  $M_2$  и  $M_3$ , месячная выработка каждого из которых составляет  $a_1$ ,  $a_2$  и  $a_3$ , соответственно. Этот уголь нужно доставить в пункты потребления  $\Pi_1$ ,  $\Pi_2$  и  $\Pi_3$  с ежемесячными потребностями  $b_1$ ,  $b_2$  и  $b_3$ . Пусть задана  $c_{ij}$  — стоимость перевозки тонны угля из любого  $M_i$  в любой  $\Pi_j$ . Дополнительно будем считать, что излишков или недостатков угля нет:

$$a_1 + a_2 + a_3 = b_1 + b_2 + b_3.$$

Задача состоит в определении такого плана перевозок, при котором:

- точно удовлетворен спрос в угле в каждом из пунктов потребления;
- вывезен весь уголь из месторождений (что, однако неизбежно вытекает из предыдущего условия и положения о совпадении суммарного потребления и суммарной выработки);
- общая стоимость перевозок наименьшая.

Через  $X_{ij}$  обозначим количество угля, отправляемого из месторождения  $M_i$  в пункт  $\Pi_j$  (например, из  $M_1$  в  $\Pi_1$  перевозится  $X_{11}$  тонн угля, а из  $M_2$  в  $\Pi_3$  —  $X_{23}$  тонн).

Составим схему перевозок:

	в $\Pi_1$	в $\Pi_2$	в $\Pi_3$	всего отправлено
из $M_1$	$X_{11}$	$X_{12}$	$X_{13}$	$a_1$
из $M_2$	$X_{21}$	$X_{22}$	$X_{23}$	$a_2$
из $M_3$	$X_{31}$	$X_{32}$	$X_{33}$	$a_3$
всего вывезено	$b_1$	$b_2$	$b_3$	

Превратим схему перевозок в систему линейных ограничений:

$$\begin{cases} X_{11} + X_{21} + X_{31} = b_1 \\ X_{12} + X_{22} + X_{32} = b_2 \\ X_{13} + X_{23} + X_{33} = b_3 \\ X_{11} + X_{12} + X_{13} = a_1 \\ X_{21} + X_{22} + X_{23} = a_2 \\ X_{31} + X_{32} + X_{33} = a_3 \end{cases}$$

Опишем стоимость всех перевозок, учитывая, что стоимость перевозки одной тонны угля равна  $c_{ij}$ :

$$S = c_{11}X_{11} + c_{12}X_{12} + c_{13}X_{13} + c_{21}X_{21} + c_{22}X_{22} + c_{23}X_{23} + c_{31}X_{31} + c_{32}X_{32} + c_{33}X_{33}$$

В этом случае  $S$  — и есть та самая целевая функция, которую нужно минимизировать, то есть, найти такие неотрицательные и удовлетворяющие ограничениям значения всех переменных  $X_{ij}$ , при которых  $S$  будет минимальна.

Эта задача относится к разряду типовых и носит название «транспортной задачи». По ее схеме можно поставить и задачу о передаче электроэнергии.

## 6.2. Геометрическая интерпретация задачи линейного программирования

Рассмотрим геометрическую интерпретацию на числовом примере:

Определить  $\max \{2x_1 + 3x_2\}$  при ограничениях :

$$\begin{aligned} x_1 + x_2 &\leq 6 \\ x_1 + 2x_2 &\leq 8,5 \\ x_1 &\leq 4 \\ x_2 &\leq 3 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$



Система ограничений рассматриваемой задачи образует область допустимых решений, среди которых требуется выбрать такое, которое доставляет максимум целевой функции.

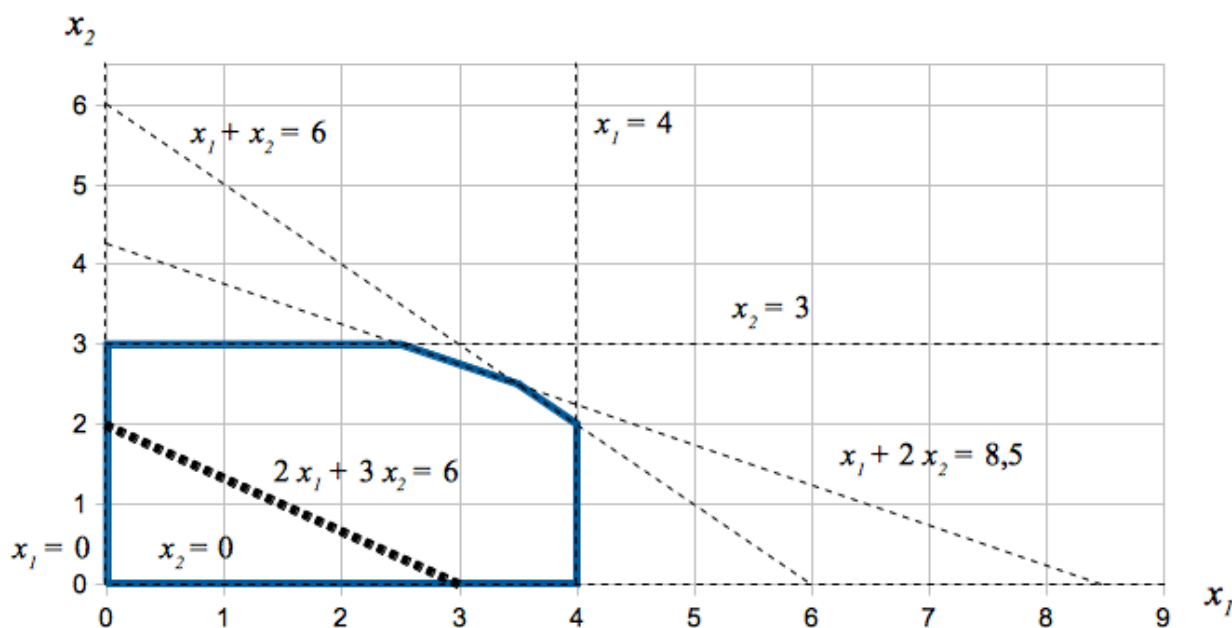


Рис. 35

Для построения области допустимых значений переменных возьмем границы условий-неравенств:

$$\begin{cases} x_1 + x_2 = 6 \\ x_1 + 2x_2 = 8,5 \\ x_1 = 4 \\ x_2 = 3 \\ x_1 = 0 \\ x_2 = 0, \end{cases}$$

Построим прямые этих функций на плоскости и, учитывая то, по какую сторону от прямой лежат области каждого из ограничений, выделим область их пересечения.

Возможные варианты при построении области допустимых значений:

- рассмотренный выше ограниченный выпуклый многоугольник — это основной случай, только в этом случае задача имеет решение
- неосновной случай — выпуклый неограниченный многоугольник
- другой неосновной случай — неравенства ограничений противоречат друг другу, и область допустимых значений переменных оказывается пустой

Вернемся к примеру: для определения максимума целевой функции придадим ей какое-либо произвольное значение, например:

$$2x_1 + 3x_2 = 6.$$

Далее, увеличивая это произвольное значение, мы передвигаем прямую этой функции вверх параллельно самой себе (задающие угол наклона коэффициенты при переменных остаются постоянными). Делаем это до тех пор, пока прямая не окажется в том положении, где на ней будет лежать только одна точка ограниченной области — эта точка и есть искомый максимум функции. В рассмотренном примере это точка пересечения прямых

$$x_1 + x_2 = 6 \text{ и } x_1 + 2x_2 = 8,5:$$

$$\begin{aligned} x_1 &= 3,5 \\ x_2 &= 2,5 \\ 2 \cdot 3,5 + 3 \cdot 2,5 &= 14,5 \\ 2x_1 + 3x_2 &= 14,5 \end{aligned}$$

### 6.3. Симплекс-метод

Для применения метода необходимо привести задачу линейного программирования к каноническому виду, то есть все ограничения записать в виде равенств, для чего в каждое из неравенств-ограничений вводится дополнительная переменная:

$\sum_{j=1}^n a_{ij} \cdot x_j \leq b_i$  — исходное неравенство, в которое добавляется переменная  $x_{n+1}$ :

$\sum_{j=1}^n a_{ij} \cdot x_j + x_{n+1} = b_i$  — канонический вид ограничения.

Рассмотрим метод на предыдущем примере:

Определить  $\max \{2x_1 + 3x_2\}$  при ограничениях:

$$\begin{cases} x_1 + x_2 \leq 6 \\ x_1 + 2x_2 \leq 8,5 \\ x_1 \leq 4 \\ x_2 \leq 3 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{cases}$$

Ограничения перепишем в каноническом виде:

$$\begin{cases} x_1 + x_2 + x_3 = 6 \\ x_1 + 2x_2 + x_4 = 8,5 \\ x_1 + x_5 = 4 \\ x_2 + x_6 = 3 \end{cases}$$

Выделим дополнительные переменные в так называемый базис, переписав уравнения следующим образом:

$$\begin{cases} x_3 = 6 - x_1 - x_2 \\ x_4 = 8,5 - x_1 - 2x_2 \\ x_5 = 4 - x_1 \\ x_6 = 3 - x_2 \end{cases}$$

При такой записи в левой части мы имеем так называемые базисные переменные, а в правой — свободные. Значения свободных переменных можно варьировать, при этом базисные переменные приобретают соответствующие значения. Допустим,  $x_1 = 0$ ,  $x_2 = 0$ . Тогда  $x_3 = 6$ ,  $x_4 = 8,5$ ,  $x_5 = 4$ ,  $x_6 = 3$ .

Решение системы ограничений, соответствующее нулевым значениям свободных переменных, называется базисным. Оно допустимо, так как все переменные удовлетворяют требованию неотрицательности. Целевая функция при этом приобретает нулевое значение. Следует отметить, что базисное решение соответствует вершине многогранника ограничений. Отсюда идея симплекс-метода заключается в последовательном переходе от одной вершины многогранника к другой в направлении улучшения целевой функции.

Ясно, что увеличение  $x_1$  или  $x_2$  приводит к возрастанию целевой функции.

- будем увеличивать только  $x_1$  до тех пор, пока одна из базисных переменных не окажется равной нулю. Ясно,  $x_1$  можно увеличивать до 4: при  $x_1 = 4$  базисная переменная  $x_5$  приобретает нулевое значение.
- меняем местами  $x_1$  и  $x_5$ , в результате получим новую систему свободных переменных  $x_2$  и  $x_5$ .
- выразим новые базисные переменные ( $x_1, x_3, x_4, x_6$ ) и целевую функцию через  $x_2$  и  $x_5$ , получим:

$$\begin{cases} \max \{8 - 2x_5 + 3x_2\} \\ x_3 = 2 + x_5 - x_2 \\ x_4 = 4,5 + x_5 - 2x_2 \\ x_1 = 4 - x_5 \\ x_6 = 3 - x_2 \end{cases}$$

Новое базисное решение равно:  $x_2 = 0, x_5 = 0, x_3 = 2, x_4 = 4,5, x_1 = 4, x_6 = 3$ .  
Значение целевой функции — 8.

Теперь будем увеличивать  $x_2$ . При  $x_2 = 2$  получим  $x_3 = 0$ . Меняем их местами, выражаем целевую функцию через  $x_3$  и  $x_5$ , получаем:

$$\begin{cases} \max \{14 + x_5 - 3x_3\} \\ x_2 = 2 + x_5 - x_3 \\ x_4 = 0,5 - x_5 + 2x_3 \\ x_1 = 4 - x_5 \\ x_6 = 1 - x_5 + x_3 \end{cases}$$

Базисное решение:  $x_3 = 0, x_5 = 0, x_2 = 2, x_4 = 0,5, x_1 = 4, x_6 = 1$ .  
Целевая функция — 14.

Далее, увеличивая значение  $x_5$ , меняя местами  $x_4$  и  $x_5$ , после преобразований получим:

$$\begin{cases} \max \{14,5 - x_3 - x_4\} \\ x_2 = 2,5 + x_3 - x_4 \\ x_5 = 0,5 + 2x_3 - x_4 \\ x_1 = 3,5 - 2x_3 + x_4 \\ x_6 = 0,5 - x_3 + x_4 \end{cases}$$

Базисное решение:  $x_3 = 0, x_4 = 0, x_2 = 2,5, x_5 = 0,5, x_1 = 3,5, x_6 = 0,5$ .  
Значение целевой функции равно 14,5.

Все коэффициенты при свободных переменных в целевой функции отрицательны, следовательно, невозможно продолжать увеличивать ее значение — любые положительные значения свободных переменных ее уменьшат. Значит, полученное решение является оптимальным.

Если проследить полученную процедуру перехода от одного базисного решения к другому на графике геометрической интерпретации задачи линейного программирования (рисунок 35), можно убедиться в том, что мы перебрали некоторое число вершин многогранника, причем каждый переход от одной вершины к следующей сопровождался увеличением целевой функции.

Для упрощения решения задачи применяют так называемые «симплекс-таблицы», в которых наглядно осуществляется вся изложенная выше система перехода от одного базисного решения к другому.

## Раздел 7

# Нелинейное программирование

Задачи нелинейного программирования подразделяются на:

- численные методы отыскания экстремума функций одной переменной (оптимизации функций одной переменной);
- численные методы отыскания экстремума функций нескольких переменных.

### 7.1. Численные методы оптимизации функций одной переменной

Задачи этого раздела решаются методами:

- нулевого порядка
  - метод золотого сечения
  - метод трех точек
  - метод Фибоначчи
  - адаптивный метод
  - метод равномерного распределения точек по отрезку
- первого порядка
- второго порядка.

#### 7.1.1. Методы первого порядка

Методы первого порядка более эффективны, нежели нулевого, и построены на определении корня производной целевой функции ( $f'(x) = 0$ ) на локальном отрезке  $[a, b]$ , когда на концах этого отрезка производная целевой функции принимает разные знаки:

$$f'(a) < 0, f'(b) > 0.$$

Необходимое условие применения методов первого порядка:

$$\exists! x \in [a, b]: (f'(x) = 0).$$

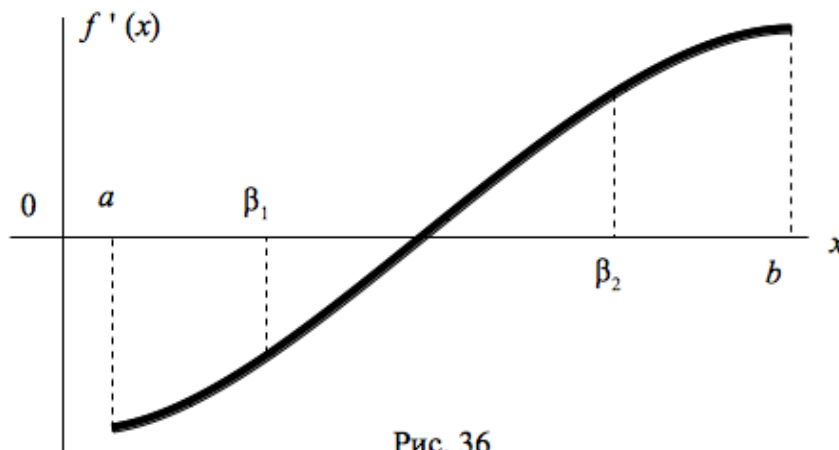
Достаточное условие:

$$\forall x \in [a, b]: (f''(x) > 0).$$

#### Условие ограничения отрезка локализации в методах первого порядка

При выполнении условий  $f'(a) < 0, f'(b) > 0, f''(x) > 0$ , если в точке  $\beta \in [a, b]$ :

- $f'(\beta) < 0$ , то искомый корень  $f'(x) = 0$  окажется на отрезке  $[\beta, b]$ ;
- $f'(\beta) > 0$ , то искомый корень  $f'(x) = 0$  окажется на отрезке  $[a, \beta]$ .



Примечание: в дальнейшем будут рассматриваться только вышеописанные условия, справедливые для отыскания именно минимума функции; в случае поиска максимума принцип остается тем же, только производная функции в начале отрезка положительна, а в конце отрицательна.

### Метод средней точки (метод биссекции)

Суть метода:

- найти производную  $f'(x)$  целевой функции;
- вычислить значения производной на концах отрезка  $[a, b]$ ;
- взять точку  $\beta$  посередине этого отрезка и вычислить значение производной в ней;
- выбрать новый отрезок  $[a^*, b^*]$ , где одна из точек  $a^*$  и  $b^*$  — это точка  $\beta$ , а другая — та из точек  $a$  и  $b$ , значение производной в которой отличается по знаку от значения производной в точке  $\beta$ , то есть выбрать такой отрезок, в котором имеет место переход значений производной через нуль;
- повторять пункты 2–4 до тех пор, пока не будет выполнено одно из условий точности:
  - $\Delta = |a^* - b^*| \leq \varepsilon_x$  — если в ходе поиска локальный отрезок в достаточной мере сузился (ширина его перестала превышать заданную «эпсилон икс»), то на этом поиск можно закончить;
  - $|f'(\beta)| \leq \varepsilon_y$  — если в ходе поиска была найдена точка, где значение производной целевой функции достаточно близко к нулю (не превышает заданного «эпсилон игрек»), то на этом тоже можно остановиться.

Рассмотрим более подробно ход решения задачи при применении метода биссекции.

Дано:

Целевая функция  $f(x)$  унимодальна (имеет только один экстремум) на отрезке  $[a, b]$ .

Заданы точности  $\varepsilon_x$  и  $\varepsilon_y$ .

1. Находим производную  $f'(x)$  целевой функции.
2. Вычисляем  $f'(a)$  и  $f'(b)$ , дабы убедиться, что они имеют разные знаки на концах отрезка.
3. Вычисляем  $\beta = \frac{b+a}{2}$ , вычисляем  $f'(\beta)$ .
4. Проверяем вертикальную точность,  $|f'(\beta)| \leq \varepsilon_y$  — если это условие не выполняется, продолжаем дальше.
5. Проверяем горизонтальную точность,  $|a^* - b^*| \leq \varepsilon_x$  — если это условие не выполняется, продолжаем дальше.
6. Выбираем новый отрезок  $[a^*, b^*]$ :

- если  $f'(\beta) < 0$ , то  $a^* = \beta$ ,  $b^* = b$ ;
  - если  $f'(\beta) > 0$ , то  $a^* = a$ ,  $b^* = \beta$ ;
1. Продолжаем с пункта 3, используя  $a^*$  и  $b^*$  вместо  $a$  и  $b$ .

....

п. Когда, наконец, выполняется какое-то из условий точности, заканчиваем поиск и принимаем  $f(\beta)$  в качестве искомого экстремума целевой функции.

Пример: найти методом биссекции экстремум функции  $y = x^3 - x + e^{-x}$  на отрезке  $[0, 1]$ , при этом точности по обеим осям заданы числом 0,1.

1. Найдем производную целевой функции:  $y' = 3x^2 - 1 - e^{-x}$ .
2. Вычислим:

$$y'(a) = 3 \cdot 0^2 - 1 - e^{-0} = -2$$

$$y'(b) = 3 \cdot 1^2 - 1 - e^{-1} = 1,63$$

3. Вычислим:

$$\beta = \frac{1 + 0}{2} = 0,5$$

$$y'(\beta) = 3 \cdot 0,5^2 - 1 - e^{-0,5} = -0,856$$

4. Для сокращения записи дальнейшие итерации представим таблицей:

$i$	$a^*$	$b^*$	$\beta$	$f'(\beta)$	$\Delta$	комментарии
1	0,000	1,000	0,500	-0,856	1,000	$f'(\beta) < 0$ $a^* = \beta$ $b^* = b$
2	0,500	1,000	0,750	0,215	0,500	$f'(\beta) > 0$ $a^* = a$ $b^* = \beta$
3	0,500	0,750	0,625	-0,363	0,250	$f'(\beta) < 0$ $a^* = \beta$ $b^* = b$
4	0,625	0,750	0,688	-0,084	0,125	$f'(\beta) < 0$ $a^* = \beta$ $b^* = b$  $f'(\beta) < \varepsilon_Y$ — можем закончить на этом, но, примера ради, сделаем еще одну итерацию
5	0,688	0,750	0,718	0,062	0,063	$f'(\beta) < \varepsilon_Y$ , $\Delta < \varepsilon_X$ — выполнены оба условия точности, на этом поиск экстремума закончим

Найденное экстремальное значение целевой функции:  $y(\beta) = y(0,718) = 0,1399$ .

## Метод хорд

От метода биссекции отличается только принципом выбора точки  $\beta$ .

В методе хорд производная  $f'(x)$  на отрезке  $[a, b]$  аппроксимируется прямой, проходящей через эти крайние точки, а в качестве точки  $\beta$  выбирается место пересечения этой хорды и оси абсцисс.

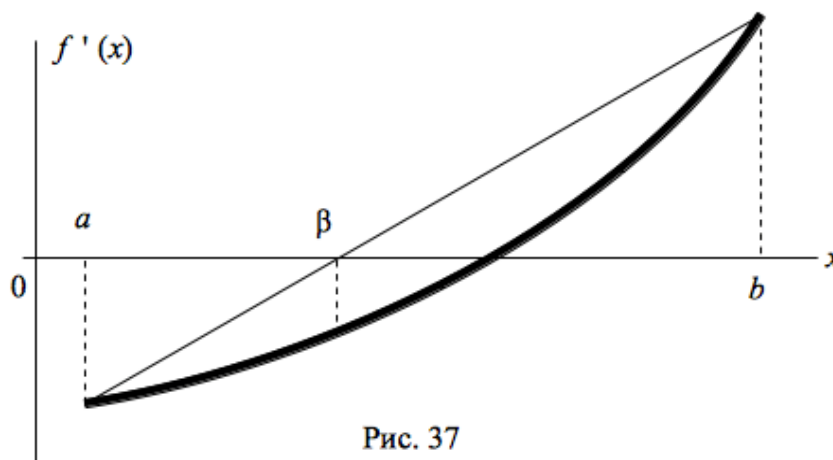


Рис. 37

Выведем выражение для нахождения точки  $\beta$ , чтобы использовать его так же, как  $\beta = \frac{b+a}{2}$  в методе средней точки. Для этого напишем уравнение, описывающее хорду, начав с общего вида уравнения линейной аппроксимации:

$$y = y_k + \frac{y_{k+1} - y_k}{x_{k+1} - x_k}(x - x_k).$$

Здесь подставим:

$$\begin{aligned}x_k &= a \\x_{k+1} &= b \\y_k &= f'(a) \\y_{k+1} &= f'(b)\end{aligned}$$

Тогда:

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

Найдем  $\beta$ , как тот  $x$ , при котором это выражение обращается в ноль (прямая пересекается с осью  $x$ ):

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

Мы получили готовое выражение, пригодное для выбора точки деления отрезка в методе хорд.

### **7.1.2. Методы второго порядка**

Методы второго порядка отличаются тем, что используют вторую производную целевой функции.

#### Метод Ньютона (метод касательных)

Заключается метод в следующем:

- находится производная целевой функции;
- выбирается начальное приближение вблизи предположительного корня уравнения  $f'(x) = 0$ ;

- в точке текущего приближения строится касательная к производной;
- точка пересечения касательной и оси абсцисс выбирается в качестве следующего приближения;
- пункты 3-4 повторяются, пока не будет достигнута заданная точность.

В этом методе нет нужды опираться на границы интервала и выбирать какую-то часть интервала в зависимости от изменения знака производной (она свой знак сохраняет от самого начального приближения и вплоть до достижения искомой нулевой точки), вся задача сводится к поитерационному перенахождению точки  $\beta$ .

В общем случае для устранения неопределенности в первой итерации полагается  $\beta = b$ .

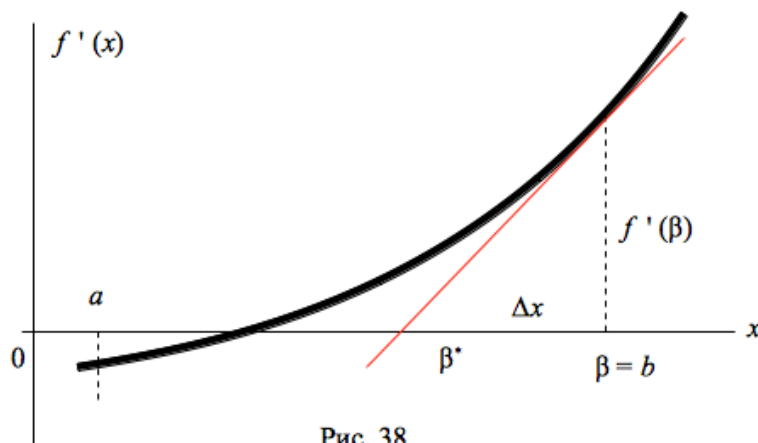


Рис. 38

Получим же формулу для вычисления  $\beta^*$  на каждой следующей итерации. Тангенс угла наклона касательной:

$$\operatorname{tg} \alpha = f''(\beta) = \frac{\Delta y}{\Delta x} = \frac{f'(\beta)}{\beta - \beta^*}$$

Отсюда:

$$\beta^* = \beta - \frac{f'(\beta)}{f''(\beta)}.$$

В этом методе в качестве условия остановки может применяться как уже использованная ранее вертикальная точность,

$$f'(\beta^*) \leq \varepsilon_y,$$

так и горизонтальная точность немного другого вида:

$$|\beta - \beta^*| \leq \varepsilon_x.$$

## **7.2. Численные методы оптимизации функций нескольких переменных**

Численные методы оптимизации функций нескольких переменных так же, как и для функций одной переменной, делятся на:

- методы нулевого порядка (прямой поиск);
- методы первого порядка (градиентные методы);
- методы второго порядка (метод Ньютона).

### **7.2.1. Методы нулевого порядка**

Эти методы сводятся к построению траектории спуска, вдоль которой целевая функция убывает, и различаются по способу построения этой траектории.



## Метод покоординатного спуска

В основе метода лежит последовательная оптимизация  $f(x)$  по каждому из параметров. Рассмотрим метод на примере функции двух переменных.

Пусть  $X_0 = (x_1^{(0)}; x_2^{(0)})$  — начальная точка поиска, которой задаются.

Каждая итерация разбивается на количество шагов, равное количеству переменных — в данном случае на два.

На каждом шаге изменяются значения одной конкретной переменной, все остальные переменные на это время становятся фиксированными параметрами. Таким образом, задача сводится к оптимизации функции одной переменной.

На примере:

Итерация первая, шаг первый:

$$x_1 = \text{var}, x_2^{(0)} = \text{const.}$$

Оптимизируем функцию  $f(x_1, x_2^{(0)})$ , находим ее минимум в точке  $x_1^*$ .

Итерация первая, шаг второй:

$$x_1^* = \text{const}, x_2 = \text{var.}$$

Оптимизируем функцию  $f(x_1^*, x_2)$ , находим ее минимум в точке  $x_2^*$ .

Итерация первая, проверка точности:

Проверяем поставленные условия по точности для  $f(x_1^*, x_2^*)$ , они могут быть следующими:

- $\Delta x_i = |x_i^* - x_i^{(0)}| \leq \varepsilon_{xi}$  — по каждой отдельной координате

- $\Delta X = \sqrt{\sum (x_i^* - x_i^{(0)})^2} \leq \varepsilon_X$  — по общему расстоянию

- $f'(x_i^*) \leq \varepsilon_{yi}$  — по значению частной производной целевой функции по каждой отдельной координате

Итерация вторая (и любая последующая) производится, если нужная точность не достигнута, и состоит из тех же действий, что и первая, только в качестве начальной точки поиска выступает найденная в предыдущей итерации.

Важно отметить, что метод покоординатного спуска неприменим, если в целевую функцию входит произведение каких-либо двух переменных.

Пример: определить экстремум функции  $f(x_1, x_2) = 4(x_1 - 5)^2 + (x_2 - 6)^2$  с точностью  $\varepsilon_{x1} = \varepsilon_{x2} = \varepsilon_{y1} = \varepsilon_{y2} = 0,1$ .

Зададимся начальным приближением:

$$x_1^{(0)} = 8$$

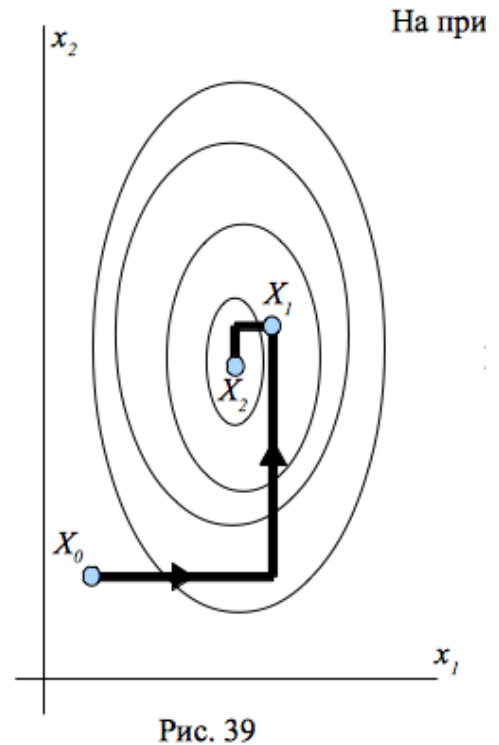
$$x_2^{(0)} = 8$$

$$f(X_0) = 4(8 - 5)^2 + (8 - 6)^2 = 40$$

Итерация первая, шаг первый:

$$x_1 = \text{var}, x_2^{(0)} = 8.$$

Найдем минимум функции  $f(x_1) = 4(x_1 - 5)^2 + (8 - 6)^2 = 4(x_1 - 5)^2 + 4$  с помощью метода Ньютона:



$$\begin{aligned}f'(x_1) &= 8(x_1 - 5) \\f''(x_1) &= 8\end{aligned}$$

В качестве начального приближения возьмем  $x_1^{(0)} = 8$ .

$$x_1^* = x_1^{(0)} - \frac{f'(x_1^{(0)})}{f''(x_1^{(0)})} = 8 - \frac{24}{8} = 5.$$

Проверим точность:

$$f'(x_1^*) = 8(5 - 5) = 0 < \varepsilon_{y1}.$$

Примечание: полезно знать, что если функция имеет квадратичную форму, то метод Ньютона находит ее экстремум с предельной точностью за одну итерацию.

Итерация первая, шаг второй:

$$x_1^* = 5, x_2 = var.$$

Найдем минимум функции  $f(x_2) = 4(5 - 5)^2 + (x_2 - 6)^2 = (x_2 - 6)^2$  с помощью метода Ньютона:

$$\begin{aligned}f'(x_2) &= 2(x_2 - 6) \\f''(x_2) &= 2\end{aligned}$$

В качестве начального приближения возьмем  $x_2^{(0)} = 8$ .

$$x_2^* = x_2^{(0)} - \frac{f'(x_2^{(0)})}{f''(x_2^{(0)})} = 8 - \frac{4}{2} = 6.$$

Проверим точность:

$$f'(x_2^*) = 2(6 - 6) = 0 < \varepsilon_{y2}.$$

Учитывая тот факт, что точности в каждом шаге были предельными, дальнейшие итерации производить не представляется возможным. Целевая функция имеет минимум, равный нулю, в точке (5; 6).

## 7.2.2. Методы первого порядка. Градиентные методы

### Понятие градиента

Если  $f(x_1, x_2, \dots, x_n)$  непрерывна и дифференцируема, то градиентом этой функции называется  $n$ -мерный вектор, элементами которого являются первые частные производные по каждой переменной:

$$\text{grad} f(x_1, x_2, \dots, x_n) = \left\{ \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right\}.$$

Если  $\text{grad} f(X_0) \neq 0$ , то он ортогонален изолинии  $f(x) = \text{const} = f(X_0)$  и направлен в

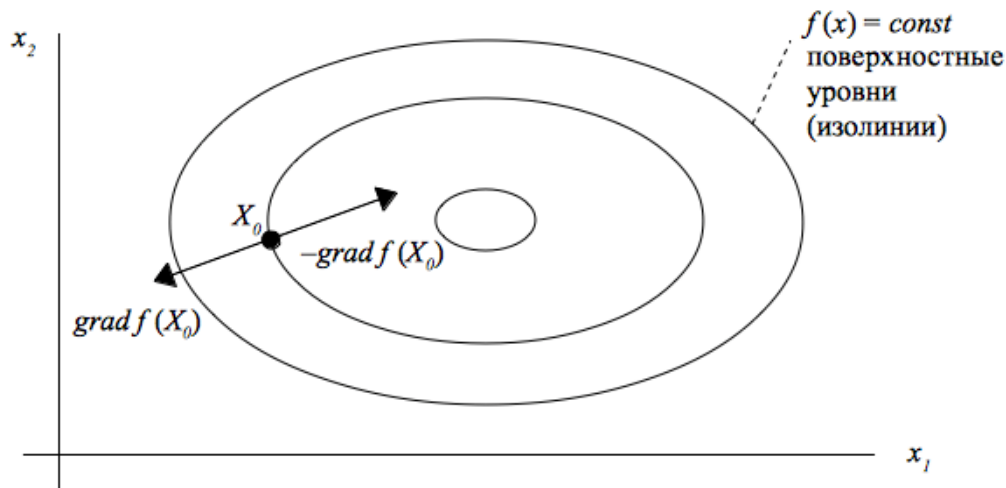


Рис. 40

сторону наискорейшего возрастания функции. Тогда антиградиент,  $-grad f(X_0)$ , направлен в сторону ее скорейшего убывания.

Абсолютная величина градиента равна геометрической сумме векторов, равных значениям частных производных и направленных по соответствующим координатным осям:

$$|grad f| = \sqrt{\left(\frac{\partial f}{\partial x_1}\right)^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 + \dots + \left(\frac{\partial f}{\partial x_n}\right)^2}.$$

### Градиентный метод

Рассматривается задача оптимизации непрерывной выпуклой функции:

$$f(X) = f(x_1, x_2, \dots, x_n).$$

Поиск оптимума начинается с некоторого исходного допустимого решения  $X_0$ . Если  $grad f(X_0) \neq 0$ , то это означает, что его можно улучшить, то есть, найти такую точку  $X^*$ , где  $|grad f(X^*)| < |grad f(X_0)|$ .

Алгоритм метода сводится к нахождению последовательности допустимых решений  $X_k$  таких, чтобы  $f(X_{k+1}) < f(X_k)$ .

Очередное решение задачи получается из уравнения:  $X_{k+1} = X_k \pm \Delta h \cdot grad f(X_k)$  (знак зависит от того, максимум ищем, +, или минимум, -). Очевидно, встает задача выбора шага  $\Delta h$ . Существуют различные способы этого выбора, это и определяет вариант градиентного метода.

#### 1. Градиентный метод с конечным шагом

Величина шага  $\Delta h$  задается постоянной в начале расчета. В таком случае длина смещения расчетной точки зависит только от значения градиента в этой точке (направление смещения в любом случае задается градиентом, как вектором):

$$\Delta X = X_{k+1} - X_k = \Delta h \cdot grad f(X_k).$$

Такой метод выбора шага отличается тем, что по мере приближения к экстремуму целевой функции ее частные производные уменьшаются, следовательно, уменьшается градиент, а с ним уменьшается и смещение расчетной точки.

Суть метода сводится к постоянному перенахождению следующей расчетной точки по формуле:

$$X_{k+1} = X_k \pm \Delta h \cdot grad f(X_k).$$

В качестве критерия окончания расчетов выбирается условие того, что длина градиента становится много меньше заданной величины точности:

$$\sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}\right)^2} \ll \varepsilon_y.$$

Особенность траектории движения расчетной точки в том, что каждый отрезок ее траектории перпендикулярен к той поверхности уровня (изолинии), от которой он начался.

Важный момент в реализации данного метода — выбор величины шага  $\Delta h$ . Если взять шаг слишком малым, то нахождение оптимума потребует долгих вычислений градиента в каждой из многочисленных точек. Если же взять его очень большим — возможен незатухающий процесс блуждания расчетной точки вокруг истинного оптимума (она будет постоянно его «перескакивать»), и это в особенности касается таких функций, у которых близкие к оптимуму поверхности уровней сильно вытянуты по одной из осей.

## 2. Метод наискорейшего спуска

Заключается в том, что на каждой итерации  $i$  вычисляется такое значение  $\Delta h_i$ , чтобы минимизировалась функция:

$$f(x_i - \Delta h_i \text{grad} f(x_i)).$$

Это — вложенная в каждую итерацию исходной задачи дополнительная задача одномерной оптимизации. Решение ее осуществляется любым подходящим методом на усмотрение инженера.

Чаще всего используют самый простой — метод дробления шага. В нем шаг  $\Delta h_i$  разбивается на  $m$  частей, вычисляются значения  $f\left(x_i - \frac{j}{m} \cdot \Delta h_i \text{grad} f(x_i)\right)$  для всех  $j$  от единицы до  $m$ , из них выбирается минимальное, и соответствующее ему  $j$  пускается в дальнейший расчет.

Метод наискорейшего спуска требует меньшего количества расчетов и имеет более быструю сходимость, нежели градиентный метод с конечным шагом. Однако, они разделяют общий недостаток: из-за того, что градиент всегда перпендикулярен поверхности уровня, градиентные методы работают медленно в случаях, когда поверхности уровня функции вытянуты по какой-то оси.

Пример: методом наискорейшего спуска оптимизировать функцию:

$$f(X) = 4(x_1 - 5)^2 + (x_2 - 6)^2$$

с точностью  $\varepsilon_y = 1$ .

$$\frac{\partial f}{\partial x_1} = 8x_1 - 40, \quad \frac{\partial f}{\partial x_2} = 2x_2 - 12.$$

Начальное приближение:  $X_0 = (8, 10)$ .

$$f(X_0) = 52.$$

$$X_{k+1} = X_k \pm \Delta h_i \cdot \text{grad} f(X_k).$$

Первая итерация:

$$x_1^{(1)} = x_1^{(0)} - \Delta h_1 \cdot \frac{\partial f}{\partial x_1^{(0)}} = 8 - 24\Delta h_1$$

$$x_2^{(1)} = x_2^{(0)} - \Delta h_1 \cdot \frac{\partial f}{\partial x_2^{(0)}} = 10 - 8\Delta h_1$$

Нужно подобрать такое  $\Delta h_1$ , чтобы  $f(x_1^{(1)}, x_2^{(1)})$  было минимальным. Подставим их выражения в целевую функцию, и минимизируем ее по переменной  $\Delta h_1$ :

$$f(\Delta h_1) = 4(8 - 24\Delta h_1 - 5)^2 + (10 - 8\Delta h_1 - 6)^2 = 2304(\Delta h_1 - 0,125)^2 + 64(\Delta h_1 - 0,5)^2$$

$$f'(\Delta h_1) = 4608(\Delta h_1 - 0,125) + 128(\Delta h_1 - 0,5) = 4736\Delta h_1 - 640$$

$$f'(\Delta h_1) = 0: \Delta h_1 = 0,135$$

$$x_1^{(1)} = 8 - 24 \cdot 0,135 = 4,76$$

$$x_2^{(1)} = 10 - 8 \cdot 0,135 = 8,92$$

Проверка точности:

$$\sqrt{\sum_1^2 \left( \frac{\partial f}{\partial x_i^{(1)}} \right)^2} = \sqrt{(8 \cdot 4,76 - 40)^2 + (2 \cdot 8,92 - 12)^2} = 6,148 > \varepsilon_y, \text{ продолжаем}$$

расчет.

Вторая итерация:

$$x_1^{(2)} = x_1^{(1)} - \Delta h_2 \cdot \frac{\partial f}{\partial x_1^{(1)}} = 4,76 + 1,92\Delta h_2$$

$$x_2^{(2)} = x_2^{(1)} - \Delta h_2 \cdot \frac{\partial f}{\partial x_2^{(1)}} = 8,92 - 5,84\Delta h_2$$

$$\begin{aligned} f(\Delta h_2) &= 4(4,76 + 1,92\Delta h_2 - 5)^2 + (8,92 - 5,84\Delta h_2 - 6)^2 = \\ &= 14,746(\Delta h_2 - 0,125)^2 + 34,106(\Delta h_2 - 0,5)^2 \end{aligned}$$

$$f'(\Delta h_2) = 0: \Delta h_2 = 0,387$$

$$x_1^{(2)} = 5,503$$

$$x_2^{(2)} = 6,660$$

$$\sqrt{(8 \cdot 5,503 - 40)^2 + (2 \cdot 6,66 - 12)^2} = 4,235 > \varepsilon_y, \text{ продолжаем расчет.}$$

Третья итерация:

$$x_1^{(3)} = x_1^{(2)} - \Delta h_3 \cdot \frac{\partial f}{\partial x_1^{(2)}} = 5,503 - 4,024\Delta h_3$$

$$x_2^{(3)} = x_2^{(2)} - \Delta h_3 \cdot \frac{\partial f}{\partial x_2^{(2)}} = 6,66 - 1,32\Delta h_3$$

$$\begin{aligned} f(\Delta h_3) &= 4(5,503 - 4,024\Delta h_3 - 5)^2 + (6,66 - 1,32\Delta h_3 - 6)^2 = \\ &= 64,77(\Delta h_3 - 0,125)^2 + 1,74(\Delta h_3 - 0,5)^2 \end{aligned}$$

$$f'(\Delta h_3) = 0: \Delta h_3 = 0,135$$

$$x_1^{(3)} = 4,960$$

$$x_2^{(3)} = 6,482$$

$$\sqrt{(8 \cdot 4,96 - 40)^2 + (2 \cdot 6,482 - 12)^2} = 1,016 > \varepsilon_y, \text{ продолжаем расчет.}$$

Четвертая итерация:

$$x_1^{(4)} = x_1^{(3)} - \Delta h_4 \cdot \frac{\partial f}{\partial x_1^{(3)}} = 4,96 + 0,32\Delta h_4$$

$$x_2^{(4)} = x_2^{(3)} - \Delta h_4 \cdot \frac{\partial f}{\partial x_2^{(3)}} = 6,482 - 0,964\Delta h_4$$

$$\begin{aligned} f(\Delta h_4) &= 4(4,96 + 0,32\Delta h_4 - 5)^2 + (6,482 - 0,964\Delta h_4 - 6)^2 \\ &= 0,410(\Delta h_4 - 0,125)^2 + 0,929(\Delta h_4 - 0,5)^2 \\ f'(\Delta h_4) &= 0,82(\Delta h_4 - 0,125) + 1,858(\Delta h_4 - 0,5) = 2,678\Delta h_4 - 1,032 \end{aligned}$$

$$f'(\Delta h_4) = 0: \Delta h_4 = 0,385$$

$$x_1^{(4)} = 5,083$$

$$x_2^{(4)} = 6,110$$

$$\sqrt{(8 \cdot 5,083 - 40)^2 + (2 \cdot 6,11 - 12)^2} = 0,699 < \varepsilon_y, \text{ расчет закончен.}$$

Оптимум целевой функции:

$$f(5,083, 6,11) = 4(5,083 - 5)^2 + (6,11 - 6)^2 = 0,040.$$

### 7.2.3. Методы второго порядка

#### Метод Ньютона

Оптимизация непрерывной выпуклой функции нескольких переменных  $f(X) = f(x_1, x_2, \dots, x_n)$  методом Ньютона сводится к вычислению последовательности допустимых решений  $X_k$  таких, чтобы  $f(X_{k+1}) < f(X_k)$ , причем каждое  $X_{k+1}$  получается из выражения:

$$X_{k+1} = X_k - \Delta h \cdot \frac{f'(X_k)}{f''(X_k)}.$$

В качестве критерия окончания расчетов применяем уже использованное ранее условие:

$$\sqrt{\sum_{i=1}^n \left( \frac{\partial f}{\partial x_i} \right)^2} \ll \varepsilon_y.$$

Пример: определить экстремум функции  $f = 4(x_1 - 5)^2 + (x_2 - 6)^2$  с точностью  $\varepsilon_y = 0,1$

Определим производные функции:

$$\frac{\partial f}{\partial x_1} = 8x_1 - 40, \frac{\partial f}{\partial x_2} = 2x_2 - 12$$

$$\frac{\partial^2 f}{\partial x_1^2} = 8, \frac{\partial^2 f}{\partial x_2^2} = 2$$

Зададимся начальным приближением:

$$X_0 = (8, 10)$$

Итерация первая:

$$\frac{\partial f}{\partial x_1^{(0)}} = 8 \cdot 8 - 40 = 24, \frac{\partial f}{\partial x_2^{(0)}} = 2 \cdot 10 - 12 = 8$$

$$\frac{\partial^2 f}{\partial x_1^{(0)2}} = 8, \frac{\partial^2 f}{\partial x_2^{(0)2}} = 2$$

$$x_1^{(1)} = x_1^{(0)} - \Delta h_1 \cdot \frac{\frac{\partial f}{\partial x_1^{(0)}}}{\frac{\partial^2 f}{\partial x_1^{(0)2}}} = 8 - \frac{24}{8} \Delta h_1 = 8 - 3\Delta h_1$$

$$x_2^{(1)} = x_2^{(0)} - \Delta h_1 \cdot \frac{\frac{\partial f}{\partial x_2^{(0)}}}{\frac{\partial^2 f}{\partial x_2^{(0)2}}} = 10 - \frac{12}{2} \Delta h_1 = 10 - 6\Delta h_1$$

$$f(X_1) = 4(8 - 3\Delta h_1 - 5)^2 + (10 - 6\Delta h_1 - 6)^2 = 36(\Delta h_1 - 1)^2 + 36(\Delta h_1 - \frac{2}{3})^2$$

$$f'(X_1) = 72(\Delta h_1 - 1 + \Delta h_1 - \frac{2}{3})$$

$$\Delta h_1 = \frac{5}{6}: \quad x_1^{(1)} = 5,5, x_2^{(1)} = 5, f(X_1) = 2$$

$$\sqrt{\left(\frac{\partial f}{\partial x_1^{(1)}}\right)^2 + \left(\frac{\partial f}{\partial x_2^{(1)}}\right)^2} = \sqrt{(8 \cdot 5,5 - 40)^2 + (2 \cdot 5 - 12)^2} = 4,47 > \varepsilon_y$$

Итерация вторая:

$$\frac{\partial f}{\partial x_1^{(1)}} = 8 \cdot 5,5 - 40 = 4, \frac{\partial f}{\partial x_2^{(1)}} = 2 \cdot 5 - 12 = -2$$

$$\frac{\partial^2 f}{\partial x_1^{(1)2}} = 8, \frac{\partial^2 f}{\partial x_2^{(1)2}} = 2$$

$$x_1^{(2)} = x_1^{(1)} - \Delta h_2 \cdot \frac{\frac{\partial f}{\partial x_1^{(1)}}}{\frac{\partial^2 f}{\partial x_1^{(1)2}}} = 5,5 - \frac{4}{8} \Delta h_2 = 5,5 - 0,5\Delta h_2$$

$$x_2^{(2)} = x_2^{(1)} - \Delta h_2 \cdot \frac{\frac{\partial f}{\partial x_2^{(1)}}}{\frac{\partial^2 f}{\partial x_2^{(1)2}}} = 5 + \frac{2}{2} \Delta h_2 = 5 + \Delta h_2$$

$$f(X_2) = 4(5,5 - 0,5\Delta h_2 - 5)^2 + (5 + \Delta h_2 - 6)^2 = (\Delta h_2 - 1)^2 + (\Delta h_2 - 1)^2$$

$$f'(X_2) = 4(\Delta h_2 - 1)$$

$$\Delta h_2 = 1: \quad x_1^{(2)} = 5, x_2^{(2)} = 6, f(X_2) = 0$$

$$\sqrt{\left(\frac{\partial f}{\partial x_1^{(2)}}\right)^2 + \left(\frac{\partial f}{\partial x_2^{(2)}}\right)^2} = \sqrt{(8 \cdot 5 - 40)^2 + (2 \cdot 6 - 12)^2} = 0 \ll \varepsilon_y \text{ — поиск закончен.}$$

$$\text{Минимум } f: f(5,6) = 4(5 - 5)^2 + (6 - 6)^2 = 0.$$

