

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ  
ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

---

МОСКОВСКИЙ ЭНЕРГЕТИЧЕСКИЙ ИНСТИТУТ  
(ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ)

---

А.А. Белогловский

**КОНСПЕКТ ЛЕКЦИЙ ПО КУРСУ «МАТЕМАТИЧЕСКОЕ  
МОДЕЛИРОВАНИЕ»  
ЧАСТЬ I: ОСНОВЫ ЧИСЛЕННЫХ МЕТОДОВ**

УЧЕБНОЕ ПОСОБИЕ  
по курсу  
«Математическое моделирование»  
для студентов, обучающихся по специальности  
«Высоковольтные электроэнергетика и электротехника»

## **ВВЕДЕНИЕ**

Одной из важных составляющих частей работы специалистов в области высоковольтной электроэнергетики и электротехники является изучение электрофизических процессов, протекающих в установках высокого напряжения. Одним из инструментов такого исследования является математическое моделирование физических процессов. В сочетании с экспериментальным исследованием оно дает возможность более детально изучить протекающие в оборудовании процессы и, тем самым, создает предпосылки создания более эффективных высоковольтных устройств. Поэтому курс «Математическое моделирование» является важной составной частью подготовки инженеров по специальности «Высоковольтные электроэнергетика и электротехника».

Благодаря увеличению производительности современной вычислительной техники и повышению ее доступности для студентов на кафедре ТЭВН МЭИ (ТУ), открывается реальная возможность не только освоения теории математического моделирования, но и выработки навыков применения изучаемых методов к решению некоторых практически значимых задач.

В данном учебном пособии представлен краткий конспект лекций по курсу «Математическое моделирование», посвященных изучению элементарных численных методов решения нелинейных уравнений и их систем, решения оптимизационных задач (задач математического программирования), задач интерполяции и сглаживания функций, вычисления определенных интегралов функции одной переменной, численного дифференцирования и приближенного решения обыкновенных дифференциальных уравнений. Изложение теоретического материала сопровождается примерами решения простых задач, иллюстрирующих применение изучаемых численных методов на практике.

Этот конспект будет дополнен подборкой учебных примеров решения характерных для техники высоких напряжений задач отыскания экстремумов функций, численного интегрирования и приближенного решения обыкновенных дифференциальных уравнений, а также набором задач для самостоятельного решения.

Материалы, приведенные в данном учебном пособии, подготовлены на основе описания упомянутых численных методов, представленного в книгах [1-12].

Автор благодарит ассистента кафедры ТЭВН МЭИ (ТУ) Черненко Л.Л. за полезные советы, позволившие улучшить содержание данного пособия, и студентку кафедры Гусарову Ю.Э. за помощь в его оформлении.

## 1. ЧИСЛЕННОЕ РЕШЕНИЕ УРАВНЕНИЙ И ИХ СИСТЕМ

### 1.1. Некоторые сведения из функционального анализа

Прежде всего, приведем некоторые сведения из функционального анализа, необходимые для понимания исследуемого материала данного параграфа.

Пространством называется совокупность объектов (чисел, функций, наборов чисел – матриц, векторов – и т.д.), между которыми установлены соотношения, аналогичные пространственным соотношениям между объектами обычного геометрического пространства (совпадение, величина, расстояние в определенной области и т.д.). Принадлежность элемента  $z$  пространству  $Q$  обозначается  $z \in Q$ .

Пусть в пространстве  $Q$  определены операции сложения элементов  $z_1$  и  $z_2$  и их умножения на числовые множители  $a_1$  и  $a_2$ , так что  $z_1 \cdot a_1 + z_2 \cdot a_2 = z_b$ .

Пространство называется линейным, если выполняются сочетательный, переместительный и распределительный законы сложения и умножения на числовые множители. В линейном пространстве имеется нулевой элемент  $\theta$ , такой, что  $z \cdot 0 = \theta$ , где  $0$  – нулевой числовой множитель. В дальнейшем мы будем иметь дело только с линейными пространствами и не будем оговаривать это специально.

Важнейшей характеристикой элемента пространства является норма – действительное неотрицательное число, которое ставится в соответствие элементу  $z$  (обозначается  $\|z\|$ ) и обладает следующими свойствами:

- $\|z\| = 0$  только при  $z = \theta$ ,
- $\|z_1 + z_2\| \leq \|z_1\| + \|z_2\|$ ;
- $\|\alpha \cdot z\| = |\alpha| \cdot \|z\|$ .

Нормой действительного числа является его модуль:  $\|z\| = |z|$ .

Расстоянием между двумя элементами пространства  $z_1$  и  $z_2$  называется действительное число  $\rho(z_1, z_2)$ , удовлетворяющее следующим условиям:

- $\rho(z_1, z_2) \geq 0$ , а условие равенства нулю выполняется при совпадении элементов  $z_1$  и  $z_2$ ;
- $\rho(z_1, z_2) = \rho(z_2, z_1)$ ;
- $\rho(z_1, z_2) \leq \rho(z_1, z_3) + \rho(z_2, z_3)$ .

Так в пространстве  $n$ -мерных векторов расстояние можно определить соотношением

$$\rho(z, y) = \sqrt{\sum_{i=1}^n (z_i - y_i)^2}.$$

Если пространство нормировано, то есть для его элементов определено понятие нормы, то можно принять  $\rho(z, y) = \|z - y\|$ , и, наоборот, в качестве нормы можно выбрать расстояние между любым элементом  $z$  и нулевым элементом  $\theta$ , т.е.  $\|z\| = \rho(z, \theta)$ .

Оператор – это правило, по которому каждому элементу  $z \in Q_1$ , ставится в соответствие элемент  $y \in Q_2$ . Оператор обозначается следующим образом:  $y = A(z)$  или  $y = A \cdot z$ .

Например, операцию интегрирования  $y = \int_a^b f(x) dx$  можно представить как действие оператора  $A = \int_a^b dx$ , отображающего функцию  $f(x)$  из функционального пространства на множество действительных чисел. Такие операторы называются функционалами.

Оператор  $A$  называется линейным, если  $A(\alpha_1 z_1 + \alpha_2 z_2) = \alpha_1 A(z_1) + \alpha_2 A(z_2)$ , где  $\alpha_1$  и  $\alpha_2$  – числовые множители.

Норма оператора – это наименьшее число  $M = \|A\|$ , при котором в данной области пространства выполняется соотношение  $\rho(A \cdot z_1, A \cdot z_2) \leq M \rho(z_1, z_2)$ . Устремляя  $z_2$  к  $z_1$  получаем частный случай нормы в точке, когда

$$\|A\| = M = \lim_{dz \rightarrow 0} \frac{\rho[A \cdot z_1, A \cdot (z_1 + dz)]}{\rho(z_1, z_1 + dz)}.$$

В частном случае оператора-функции  $y = f(x)$  имеем

$$\|f\| = \lim_{dz \rightarrow 0} \frac{|f(z_1 + dz) - f(z_1)|}{|z_1 + dz - z_1|} = \lim_{dz \rightarrow 0} \frac{|f(z_1 + dz) - f(z_1)|}{|dz|} = |f'|,$$

т.е. норма оператора-функции равна модулю ее производной.

Суперпозиция (произведение) операторов  $W = B \cdot y = B(A \cdot z) = B \cdot A(z)$ . В общем случае  $B \cdot A \neq A \cdot B$ . Для нормы произведения операторов должно выполняться дополнительное условие  $\|A \cdot B\| \leq \|A\| \cdot \|B\|$ .

Степень оператора  $A \cdot (A \cdot z) = A^2 \cdot z$ . Из сформулированного выше свойства нормы произведения операторов вытекает, что  $\|A^n\| \leq \|A\|^n$ .

## 1.2. Решение нелинейного уравнения методом простых итераций

Пусть дано операторное уравнение  $x=P \cdot x$ , имеющее точное решение  $x^*$ . Его численное решение организуется в виде итерационного процесса, в котором очередное приближение  $x_k$  вычисляется из предыдущего приближения  $x_{k-1}$  по формуле

$$x_k = P \cdot x_{k-1} = P \cdot (P \cdot x_{k-2}) = P^2 \cdot x_{k-2} = P^k \cdot x_0,$$

Где  $x_0$  – начальное приближение к решению.

Определим условие сходимости итераций. По определению нормы оператора

$$\rho(x_k, x^*) \leq \|P^k\| \rho(x_0, x^*).$$

Принимая  $\rho(x_1, x_2) = \|x_1 - x_2\|$ , получаем

$$\|x_1 - x_2\| \leq \|P\| \cdot \|x_0 - x^*\| \leq \|P\|^k \cdot \|x_0 - x^*\|.$$

Из полученного соотношения следует, что если  $\|P\| < 1$ , то  $\|x_0 - x^*\| \rightarrow 0$ , т.е. итерации сходятся.

Таким образом, достаточным условием сходимости итерационного процесса решения нелинейного уравнения методом простых итераций является выполнение соотношения  $\|P\| < 1$  в области, включающей все значения последовательных приближений  $x_k$ .

Необходимое условие сходимости итераций заключается в том, что  $\|P\| < 1$  в некоторой малой окрестности точки точного решения  $x^*$ .

В случае оператора-функции необходимо найти решение уравнения  $x=f(x)$ . Поскольку  $\|f\| = |f'|$ , то достаточное условие сходимости итерационного процесса  $x_k = f(x_{k-1})$  записывается в виде  $|f'| < 1$  в области, включающей все приближения  $x_k$ .

Полученное достаточное условие сходимости итераций соблюдается далеко не всегда. Например, вблизи точного решения  $x^*$  модуль производной  $|f'|$  может быть больше 1.

Один из способов улучшения сходимости итераций заключается в следующем. Обозначим  $F(x)=f(x)-x$ . Тогда справедливо уравнение  $\alpha F(x)=0$ , где  $\alpha$  – произвольный числовой множитель. Тогда исходное уравнение  $x=f(x)$  эквивалентно уравнению  $x=\alpha F(x)+x$ .

Для сходимости итераций по формуле  $x_k=\alpha F(x_{k-1})+x_{k-1}$  достаточно, чтобы  $|(\alpha F(x) + x)'| = |\alpha F'(x) + 1| < 1$ , или  $-2 < \alpha F'(x) < 0$ . Отсюда находим значение  $\alpha$ .

Описанный метод улучшения сходимости неудобен тем, что требуется предварительно оценить границы отрезка, в которых лежит корень. Этот недостаток устраняется в рамках метода Ньютона.

### 1.3. Решение нелинейных уравнений и их систем методом Ньютона

Для решения нелинейных уравнений и их систем очень часто используется метод Ньютона. Его итерационная формула основана на разложении функции  $F(x)$ , являющейся левой частью уравнения  $F(x)=0$  в ряд Тейлора в окрестностях приближенного значения  $x_{k-1}$ :

$$F(x) = F(x_{k-1}) + F'(x_{k-1})(x - x_{k-1}) + o((x - x_{k-1})^2) \approx F(x_{k-1}) + F'(x_{k-1})(x - x_{k-1}).$$

Приравняв нулю полученную линейную аппроксимацию функции  $F(x)$  и подставив в нее  $x=x_k$ , найдем, что

$$x_k = x_{k-1} - \frac{F(x_{k-1})}{F'(x_{k-1})}.$$

Полученную итерационную формулу можно также вывести из условия сходимости усовершенствованного метода простых итераций. Для этого нужно потребовать, чтобы вблизи корня  $x^*$  выполнялось бы условие  $|\alpha F'(x^*) + 1| = 0$ , откуда  $\alpha = -1/F'(x^*)$ . Заменяя производную в точке  $x^*$  производной в промежуточной точке  $x_{k-1}$ , получим итерационную формулу метода Ньютона.

Следует заметить, что метод Ньютона не гарантирует сходимости, если начальное приближение выбрано далеко от корня  $x^*$ .

Метод Ньютона обобщается на случай решения системы нелинейных уравнений

$$F_i(\bar{x}) = 0, \quad i = 1, 2, \dots, n,$$

где  $\bar{x} = [x_1, x_2, \dots, x_n]^T$  – вектор переменных размерности  $n$ . Тогда имеем итерационную формулу

$$\bar{x}_k = \bar{x}_{k-1} - [F'(\bar{x}_{k-1})]^{-1} \cdot \bar{F}(\bar{x}_{k-1}).$$

Здесь  $\bar{F}(\bar{x}) = [F_1(\bar{x}), F_2(\bar{x}), \dots, F_n(\bar{x})]^T$  – вектор функций левых частей системы уравнений, записанный в промежуточной точке  $\bar{x}_{k-1}$ , а  $[F'(\bar{x}_{k-1})]$  – матрица первых производных функции

$$[F'] = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \dots & \frac{\partial F_1}{\partial x_n} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \dots & \frac{\partial F_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial F_n}{\partial x_1} & \frac{\partial F_n}{\partial x_2} & \dots & \frac{\partial F_n}{\partial x_n} \end{bmatrix},$$

также записанная в промежуточной точке  $\bar{x}_{k-1}$ .

#### Пример 1.

Решить численно уравнение  $x - e^{-x} + 2 = 0$  с точностью  $\varepsilon=0,001$ .

Обозначим  $F(x) = x - e^{-x} + 2$ , тогда уравнение приобретает вид  $F(x) = 0$ . Вычислим первую производную функции  $F(x)$ :  $F'(x) = 1 + e^{-x}$ . Тогда итерационная формула метода Ньютона приобретает вид:

$$x_k = x_{k-1} - \frac{x_{k-1} - e^{-x_{k-1}} + 2}{1 + e^{-x_{k-1}}}, \quad k = 1, 2, \dots$$

В качестве начального приближения возьмем  $x_0 = 1$ . Итерации будем продолжать до тех пор, пока не будет выполнено условие

$$\delta x_k = \left| \frac{x_k - x_{k-1}}{x_k} \right| \leq \varepsilon, \quad k = 1, 2, \dots$$

где  $\varepsilon$  – предельное допустимое расхождение значений  $x_k$  между двумя последовательными итерациями, заданное в относительных единицах. Именно такое расхождение здесь и далее рассматривается в качестве точности численного расчета.

Результаты расчета сведены в табл. 1.

Таблица 1

Результаты решения уравнения  $x - e^{-x} + 2 = 0$  методом Ньютона

$k$	0	1	2	3	4	5
$x_k$	1	-0,92423	-0,51395	-0,44399	-0,44285	-0,44285
$\delta x_k$	—	2,08198	0,79829	0,15757	0,00257	$\approx 0$

Для сравнения в табл.2 приведем результаты расчетов по схеме метода простых итераций:

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

где  $f(x) = e^{-x} - 2$ . При этом вблизи корня  $|f'(x)| = |-e^{-x}| = e^{-x} > 1$ , и итерационный процесс расходится (табл. 2).

Таблица 2

Результаты решения уравнения  $x - e^{-x} + 2 = 0$  методом простых итераций.

$k$	0	1	2	3	4	5
$x_k$	1	-1,63212	3,11471	-1,95561	5,06823	-1,99371

#### 1.4. Список контрольных вопросов к 1-й главе

1. Что такое оператор? Что такое норма оператора вообще и оператора-функции в частности?

2. Сформулируйте метод простых итераций численного решения нелинейных уравнений. Предложите его графическую интерпретацию.

3. В чем заключается условие сходимости метода простых итераций? Укажите пути ее улучшения.
4. Поясните основные принципы численного решения нелинейных уравнений методом Ньютона. Предложите его графическую интерпретацию.
5. Выведите итерационную формулу метода Ньютона для случая решения одного нелинейного уравнения.
6. Запишите итерационную формулу метода Ньютона для случая решения системы нелинейных уравнений.
7. В чем заключаются достоинства и недостатки метода Ньютона?



## 2. ОСНОВЫ РЕШЕНИЯ НЕЛИНЕЙНЫХ ЗАДАЧ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

### 2.1. Постановка задачи математического программирования

В ходе создания энергетического оборудования часто приходится сталкиваться с необходимостью решения задач оптимизации. Например, следует выбрать геометрические размеры разрядного промежутка так, чтобы, с одной стороны, его пробивное напряжение было бы наибольшим, и, с другой стороны, масса и габариты проектируемого устройства укладывались бы в заданные рамки. Таким образом, эта задача содержит целевую функцию, оптимум которой следует определить, и систему равенств и неравенств, накладывающих ограничения на допустимые значения переменных. Эта задача называется задачей математического программирования и формулируется следующим образом:

- найти оптимум (например, минимум) целевой функции  $f(\bar{x})$

$$f(\bar{x}) \rightarrow \min \quad (1)$$

- при условии выполнения ограничений

$$g_i(\bar{x}) \geq 0, i = \overline{1, m}, \quad (2)$$

$$h_j(\bar{x}) = 0, j = \overline{1, p}, \quad (3)$$

где  $\bar{x} = [x_1, x_2, \dots, x_n]^T$  – вектор переменных задачи размерности  $n$ .

Поскольку в задачах математического программирования требуется найти максимальное или минимальное значение целевой функции (1) при условии выполнения ограничений (2) и (3), то эти задачи часто называют задачами условной оптимизации, а их решение – условным максимумом (минимумом) целевой функции  $f(\bar{x})$ .

### 2.2. Особенности поиска решений в задачах математического программирования

Решение задачи математического программирования (1)–(3) в общем виде представляет собой достаточно сложную проблему. Поэтому прежде чем перейти к изучению общего алгоритма ее решения, следует рассмотреть частные формулировки задачи с точки зрения возможности упрощения ее решения. Прежде всего, такими частными случаями являются задачи, включающие в себя ограничения только в виде равенств или только в виде неравенств.

### 2.2.1. Задача математического программирования с ограничениями в виде равенств

Рассмотрим задачу математического программирования, в которой есть ограничения только в виде равенств. Пусть целевая функция задачи является функцией двух переменных

$$z = f(\bar{x}) = f(x_1, x_2),$$

а ее аргументы связаны уравнением-ограничением

$$\varphi(x_1, x_2) = 0.$$

Если функции  $z=f(x_1, x_2)$  поставить в соответствие некоторую поверхность, то в данной задаче необходимо найти следующие точки (рис. 1):

1. Точки, принадлежащие линии пересечения этой поверхности с цилиндром, направляющая которого описывается уравнением-ограничением  $\varphi(x_1, x_2)$ , а образующая параллельна оси  $OZ$ .

2. Точки этой линии пересечения, в которых функция  $z=f(x_1, x_2)$  принимает экстремальные значения (точки  $A$  и  $B$  на рис. 1).

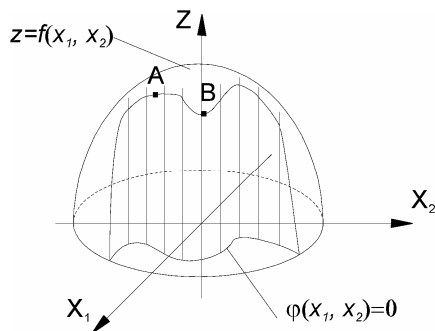


Рис. 1. Решение задачи математического программирования при ограничениях в виде равенств

Как видно из рис. 1, точки условного экстремума  $A$  и  $B$  не совпадают с максимумом или минимумом функции  $z=f(x_1, x_2)$ , т.е. с безусловным экстремумом функции  $f(x_1, x_2)$ .

Если из уравнения связи  $\varphi(x_1, x_2)=0$  выразить в явном виде одну переменную через другую, например  $x_2=\psi(x_1)$ , то функция  $z=f(x_1, x_2)=f[x_1, \psi(x_1)]$  становится функцией одной переменной  $x_1$ . Ее безусловный экстремум можно найти традиционным методом, приравняв нулю первую производную функции  $f[x_1, \psi(x_1)]$  по  $x_1$ .

Таким образом, безусловный экстремум функции  $f[x_1, \psi(x_1)]$  является условным экстремумом для функции  $f(x_1, x_2)$  при ограничении  $\varphi(x_1, x_2)=0$ .

Из сказанного следует, что решение задач математического программирования, включающих в себя только ограничения в виде равенств, может быть упрощено за счет понижения

размерности задачи. Для этого нужно выразить из ограничений-равенств одни переменные через другие.

### 2.2.2. Задача математического программирования с ограничениями в виде неравенств

Наличие в задаче математического программирования ограничений в виде неравенств еще более усложняет задачу. Рассмотрим возникающие при этом особенности решения на следующем примере:

$$f(\bar{x}) \rightarrow \max, \quad g_1(\bar{x}) = x_1 - x_2^2 \geq 0, \quad g_2(\bar{x}) = 1 - x_1^2 - x_2^2 \geq 0.$$

Область допустимых значений  $D$  переменных  $x_1$  и  $x_2$ , задаваемая системой ограничений, в этой задаче есть пересечение области, лежащей «внутри» параболы  $x_1 = x_2^2$ , с кругом единичного радиуса, уравнение окружности которого имеет вид  $x_1^2 + x_2^2 = 1$  (рис. 2).

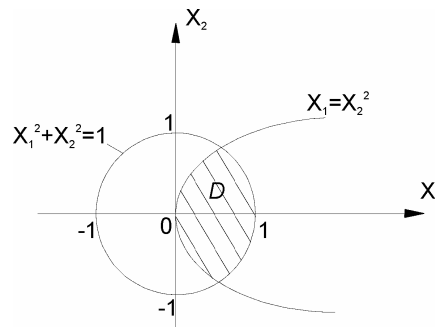


Рис. 2. Область допустимых значений переменных  $x_1$  и  $x_2$

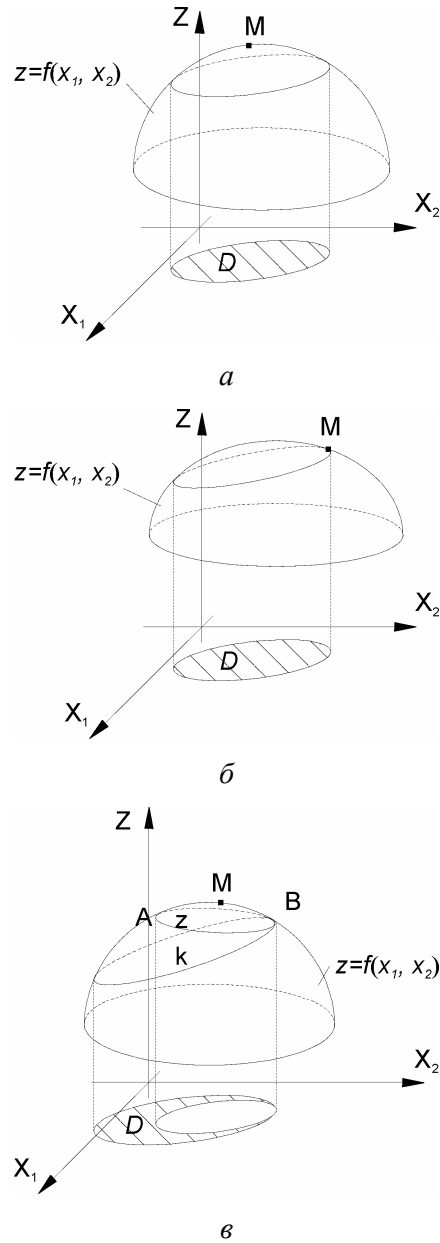
В общем случае пересечение цилиндра, направляющей которого является граница области допустимых значений переменных  $D$ , с поверхностью целевой функции  $z=f(x_1, x_2)$  может давать три варианта расположения безусловного экстремума функции  $f(x_1, x_2)$  относительно области  $D$ .

Эти варианты показаны на рис. 3 и сводятся к следующему:

а) Точка  $M$  безусловного экстремума функции  $z=f(x_1, x_2)$  лежит внутри области  $D$  допустимых значений переменных и является также точкой условного экстремума задачи.

б) Точка  $M$  лежит на границе области  $D$ , и, как и в предыдущем случае, в ней достигается условный экстремум целевой функции.

в) Точка  $M$  лежит вне области допустимых значений переменных, а целевая функция достигает наибольшего значения в одной из точек линии  $AB$ , причем не исключено, что оно будет достигаться одновременно в нескольких точках линии  $AB$  или на всей линии.



**Рис 3.** Положение точки  $M$  безусловного экстремума целевой функции  $f(x_1, x_2)$  относительно области допустимых значений переменных  $D$ :  $a$  – точка  $M$  лежит внутри области  $D$ ;  $б$  – точка  $M$  лежит на границе области  $D$ ;  $в$  – точка  $M$  лежит вне пределов области  $D$

Эти выводы справедливы и в случае трех и более переменных.

Из сказанного следует, что решение задач математического программирования с ограничениями в виде неравенств представляет собой достаточно сложную проблему. Алгоритм их решения должен включать в себя не только поиск точек экстремумов целевой функции, но и отбор среди них тех точек, которые лежат в области допустимых значений переменных. Затем должны быть определены максимальное и минимальное значения целевой функции на границах области допустимых значений. Лишь после этого максимальное или минимальное значение

функции во всей области может быть определено при помощи сравнения ее значений в отобранных точках экстремумов с максимальным или минимальным значением на границе.

### **2.3. Общий алгоритм решения задачи математического программирования**

подавляющее большинство задач математического программирования, встречающихся на практике, включает в себя ограничения не только в виде равенств, но и неравенств. Многие задачи включают в себя только ограничения-неравенства. Поэтому, как правило, даже если размерность задачи удастся сократить при помощи описанной в предыдущем параграфе процедуры подстановки, которая использует ограничения-равенства в качестве уравнений связи между переменными, приходится иметь дело с задачей математического программирования, содержащей ограничения-неравенства. Алгоритм их решения, основанный на отборе среди всех точек экстремумов целевой функции тех точек, которые лежат внутри области допустимых значений переменных, требует очень большого объема вычислений, и потому неприменим в случае функций многих переменных.

Поэтому на практике используется иной алгоритм решения задач математического программирования, основанный на составлении вспомогательной задачи безусловной оптимизации, целевая функция которой включает в себя как целевую функцию исходной задачи (1)–(3), так и ее ограничения.

Этот алгоритм состоит из трех этапов.

- Формулируется задача математического программирования (1)–(3), при необходимости производится уменьшение размерности задачи с помощью процедуры подстановки, которая использует ограничения-равенства в качестве уравнений связи между переменными. Здесь необходимо учитывать, целесообразна ли такая подстановка, поскольку в некоторых случаях она может неоправданно усложнить целевую функцию, а, следовательно, и усложнить последующую процедуру решения.

- Основываясь на исходной задаче математического программирования (1)–(3), которая является задачей безусловной оптимизации, составляется вспомогательная задача условной оптимизации (т.е. оптимизационная задача без ограничений), в целевую функцию которой включаются как целевая функция исходной задачи, так и ограничения исходной задачи.

- Полученная вспомогательная задача решается одним из аналитических или численных методов решения задач безусловной оптимизации.

Основные принципы составления вспомогательных оптимизационных задач и численного решения задач безусловной оптимизации рассматриваются в следующих параграфах главы.

## 2.4. Метод Лагранжа

Метод Лагранжа является одним из методов сведения задач математического программирования (условной оптимизации) к задачам безусловной оптимизации.

Начнем его изучение со случая задачи математического программирования, содержащей ограничения только в виде равенств.

Пусть требуется решить следующую задачу:

$$f(\bar{x}) \rightarrow \min, \quad h_j(\bar{x}) = 0, \quad j = \overline{1, p}. \quad (4)$$

По условию задачи составляют функцию Лагранжа

$$L(\bar{x}, \bar{\lambda}) = f(\bar{x}) - \sum_{j=1}^p \lambda_j h_j(\bar{x}), \quad (5)$$

где  $\lambda_j, j=1, 2, \dots, p$  – неизвестные множители, подлежащие определению (множители Лагранжа).

В результате исходная задача определения условного минимума (4), содержащая  $n$  неизвестных  $x_1, x_2, \dots, x_n$ , составляющих вектор  $\bar{x}$ , и  $p$  ограничений, сводится к эквивалентной ей задаче безусловной минимизации функции Лагранжа (5), зависящей от  $n+p$  переменных ( $n$  неизвестных  $x_1, x_2, \dots, x_n$  и  $p$  множителей Лагранжа  $\lambda_1, \lambda_2, \dots, \lambda_p$ ).

Точки, в которых возможен экстремум, находят как решение системы алгебраических уравнений, полученной приравниванием нулю частных производных от функции Лагранжа по искомым переменным  $x_1, x_2, \dots, x_n$  и  $\lambda_1, \lambda_2, \dots, \lambda_p$ :

$$\begin{cases} \frac{\partial L}{\partial x_k} = 0, \quad k = \overline{1, n}, \\ \frac{\partial L}{\partial \lambda_j} = 0, \quad j = \overline{1, p}. \end{cases} \quad (6)$$

Таким образом, метод Лагранжа сводит задачу отыскания условного экстремума функции  $f$  к эквивалентной ей задаче отыскания безусловного экстремума функции Лагранжа  $L$ .

В общем случае для задачи (1)–(3), содержащей ограничения в виде равенств и неравенств, функция Лагранжа вводится следующим образом. Ограничения-неравенства можно преобразовать в равенства, если вычесть из левой части дополнительные неотрицательные переменные  $s_i, i=1, 2, \dots, m$ . Условие их неотрицательности можно учесть автоматически, если в качестве таких дополнительных переменных ввести  $(s_i)^2, i=1, 2, \dots, m$ . Теперь в решаемой задаче имеются только ограничения-равенства и функция Лагранжа для нее записывается в виде

$$L(\bar{x}, \bar{s}, \bar{u}, \bar{\lambda}) = f(\bar{x}) - \sum_{i=1}^m u_i (g_i(\bar{x}) - s_i^2) - \sum_{j=1}^p \lambda_j h_j(\bar{x}), \quad (7)$$

где  $\lambda_j, j=1,2,\dots,p$  и  $u_i, i=1,2,\dots,m$  – множители Лагранжа, подлежащие определению наряду с координатами вектора  $\bar{x}$ . Множители  $\lambda_j$  для ограничений в виде равенств могут иметь любой знак, множители  $u_i$  для ограничений в виде неравенств должны быть неотрицательными.

Таким образом, при помощи метода Лагранжа задача (1)–(3) отыскания условного экстремума функции  $f$ , зависящей  $n$  от переменных, сводится к задаче отыскания безусловного экстремума функции  $L$ , которая зависит от  $n+m+p$  переменных. Этот экстремум в общем случае находится одним из численных методов безусловной оптимизации.

## 2.5. Основы численного решения задач безусловной оптимизации

После того, как исходная задача условной оптимизации (1)–(3) сведена к задаче безусловной оптимизации, последней проблемой, которую необходимо решить, является отыскание безусловного экстремума этой задачи. Эта проблема усугубляется тем, что, как следует из сказанного выше, ее целевая функция является функцией многих переменных. В большинстве из встречающихся на практике случаев сделать это можно только численно.

В задачах многомерной безусловной оптимизации требуется отыскать минимум некоторой целевой функции

$$f(\bar{x}) \rightarrow \min, \quad (8)$$

Дополнительные ограничения на значения переменных в задаче (8) отсутствуют.

Суть численных методов решения задачи безусловной оптимизации заключается в том, что строится последовательность векторов, которые вычисляются по следующему правилу:

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} + a_k \cdot \bar{h}^{(k)}.$$

Здесь  $k$  – номер итерации, вектор  $\bar{h}^{(k)}$  определяет направление движения на  $k+1$ -й итерации, параметр  $a_k$  определяет величину шага в этом направлении. Начальное приближение  $\bar{x}^{(0)}$  выбирается в зависимости от вида целевой функции.

В качестве примера численного метода решения задач безусловной оптимизации рассмотрим метод Ньютона 2-го порядка точности. Он основан на квадратичной аппроксимации минимизируемой функции. Для успешного применения метода Ньютона необходимо чтобы матрица вторых производных функции  $f(\bar{x})$  была невырожденной.

Разложим функцию  $f(\bar{x})$  в ряд Тейлора в окрестностях точки  $\bar{x}^{(k)}$ , ограничиваясь при этом членами ряда не выше второго:

$$f(\bar{x}) \approx f(\bar{x}^{(k)}) + f'(\bar{x}^{(k)}) \cdot (\bar{x} - \bar{x}^{(k)}) + \frac{1}{2} [f''(\bar{x}^{(k)})] \cdot (\bar{x} - \bar{x}^{(k)}) \cdot (\bar{x} - \bar{x}^{(k)}),$$

где  $f'(\bar{x})$  – вектор первых производных функции  $f$ ,  $[f''(\bar{x})]$  – матрица ее вторых производных.

Предположим, что  $\bar{x}^{(k+1)}$  – точка экстремума функции  $f(\bar{x})$ . Тогда, дифференцируя полученное разложение по  $\bar{x}$ , в точке  $\bar{x}^{(k+1)}$  получаем

$$f'(\bar{x}^{(k)}) + f''(\bar{x}^{(k)}) \cdot (\bar{x}^{(k+1)} - \bar{x}^{(k)}) = 0.$$

Из последнего выражения находим

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} - [f''(\bar{x}^{(k)})]^{-1} \cdot f'(\bar{x}^{(k)}). \quad (9)$$

### Пример 2.1.

С точностью  $\varepsilon = 0,001$  найти точку максимума функции  $f(x) = xe^{-x}$ .

Прежде всего, определим первую  $f'(x)$  и вторую производную  $f''(x)$  целевой функции  $f(x)$ :

$$f'(x) = (1-x)e^{-x}, \quad f''(x) = -e^{-x} - (1-x)e^{-x} = (x-2)e^{-x}.$$

Тогда итерационная формула (9) приобретает вид

$$x^{(k+1)} = x^{(k)} - \frac{f'(x^{(k+1)})}{f''(x^{(k+1)})} = x^{(k)} - \frac{1-x^{(k)}}{x^{(k)}-2} = x^{(k)} + \frac{x^{(k)}-1}{x^{(k)}-2}.$$

Введем критерий останова вычислений по этой формуле следующим образом:

$$\delta x^{(k+1)} = \left| \frac{x^{(k+1)} - x^{(k)}}{x^{(k+1)}} \right| \leq \varepsilon, \quad \delta f^{(k+1)} = \left| \frac{f(x^{(k+1)}) - f(x^{(k)})}{f(x^{(k+1)})} \right| \leq \varepsilon.$$

где  $\varepsilon$  – предельная невязка значений  $x$  между последовательными итерациями.

В качестве начального приближения выберем точку  $x^{(0)} = 0$ . Результаты расчета сведены в табл. 3.

Таблица 3

Результаты максимизации функции  $f(x) = xe^{-x}$  методом Ньютона

$k$	0	1	2	3	4	5
$x^{(k)}$	0	0,50000	0,83333	0,97619	0,99945	1,00000
$\delta x^{(k)}$	—	1,00000	0,40000	0,14634	0,02327	0,00055
$f(x^{(k)})$	0	0,30327	0,36217	0,36777	0,36788	0,36788
$\delta f^{(k)}$	—	1,00000	0,16263	0,01525	0,00029	$\approx 0$

Таким образом, найдена точка максимума целевой функции  $x^{(\max)}=1,0$ . Значение целевой функции в этой точке равно  $f^{(\max)} \approx 0,36788$ .

### Пример 2.2.

С точностью  $\varepsilon=0,001$  найти точку минимума функции  $f(\bar{x}) = 9 \cdot x_1^2 + x_2^2 - 18 \cdot x_1 + 6 \cdot x_2 + 18$ .



Прежде всего, определим вектор первых производных  $f'(\bar{x})$ , и матрицу вторых производных  $[f''(\bar{x})]$  целевой функции  $f(\bar{x})$ , а также обратную матрицу  $[f''(\bar{x})]^{-1}$ :

$$f'(\bar{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 18 \cdot x_1 - 18 \\ 2 \cdot x_2 + 6 \end{bmatrix}, \quad [f''(\bar{x})] = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 18 & 0 \\ 0 & 2 \end{bmatrix}, \quad [f''(\bar{x})]^{-1} = \begin{bmatrix} 1/18 & 0 \\ 0 & 1/2 \end{bmatrix},$$

и введем критерий останова вычислений по формуле (9):

$$\delta x_1^{(k+1)} = \left| \frac{x_1^{(k+1)} - x_1^{(k)}}{x_1^{(k+1)}} \right| \leq \varepsilon, \quad \delta x_2^{(k+1)} = \left| \frac{x_2^{(k+1)} - x_2^{(k)}}{x_2^{(k+1)}} \right| \leq \varepsilon, \quad \delta f^{(k+1)} = \left| \frac{f(x_1^{(k+1)}, x_2^{(k+1)}) - f(x_1^{(k)}, x_2^{(k)})}{f(x_1^{(k+1)}, x_2^{(k+1)})} \right| \leq \varepsilon.$$

В качестве начального приближения выберем вектор

$$\bar{x}^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

и выполним 1-ю итерацию по формуле (9):

$$\bar{x}^{(1)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1/18 & 0 \\ 0 & 1/2 \end{bmatrix} \times \begin{bmatrix} 18 \cdot 0 - 18 \\ 2 \cdot 0 + 6 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}.$$

Теперь проверим, выполнен ли критерий останова вычислений:

$$\delta x_1^{(1)} = \left| \frac{x_1^{(1)} - x_1^{(0)}}{x_1^{(1)}} \right| = \left| \frac{1 - 0}{1} \right| = 1 > \varepsilon, \quad \delta x_2^{(1)} = \left| \frac{x_2^{(1)} - x_2^{(0)}}{x_2^{(1)}} \right| = \left| \frac{-3 - 0}{-3} \right| = 1 > \varepsilon.$$

Критерий останова не выполнен ни по одной из переменных, и, следовательно, необходима следующая итерация:

$$\bar{x}^{(2)} = \begin{bmatrix} 1 \\ -3 \end{bmatrix} - \begin{bmatrix} 1/18 & 0 \\ 0 & 1/2 \end{bmatrix} \times \begin{bmatrix} 18 \cdot 1 - 18 \\ 2 \cdot (-3) + 6 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}.$$

На этот раз критерий останова вычислений выполнен:

$$\delta x_1^{(2)} = \left| \frac{x_1^{(2)} - x_1^{(1)}}{x_1^{(2)}} \right| = \left| \frac{1 - 1}{1} \right| = 0 < \varepsilon, \quad \delta x_2^{(2)} = \left| \frac{x_2^{(2)} - x_2^{(1)}}{x_2^{(2)}} \right| = \left| \frac{-3 + 3}{-3} \right| = 0 < \varepsilon.$$

Таким образом, найдена точка минимума целевой функции:

$$\bar{x}^{(\min)} = \begin{bmatrix} 1 \\ -3 \end{bmatrix},$$

а значение целевой функции в данной точке равно

$$f^{(\min)} = f(\bar{x}^{(\min)}) = 9 \cdot (1)^2 + (-3)^2 - 18 \cdot (1) + 6 \cdot (-3) + 18 = 0.$$

Метод Ньютона обладает хорошей сходимостью и позволяет получить решение за небольшое число итераций с приемлемой точностью при поиске минимума квадратичной или близкой к ней функции. Но при минимизации неквадратичных функций этот метод не всегда оказывается достаточно эффективным. Это связано с тем, что если точка начального прибли-

жения находится далеко от точного решения, то шаг по методу Ньютона может оказаться чрезмерно большим, и вычислительный процесс будет расходиться.

Для устранения этого недостатка может использоваться модифицированный метод Ньютона. В нем для улучшения сходимости в итерационную формулу (9) вводится коэффициент  $\alpha^{(k)} \leq 1$ :

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} - \alpha^{(k)} \cdot [f''(\bar{x}^{(k)})]^{-1} \cdot f'(\bar{x}^{(k)}).$$

Значение  $\alpha^{(k)}$  выбирается так, чтобы на каждом шаге итерационного процесса выполнялось условие

$$f(\bar{x}^{(k+1)}) \leq f(\bar{x}^{(k)}).$$

В простейшем случае этот коэффициент можно определить по формуле

$$\alpha^{(k)} = 1/2^m, \quad m = 0, 1, 2, \dots,$$

изменяя значение  $m$  до тех пор, пока не будет выполнено заданное условие.

## 2.6. Список контрольных вопросов ко 2-й главе

1. Сформулируйте постановку задачи математического программирования. В чем заключается различие между задачами условной и безусловной оптимизации?
2. В чем заключаются особенности решения задачи математического программирования с ограничениями в виде равенств и в виде неравенств?
3. В чем состоит общий подход к решению задач математического программирования?
4. В чем заключается назначение метода Лагранжа? Поясните его основные положения.
5. В каких случаях применяются численные методы решения задач безусловной оптимизации?
6. Сформулируйте общий подход к численному решению задачи безусловной оптимизации.
7. Как выполняется численное решение задачи безусловной оптимизации методом Ньютона?
8. В чем заключаются особенности сходимости итераций при использовании этого метода? Укажите пути ее улучшения.

### 3. ИНТЕРПОЛЯЦИЯ ФУНКЦИЙ ОДНОЙ ПЕРЕМЕННОЙ

#### 3.1. Понятие интерполяции функции

Пусть на отрезке  $[a, b]$  заданы  $n+1$  опорных (узловых) точек  $a=x_0 < x_1 < x_2 < \dots < x_n=b$ . Пусть, кроме того, заданы  $n+1$  действительных чисел  $y_i, i=0, 1, \dots, n$ , например, как значения некоторой функции  $f(x)$  в узловых точках  $x_i, i=0, 1, \dots, n$  (рис. 4). Тогда имеем следующую задачу интерполяции: необходимо найти многочлен  $I_n(x)$  степени не больше  $n$  такой, что  $I_n(x_i), i=0, 1, \dots, n$ .

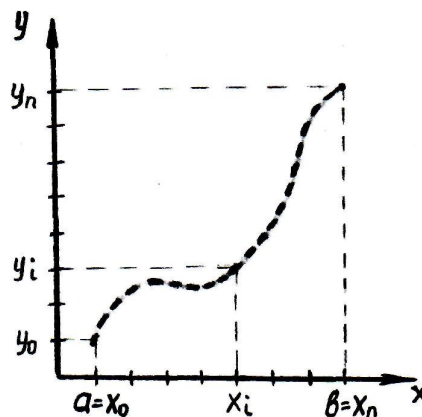


Рис. 4. К постановке задачи интерполяции функций одной переменной

Интерполяцию применяют главным образом тогда, когда известны только дискретные значения функции  $y=f(x)$ , и, чтобы вычислить другие ее значения между узловыми точками (интерполяция) или за отрезком узловых точек  $[a, b]$  (экстраполяция), ее приближают многочленом  $I_n(x)$ .

#### 3.2. Кусочно-линейная интерполяция

При использовании этого вида интерполяции на каждом отрезке  $[x_i, x_{i+1}]$  функция аппроксимируется отрезком прямой

$$f(x) = a_0 + a_1 x.$$

Тогда, исходя из условия неразрывности функции  $f(x)$  в узловых точках, для каждого из отрезков  $[x_i, x_{i+1}]$ , получаем систему из двух уравнений

$$f(x_i) = a_0 + a_1 x_i = y_i,$$

$$f(x_{i+1}) = a_0 + a_1 x_{i+1} = y_{i+1},$$

откуда находим

$$a_1 = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}, \quad a_0 = y_i - a_1 x_i = \frac{x_{i+1} y_i - x_i y_{i+1}}{x_{i+1} - x_i}.$$

Тогда аппроксимирующая функция  $f(x)$  на каждом из отрезков  $[x_i, x_{i+1}]$ ,  $i=0, 1, \dots, n-1$  описывается уравнением

$$f(x) = \frac{(x_{i+1} - x)y_i - (x - x_i)y_{i+1}}{x_{i+1} - x_i},$$

где  $x \in [x_i, x_{i+1}]$ ,  $i=0, 1, \dots, n-1$ .

Недостатком кусочно-линейной интерполяции является то, что хотя сама функция  $f(x)$  непрерывна в узловых точках, но ее первая производная в них испытывает разрывы. Для преодоления этого недостатка используется сплайн-интерполяция.

### 3.3. Сплайн-интерполяция

Происхождение сплайн-интерполяции и сам термин «сплайн» связывают с техническим приемом чертежников. При необходимости провести кривую через сеточный график на бумаге наносят точки  $(x_i, y_i)$ ,  $i=0, 1, \dots, n$ , около каждой из которых втыкают рядом друг с другом две булавки, и через образовавшийся «коридор» пропускают тонкую, гибкую и упругую стальную линейку («сплайн»). Форма, которую принимает эта линейка (вдоль нее и проводит требуемую линию), решает задачу гладкой интерполяции. Полученная кривая является дважды непрерывно дифференцируемой (т.е. ее первая и вторая производные непрерывны в узловых точках).

Теория упругости определяет форму  $y(x)$ , которую принимает «сплайн» требованием минимума энергии упругого состояния, т.е.  $y(x)$  определяется решением вариационной задачи  $\min_{y(\bullet)} \int_a^b [y''(x)]^2 dx$  при  $y(x_i) = y_i$ ,  $i=0, 1, \dots, n$ , где  $y(\bullet)$  – функция  $y(x)$ , рассматриваемая целиком как точка функционального пространства, а  $y(x)$  – в данном случае число, являющееся значением функции в точке  $x$ .

Вариационное исчисление дает следующий результат: внутри интервалов  $[x_i, x_{i+1}]$  функция  $y(x)$  удовлетворяет уравнению Эйлера  $d^4 y / dx^4 = 0$  и условиям непрерывности первых и вторых производных функции  $y(x)$  в узловых точках  $x_i$ ,  $i=0, 1, \dots, n$ . Из уравнения Эйлера следует, что  $y(x)$  является кубическим многочленом, причем на каждом интервале имеется свой кубический многочлен, и все они гладко сопрягаются друг с другом.

Таким образом, сплайн  $y(x)$  определяется формулой

$$y(x) = a_i x^3 + b_i x^2 + c_i x + d_i, \quad x \in [x_i, x_{i+1}], \quad i = 0, 1, \dots, n-1,$$

где коэффициенты  $a_i, b_i, c_i, d_i, i=0,1,\dots,n-1$  (их общее число равно  $4n$ ) определяются из следующих условий (уравнений):

– каждый кубический многочлен на концах своего интервала принимает заданные значения  $y_i$  и  $y_{i+1}$ , т.е.

$$y_i = a_i x_i^3 + b_i x_i^2 + c_i x_i + d_i, \quad i = 0, 1, \dots, n-1,$$

$$y_{i+1} = a_i x_{i+1}^3 + b_i x_{i+1}^2 + c_i x_{i+1} + d_i, \quad i = 0, 1, \dots, n-1;$$

– в каждой узловой точке сетки (кроме границ отрезка) имеем два условия, приравнивающие левые и правые значения первой и второй производных, т.е.

$$3a_{i-1}x_i^2 + 2b_{i-1}x_i + c_{i-1} = 3a_i x_i^2 + 2b_i x_i + c_i, \quad i = 1, 2, \dots, n-1,$$

$$6a_{i-1}x_i + 2b_{i-1} = 6a_i x_i + 2b_i, \quad i = 1, 2, \dots, n-1;$$

– на границах отрезка узловых точек имеем еще два уравнения, отражающие требования равенства нулю вторых производных сплайн-функции в точках  $x_0$  и  $x_n$ :

$$6a_0 x_0 + 2b_0 = 0,$$

$$6a_{n-1} x_n + 2b_{n-1} = 0.$$

Таким образом, имеем систему из  $4n$  линейных алгебраических уравнений относительно  $4n$  неизвестных значений коэффициентов  $a_i, b_i, c_i, d_i, i=0,1,\dots,n-1$ . Эта СЛАУ может быть решена одним из методов, разработанных для такого рода систем.

### 3.4. Сглаживание функции степенным многочленом с использованием метода наименьших квадратов

Пусть в плоскости  $xOy$  заданы  $N+1$  точек  $(x_i, y_i), i=0,1,\dots,N$ , причем  $x_i \neq x_j$  при  $i \neq j$  (рис. 5). Нужно найти функцию  $y=f(x, a_0, a_1, \dots, a_{n-1})$ , зависящую от  $x$  и от параметров  $a_0, a_1, \dots, a_{n-1}$  ( $n < N$ ), график которой как можно лучше приближается к заданным точкам. Согласно методу наименьших квадратов параметры  $a_0, a_1, \dots, a_{n-1}$  выбираются таким образом, чтобы величина

$$Q(a_0, a_1, \dots, a_{n-1}) = \sum_{i=0}^N [f(x_i, a_0, a_1, \dots, a_{n-1}) - y_i]^2$$

была наименьшей.

Если для величины  $Q$  воспользоваться теорией экстремумов, то при соответствующих предположениях о дифференцируемости функции  $f$  получаются следующие необходимые условия для определения параметров  $a_0, a_1, \dots, a_{n-1}$ :

$$\frac{\partial Q}{\partial a_j} = 0, \quad j = 0, 1, \dots, n-1.$$

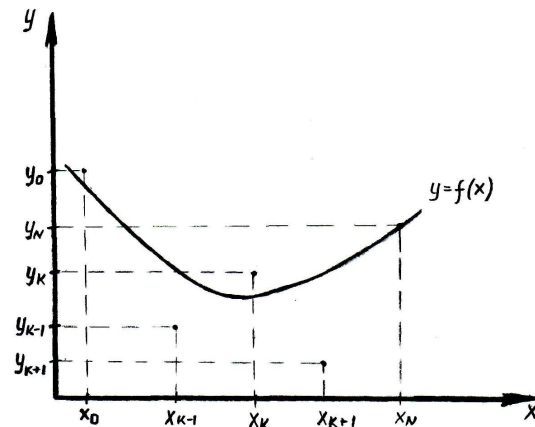


Рис. 5. К постановке задачи сглаживания функции методом наименьших квадратов

Если в качестве функции  $f$  выбрать многочлен относительно  $x$

$$f(x, a_0, a_1, \dots, a_{n-1}) = \sum_{j=0}^{n-1} a_j x^j,$$

то система уравнений относительно параметров  $a_0, a_1, \dots, a_{n-1}$  запишется в следующем виде:

$$\frac{\partial Q}{\partial a_j} = 2 \sum_{i=0}^N \left[ \left( \sum_{j=0}^{n-1} a_j x_i^j - y_i \right) x_i^j \right], \quad j = 0, 1, \dots, n-1,$$

т.е. получаем систему линейных уравнений относительно параметров  $a_j, j=0, 1, \dots, n-1$ .

### Пример 3.

Заданы следующие три точки:  $x_0=2, y_0=1,5; x_1=4, y_1=2,5; x_2=8, y_2=3,5$ . Необходимо определить параметры линейной аппроксимации зависимости  $y=f(x)$ , заданной тремя указанными точками.

Итак, пусть искомая линейная аппроксимация описывается формулой  $f(x, a_0, a_1) = a_0 + a_1 x$ . Тогда

$$Q(a_0, a_1) = \sum_{i=0}^N (a_0 + a_1 x_i - y_i)^2,$$

где, учитывая, что число заданных точек равно 3, а их нумерация начинается с 0,  $N=2$ . Неизвестные значения  $a_0$  и  $a_1$  можно определить из системы уравнений

$$\begin{cases} \frac{\partial Q}{\partial a_0} = 0, \\ \frac{\partial Q}{\partial a_1} = 0. \end{cases}$$

Здесь  $\frac{\partial Q}{\partial a_0} = 2 \cdot \sum_{i=0}^N (a_0 + a_1 x_i - y_i)$ ,  $\frac{\partial Q}{\partial a_1} = 2 \cdot \sum_{i=0}^N [(a_0 + a_1 x_i - y_i) x_i]$ .

Отсюда получаем следующую систему линейных алгебраических уравнений относительно неизвестных  $a_0$  и  $a_1$ :

$$\begin{cases} \sum_{i=0}^N (a_0 + a_1 x_i - y_i) = 0, \\ \sum_{i=0}^N [(a_0 + a_1 x_i - y_i) x_i] = 0. \end{cases}$$

Эта система преобразуется к виду

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

или, учитывая, что  $\sum_{i=0}^N a_0 = a_0 \sum_{i=0}^N 1 = (N+1)a_0$ ,

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

Тогда, введя обозначения  $A_{11} = N+1$ ,  $A_{12} = A_{21} = \sum_{i=0}^N x_i$ ,  $A_{22} = \sum_{i=0}^N x_i^2$ ,  $B_1 = \sum_{i=0}^N y_i$ ,  $B_2 = \sum_{i=0}^N (x_i y_i)$ , окончательно получаем

$$\begin{cases} A_{11}a_0 + A_{12}a_1 = B_1, \\ A_{21}a_0 + A_{22}a_1 = B_2. \end{cases}$$

Подставляя цифры, находим, что  $A_{11} = 2+1 = 3$ ,  $A_{12} = A_{21} = 2+4+8 = 14$ ,  $A_{22} = 2^2 + 4^2 + 8^2 = 84$ ,  $B_1 = 1,5 + 2,5 + 3,5 = 7,5$ ,  $B_2 = 2 \cdot 1,5 + 4 \cdot 2,5 + 8 \cdot 3,5 = 41$ , и таким образом, для определения параметров  $a_0$  и  $a_1$  имеем следующую систему алгебраических уравнений:

$$\begin{cases} 3a_0 + 14a_1 = 7,5, \\ 14a_0 + 84a_1 = 41, \end{cases}$$

откуда находим, что  $a_0=1,0$ ,  $a_1=0,3214$ .

Таким образом, заданная точками зависимость  $y = f(x)$  при помощи метода наименьших квадратов аппроксимируется линейной функцией  $y = 1 + 0,3214x$ .

Вычислим погрешности аппроксимации в заданных точках  $x_0$ ,  $x_1$  и  $x_2$ . Эти данные приведены в табл. 4. Значение  $\delta_i$  в табл. 4 вычислялось как

$$\delta_i = \left| \frac{y_i - f(x_i)}{y_i} \right| \cdot 100\%.$$

Из приведенных в табл. 4 данных видно, что, в данном случае погрешности аппроксимации не превысили 10%.

Таблица 4

Погрешности линейного сглаживания функции методом наименьших квадратов в Примере 3

$x_i$	2	4	8
$y_i$	1,5	2,5	3,5
$f(x_i)$	1,643	2,286	3,571
$\delta_i$	9,533	8,560	7,120

### 3.5. Список контрольных вопросов к 3-й главе

1. Сформулируйте постановку задачи интерполяции таблично заданной функции.
2. В чем заключается суть кусочно-линейной интерполяции таблично заданной функции?
3. В чем заключается суть сплайн-интерполяции таблично заданной функции?
4. Как построить кубический сплайн?
5. Сформулируйте постановку задачи нахождения приближающей (сглаживающей) функции для таблично заданной функции.
6. В чем заключается суть приближения таблично заданной функции по методу наименьших квадратов?
7. Какие функции могут быть использованы в качестве приближающих?
8. Как найти линейную приближающую функцию?
9. Как оценить качество полученной приближающей функции?
10. Чем принципиально различаются постановки задач приближения функции и ее интерполяции?



## 4. ВЫЧИСЛЕНИЕ ОПРЕДЕЛЕННЫХ ИНТЕГРАЛОВ

### 4.1. Понятие о приближенном вычислении определенного интеграла

Здесь речь пойдет об одной из самых распространенных операций – вычислении определенного интеграла  $\int_a^b f(x)dx$ . Самый общий подход к ее решению состоит в том, чтобы аппроксимировать функцию  $f(x)$  какой-либо другой функцией  $\tilde{f}(x)$ , для которой интеграл вычисляется аналитически. Тогда для решения задачи строим  $\tilde{f}(x)$  с оценкой погрешности

$$|f(x) - \tilde{f}(x)| \leq \varepsilon, \quad \forall x \in [a, b]$$

и полагаем приближенно

$$\int_a^b f(x)dx \approx \int_a^b \tilde{f}(x)dx$$

с очевидной оценкой погрешности  $(b-a)\varepsilon$ .

Введем на отрезке  $[a, b]$  сетку  $x_i, i=0, 1, \dots, N$ , где  $a=x_0 < x_1 < x_2 < \dots < x_N=b$ , и таблицу значений  $f_i, i=0, 1, \dots, N$ . Рассмотрим несколько простых вариантов построения функции  $\tilde{f}(x)$ , приводящих к широко распространенным формулам.

### 4.2. Формула трапеций

Функция  $\tilde{f}(x)$  строится как кусочно-линейная интерполяция значений  $f_i, i=0, 1, \dots, N$  на равномерной сетке с шагом  $\Delta x=(b-a)/N$ . Тогда

$$\int_a^b \tilde{f}(x)dx = \Delta x \sum_{i=0}^{N-1} \frac{1}{2}(f_i + f_{i+1}) = \Delta x \left( \frac{f_0}{2} + f_1 + f_2 + \dots + f_{N-1} + \frac{f_N}{2} \right) = \Delta x \left( \frac{f_0}{2} + \sum_{i=1}^{N-1} f_i + \frac{f_N}{2} \right).$$

Формулы такого рода  $(\sum c_i f_i)$  называют механическими квадратурами,  $c_i$  – коэффициентами (весами) квадратуры,  $x_i$  – ее узлами.

Точность формулы трапеций зависит от гладкости функции  $f$ . Если она на отрезке  $[a, b]$  имеет первую производную, ограниченную числом  $P_1$ , то  $|f(x) - \tilde{f}(x)| \leq P_1 \Delta x / 2$ , и погрешность формулы трапеций не превосходит  $P_1 \Delta x (b-a) / 2$ . Если  $f$  на  $[a, b]$  имеет вторую производную, ограниченную числом  $P_1$ , то погрешность формулы итераций не превосходит  $P_1 \Delta x^2 (b-a) / 2$ , поскольку  $|f(x) - \tilde{f}(x)| \leq P_1 \Delta x^2 / 2$ .

### 4.3. Формула Симпсона

Она также строится на основе равномерной сетки, содержащей, однако, четное число интервалов  $2N$ . Вычисляется таблица значений  $f_i$ ,  $i=0,1,\dots,N$ , и функция строится как кусочно-квадратичная интерполяция, т.е. на каждой паре интервалов  $(x_{2i}, x_{2i+1}, x_{2i+2})$  по значениям  $f_{2i}$ ,  $f_{2i+1}$ ,  $f_{2i+2}$  строится интерполяционный полином Лагранжа. Некоторые выкладки дают:

$$\int_{x_{2i}}^{x_{2i+2}} \tilde{f}(x) dx = \frac{1}{3} \Delta x (f_{2i} + f_{2i+1} + f_{2i+2}), \quad i = 0, 1, \dots, N-1.$$

Суммируя для  $i=0,1,\dots,N-1$  получаем:

$$\int_a^b \tilde{f}(x) dx = \frac{1}{3} \Delta x \sum_{i=0}^{N-1} (f_{2i} + 4f_{2i+1} + f_{2i+2}) = \frac{1}{3} \Delta x (f_0 + 4f_1 + 2f_2 + 4f_3 + \dots + 2f_{2N-2} + 4f_{2N-1} + f_{2N}).$$

Если функция  $f$  на отрезке  $[a,b]$  имеет третью производную ограниченную числом  $P_3$ , то  $|f(x) - \tilde{f}(x)| \leq P_3 \Delta x^3 / 3$  и погрешность формулы Симпсона не превосходит  $P_3 \Delta x^3 (b-a) / 3$ , где  $\Delta x = (b-a) / 2N$ .

### 4.4. Оценка точности численного интегрирования

Теоретические оценки погрешностей не очень популярны среди практиков. Если требуется вычислить интеграл с погрешностью  $\varepsilon$ , то мало кто сначала оценит третью производную функции  $f$  и вычислит шаг сетки  $\Delta x = \sqrt[3]{3\varepsilon / [(b-a)P_3]}$ . Дело в том, что сама эта оценка завышена и завышено значение константы  $P_3$ . Кроме того, само вычисление  $f'''$  может быть проблематичным, особенно если функция задана некоторым сложным образом.

Поступают иначе. Вычисляя интеграл с небольшим числом узлов ( $N=2\div 3$ ), получают его значение  $S_N$ ; вычисляя интеграл с удвоенным  $N$ , получают  $S_{2N}$ . Если модуль  $\delta S = |(S_{2N} - S_N) / S_{2N}| \leq \varepsilon$ , то задачу считают решенной, а в противном случае вычисляют  $S_{4N}$  и т.д. Нужно иметь в виду, что для гладких функций  $f$  часто интеграл вычисляется достаточно точно при малом числе узлов.

#### Пример 4.

Вычислим приближенно, по формуле Симпсона, интеграл

$$S = \int_0^2 e^x dx$$

с точностью  $\varepsilon=0,001$ . Данный интеграл берется аналитически:

$$S = e^x \Big|_0^2 = e^2 - 1 = 6,389056098.$$

В приведенной ниже таблице 4 представлены координаты узлов сетки  $x_i$ , значения функции  $f_i$  в узлах  $i=0,1,\dots,N$  и коэффициенты квадратур Симпсона  $c_i$  при  $N=1,2,4$ .

Таблица 4

Значения координат  $(x_i, f_i)$ ,  $i=0,1,\dots,N$  и коэффициентов квадратур Симпсона  $c_i$  в примере 4

$x_i$	$f_i$	Коэффициенты квадратур Симпсона $c_i$		
		$N=1$	$N=2$	$N=4$
0	1,0000000	1	1	1
0,25	1,2840254			4
0,5	1,6487213		4	2
0,75	2,1170000			4
1,0	2,7182818	4	2	2
1,25	3,4903425			4
1,5	4,4816891		4	2
1,75	5,7546026			4
2,0	7,3890561	1	1	1

По таблице легко вычисляются значения  $S_N = \frac{1}{3} \Delta x \left[ \sum_{i=0}^N c_i f_i \right]$ :

$$S_1 = \frac{1}{3} \cdot 1,00 \cdot (1 \cdot 1,0000000 + 4 \cdot 2,7182818 + 1 \cdot 7,3890561) = 6,4207278;$$

$$S_2 = \frac{1}{3} \cdot 0,5 \cdot (1 \cdot 1,0000000 + 4 \cdot 1,6487213 + 2 \cdot 2,7182818 + 4 \cdot 4,4816891 + 1 \cdot 7,3890561) = 6,3912102;$$

$$\delta S_2 = \left| \frac{S_2 - S_1}{S_2} \right| = \left| \frac{6,391210176 - 6,42072776}{6,391210176} \right| = 0,004618466 > \varepsilon;$$

$$S_4 = \frac{1}{3} \cdot 0,25 \cdot (1 \cdot 1,0000000 + 4 \cdot 1,2840254 + 2 \cdot 1,6487213 + 4 \cdot 2,1170000 + 2 \cdot 2,7182818 + 4 \cdot 3,4903425 + 2 \cdot 4,4816891 + 4 \cdot 5,7546026 + 1 \cdot 7,3890561) = 6,3891973;$$

$$\delta S_4 = \left| \frac{S_4 - S_2}{S_4} \right| = \left| \frac{6,3891973 - 6,391210176}{6,3891973} \right| = 0,000315044 < \varepsilon.$$

Относительные погрешности величин  $S_1$ ,  $S_2$  и  $S_4$  по отношению к точному значению  $S$  составляют 0,5, 0,03 и 0,002 %. Точность интегрирования, необходимая в инженерных расчетах, зависит от характера приложения, но зачастую даже погрешность 0,5 % является вполне приемлемой.

#### 4.5. Вычисление интегралов с особенностями

Рассмотрим методику вычисления интегралов с особенностями на простом примере:

$$\int_0^1 \frac{f(x)}{\sqrt{x}} dx,$$

где  $f(x)$  – произвольная гладкая функция. Здесь подынтегральная функция обращается в бесконечность при  $x = 0$ , но интеграл существует. Попытки его прямого вычисления по формуле

Симпсона или трапеций сразу же приведет к неудаче: первое же слагаемое  $f_0$  обращается в бесконечность.

Эта проблема решается при помощи приема регуляризации или «выделения особенности». Он состоит в том, что проделывается тождественное преобразование

$$\int_0^1 \frac{f(x)}{\sqrt{x}} dx = \int_0^1 \frac{f(x) - \varphi(x)}{\sqrt{x}} dx + \int_0^1 \frac{\varphi(x)}{\sqrt{x}} dx.$$

Функция  $\varphi(x)$  выбирается так, чтобы первый интеграл в правой части выражения не содержал бы особенности и при небольшом объеме вычислений достаточно точно определялся, например, по формуле Симпсона, а второй интеграл содержал бы особенность, но вычислялся аналитически.

В рассматриваемом примере цель будет достигнута, если в качестве  $\varphi(x)$  взять отрезок разложения  $f(x)$  в ряд Тейлора в точке особенности, т.е.  $x = 0$ :  $\varphi(x) \approx f(0) + f'(0) \cdot x$ . Это приведет к вычислению

$$\begin{aligned} \int_0^1 \frac{f(x)}{\sqrt{x}} dx &= \int_0^1 \frac{f(x) - f(0) - f'(0) \cdot x}{\sqrt{x}} dx + \int_0^1 \frac{f(0) + f'(0) \cdot x}{\sqrt{x}} dx = \\ &= \int_0^1 \frac{f(x) - f(0) - f'(0) \cdot x}{\sqrt{x}} dx + f(0) \int_0^1 \frac{dx}{\sqrt{x}} + f'(0) \int_0^1 \sqrt{x} dx. \end{aligned}$$

Если, например,  $f(x) = \cos(x)$ , то

$$\int_0^1 \frac{\cos(x)}{\sqrt{x}} dx = \int_0^1 \frac{\cos(x) - 1}{\sqrt{x}} dx + \int_0^1 \frac{dx}{\sqrt{x}}.$$

Второе слагаемое равно 2. Первое слагаемое вычислим по формуле Симпсона: сначала с шагом  $\Delta x = 0,5$ , что дает значение 1,807967, а затем с шагом  $\Delta x = 0,25$ , что дает 1,808850.

Второй способ приближенного вычисления интегралов с особенностями заключается в следующей замене интеграла:

$$\int_0^1 \frac{f(x)}{\sqrt{x}} dx \approx \int_{\varepsilon}^1 \frac{f(x)}{\sqrt{x}} dx, \quad \varepsilon > 0.$$

При достаточно малом значении  $\varepsilon$  он приближает нужное значение. Однако этот способ решения требует гораздо большего числа вычислений подынтегральной функции для достижения одной и той же точности.

#### 4.6. Вычисление интегралов от быстроосциллирующих функций

Пусть требуется вычислить интеграл

$$\int_0^{\pi} f(x)\sin(kx)dx$$

при большом значении  $k$ , например  $k = 100$ , где  $f(x)$  – гладкая функция. Вычисление этого интеграла по стандартной формуле Симпсона, конечно, возможно, но требует сетки с очень малым шагом: каждая волна должна быть описана некоторым числом узлов сетки, а волн много.

Для вычисления подобных интегралов используется следующий прием: гладкая функция  $f(x)$  аппроксимируется некоторой функцией  $\tilde{f}(x)$ , такой, чтобы интеграл от  $\tilde{f}(x)\sin(kx)$  вычислялся бы аналитически. Тогда задача сводится к тождественному преобразованию:

$$\int_0^{\pi} f(x)\sin(kx)dx = \int_0^{\pi} \tilde{f}(x)\sin(kx)dx + \int_0^{\pi} [f(x) - \tilde{f}(x)]\sin(kx)dx.$$

Второе слагаемое является малым и отбрасывается. Например, функция  $f(x)$  может быть заменена ее интерполяционным полиномом (т.е. степенным приближением).

#### 4.7. Список контрольных вопросов к 4-й главе

1. В каких случаях используется численное интегрирование?
2. Какова постановка задачи численного интегрирования?
3. Какие Вы знаете методы численного интегрирования функций?
4. Поясните графическую интерпретацию методов трапеций и Симпсона.
5. Чем отличаются методы трапеций и Симпсона по постановке и реализации алгоритма?
6. Как оценить погрешность результатов численного интегрирования?
7. Как и почему влияет на точность численного интегрирования величина шага  $\Delta x$ ?

## 5. ЧИСЛЕННОЕ ДИФФЕРЕНЦИРОВАНИЕ. ПРИБЛИЖЕННОЕ РЕШЕНИЕ ОБЫКНОВЕННЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

### 5.1. Численное дифференцирование функций

По определению производная функции  $f(x)$  определяется соотношением

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

Переходя здесь от бесконечно малых к конечным разностям, получаем приближенную формулу для численного дифференцирования функции  $f(x)$ :

$$f'(x) \approx \frac{f(x + \Delta x) - f(x)}{\Delta x}. \quad (10)$$

Эта формула позволяет приближенно вычислить производную функции  $f$  в любой точке  $x$  ее области определения, задавшись предварительно значением приращения  $\Delta x$ .

Замена бесконечно малых приращений аргумента  $x$  конечными является источником погрешностей. Для оценки их величины разложим функцию  $f$  в точке  $x$  в ряд Тейлора:

$$\begin{aligned} f(x + \Delta x) &= f(x) + \frac{f'(x)}{1!} (x + \Delta x - x) + \frac{f''(x)}{2!} (x + \Delta x - x)^2 + \dots + \frac{f^{(n)}(x)}{n!} (x + \Delta x - x)^n + \dots = \\ &= f(x) + \frac{f'(x)}{1!} \Delta x + \frac{f''(x)}{2!} \Delta x^2 + \frac{f'''(x)}{3!} \Delta x^3 + \dots + \frac{f^{(n)}(x)}{n!} \Delta x^n + \dots \end{aligned}$$

Перенеся  $f(x)$  в левую часть полученного равенства и разделив его левую и правую части на  $\Delta x$ , в соответствии с формулой (10) получим следующее приближенное равенство для  $f'(x)$ :

$$f'(x) \approx \frac{f(x + \Delta x) - f(x)}{\Delta x} = \frac{f'(x)}{1!} + \frac{f''(x)}{2!} \Delta x + \frac{f'''(x)}{3!} \Delta x^2 + \dots + \frac{f^{(n)}(x)}{n!} \Delta x^{n-1} + \dots,$$

то есть

$$f'(x) \approx f'(x) + \frac{1}{2} f''(x) \Delta x + \dots$$

Здесь все члены, начиная со второго, определяют отличие приближенного значения производной от ее точного значения. Основной член погрешности  $f''(x) \Delta x / 2$  пропорционален  $\Delta x$ . Поэтому говорят, что формула (10) имеет первый порядок точности по  $\Delta x$ .

Можно также приближенно вычислить производную, используя симметричную разностную схему:

$$f'(x) \approx \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x}. \quad (11)$$

Для оценки ее точности кроме использованного выше разложения функции  $f$  в точке  $x$  в ряд Тейлора запишем ее иное подобное разложение, используя первые четыре его члена:

$$\begin{aligned} f(x - \Delta x) &= f(x) + \frac{f'(x)}{1!}(x - \Delta x - x) + \frac{f''(x)}{2!}(x - \Delta x - x)^2 + \dots + \frac{f^{(n)}(x)}{n!}(x - \Delta x - x)^n + \dots = \\ &= f(x) - \frac{f'(x)}{1!}\Delta x + \frac{f''(x)}{2!}\Delta x^2 - \frac{f'''(x)}{3!}\Delta x^3 + \dots + \frac{f^{(n)}(x)}{n!}(-\Delta x)^n + \dots \end{aligned}$$

Вычтя из первого разложения второе и разделив полученное выражение на  $2\Delta x$ , приходим к следующему приближенному равенству:

$$\begin{aligned} f'(x) \approx \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x} &= \frac{1}{2\Delta x} \left[ f(x) + \frac{f'(x)}{1!}\Delta x + \frac{f''(x)}{2!}\Delta x^2 + \frac{f'''(x)}{3!}\Delta x^3 + \dots \right] - \\ &- \frac{1}{2\Delta x} \left[ f(x) - \frac{f'(x)}{1!}\Delta x + \frac{f''(x)}{2!}\Delta x^2 - \frac{f'''(x)}{3!}\Delta x^3 + \dots \right] = f'(x) + \frac{f'''(x)}{3!}\Delta x^2 + \dots \end{aligned}$$

Таким образом, основной член погрешности формулы (11) равен  $f'''(x)\Delta x^2/6$ . Поскольку он пропорционален  $\Delta x^2$ , то говорят, что формула (11) имеет второй порядок точности по  $\Delta x$ .

## 5.2. Численное интегрирование задачи Коши для систем обыкновенных дифференциальных уравнений

Рассмотрим задачу Коши для систем дифференциальных уравнений. Требуется найти вектор-функцию  $\bar{u}(t)$ , удовлетворяющую уравнению

$$\frac{d\bar{u}}{dt} = \bar{f}(\bar{u}, t), \quad \bar{u}(0) = \bar{u}_0, \quad 0 \leq t \leq T.$$

Здесь  $\bar{u}$  –  $p$ -мерный вектор  $\bar{u} = [u_1, u_2, \dots, u_p]^T$ ,  $\bar{f} = [f_1, f_2, \dots, f_p]^T$  – заданная вектор-функция той же размерности,  $\bar{u}_0$  – заданная точка (начальные условия, данные Коши).

Для приближенного решения поставленной задачи используем метод сеток, называемый также методом конечных разностей, который является одним из наиболее общих, хотя и достаточно трудоемких, методов решения дифференциальных уравнений.

Начнем с основных объектов метода сеток. На интервале  $[0, T]$ , на котором ищется решение, введем покрывающую его дискретную сетку точек  $0 = t_0 < t_1 < t_2 < \dots < t_n = T$ , или  $t_i, i=0, 1, \dots, n$ . Ради простоты в дальнейшем будем использовать сетку с постоянным шагом  $\tau = T/n$ ; тогда  $t_i = \tau \cdot i, i=0, 1, \dots, n$ .

Приближенно решение будем искать в виде сеточной функции, т.е. в виде функции дискретного аргумента  $i$ ; обозначим ее как  $\bar{u}_i$ ,  $i=0,1,\dots,n$ . Напомним, что каждый  $\bar{u}_i$  есть  $p$ -мерный вектор. С содержательной точки зрения  $\bar{u}_i$  приближенно представляет собой значение искомой вектор-функции в узле сетки  $t_i$ :  $\bar{u}_i \approx \bar{u}(t_i)$ ,  $i = 0,1,\dots,n$ .

Сеточная функция  $\bar{u}_i$  не может удовлетворять никакому дифференциальному уравнению, и нужно построить какие-то другие уравнения, позволяющие ее определить так, чтобы она была приближенным решением исходной задачи. Такие уравнения, называемые разностными, строятся следующим образом: входящие в дифференциальные уравнения производные заменяются соответствующими конечными разностями. Это можно сделать разными способами, и они приводят к разным уравнениям. Например:

$$\frac{\bar{u}_{i+1} - \bar{u}_i}{\tau} = \bar{f}(\bar{u}_i, t_i) \text{ – явная схема Эйлера,}$$

$$\frac{\bar{u}_{i+1} - \bar{u}_i}{\tau} = \bar{f}(\bar{u}_{i+1}, t_i) \text{ – неявная схема Эйлера,}$$

$$\frac{\bar{u}_{i+1} - \bar{u}_{i-1}}{2\tau} = \bar{f}(\bar{u}_i, t_i) \text{ – схема с центральной разностью.}$$

Во всех случаях значения  $\bar{u}_i$  определяются последовательно, слева направо. Особенно просто вычисляются  $\bar{u}_i$  при использовании явной схемы Эйлера:

$$\bar{u}_{i+1} = \bar{u}_i + \tau \cdot \bar{f}(\bar{u}_i, t_i), \quad i = 0,1,\dots,n.$$

Здесь  $\bar{u}_0$ , являющееся данными Коши, известно. В правой части этой формулы используются уже найденные значения  $\bar{u}_i$ .

В случае неявной схемы Эйлера ситуация несколько сложнее. Пусть значение  $\bar{u}_i$  уже найдено. Тогда значение  $\bar{u}_{i+1}$  находится из уравнения

$$\bar{u}_{i+1} = \bar{u}_i + \tau \cdot \bar{f}(\bar{u}_{i+1}, t_i),$$

которое является нелинейным относительно неизвестного  $\bar{u}_{i+1}$ . Правда, это слабая нелинейность, поскольку перед  $\bar{f}$  стоит малый множитель  $\tau$ . Интуитивно понятно, что  $\bar{u}_{i+1}$  мало отличается от  $\bar{u}_i$ , т.е.  $\bar{u}_i$  есть очень хорошее начальное приближение для какого-либо итерационного метода определения  $\bar{u}_{i+1}$ .

В схеме с центральной разностью мы сталкиваемся с характерным явлением: формальные порядки дифференциального и разностного уравнений не совпадают (под формальным порядком понимается число произвольных постоянных в общем решении, или, если угодно, число дополнительных данных, полностью определяющих решение). В явной и неявной схемах Эйле-



ра, как и в дифференциальном уравнении, достаточно задать  $\bar{u}_0$ , чтобы все остальные значения  $\bar{u}_1, \bar{u}_2, \dots, \bar{u}_n$  определялись однозначно. В схеме с центральной разностью ситуация иная: только задав  $\bar{u}_0$  и  $\bar{u}_1$ , удастся определить  $\bar{u}_2 = \bar{u}_0 + 2\tau \cdot \bar{f}(\bar{u}_1, t_1)$  и т.д.

Значение  $\bar{u}_0$  определяется постановкой задачи, а  $\bar{u}_1$  формально можно задать каким угодно. Фактически, чтобы получить хорошие результаты,  $\bar{u}_1$  нужно задавать достаточно аккуратно. Например, приемлемым (но не лучшим) является выбор  $\bar{u}_1 = \bar{u}_0$ . Лучше (и правильнее) вычислить  $\bar{u}_1$  по какой-либо простой разностной схеме, например, по явной схеме Эйлера:  $\bar{u}_1 = \bar{u}_0 + \tau \cdot \bar{f}(\bar{u}_0, t_0)$ .

### Пример 5.

Решить методом Эйлера дифференциальное уравнение  $\frac{du}{dt} = t^2$ ,  $t \in [0, 5]$ ,  $u_0 = u|_{t=0} = 1,3$ .

В данном случае  $f(u, t) = t^2$ . Эта задача имеет аналитическое решение  $u(t) = t^3/3 + c$ . Поскольку  $u(0) = c = u_0$ , то  $u(t) = t^3/3 + u_0$ .

Начнем с построения конечно-разностной сетки:  $t_0 = 0$ ,  $T = t_n = 5$ . Пусть  $n = 10$ . Тогда  $\tau = \frac{T - t_0}{n} = \frac{5 - 0}{10} = 0,5$ ,  $t = i \cdot \tau$ ,  $i = \overline{0, n}$ .

Используем явный метод Эйлера. Тогда значения функции  $u_i$  в узлах сетки определяются уравнением

$$u_{i+1} = u_i + \tau \cdot f(u_i, t_i) = u_i + \tau \cdot t_i^2,$$

а погрешность вычислений определим по формуле

$$\delta u_i = \left| \frac{u(t_i) - u_i}{u(t_i)} \right| \cdot 100\%.$$

Результаты расчета приведены в табл. 6:

Таблица 6

Результаты решения примера 5 с шагом сетки  $\tau = 0,5$ .

$i$	$t_i$	$u_i$	$u(t_i)$	$\delta u_i, \%$
0	0,0	1,300	1,300	0,000
1	0,5	1,425	1,342	6,211
2	1,0	1,925	1,633	17,857
3	1,5	3,050	2,425	25,773
4	2,0	5,050	3,967	27,311
5	2,5	8,175	6,508	25,608
6	3,0	12,675	10,300	23,058
7	3,5	18,800	15,592	20,577
8	4,0	26,800	22,693	18,409
9	4,5	36,925	31,675	16,575
10	5,0	49,425	42,967	15,031

Для сравнения приведем результаты расчета, полученные при  $N = 100$  и  $\tau = 0,05$ , т.е. при шаге сетки, уменьшенном в 10 раз. Соответствующие данные показаны в табл. 7 для тех же самых значений  $t$ , что и в предыдущей таблице.

Таблица 7

Результаты решения примера 5 с шагом сетки  $\tau = 0,05$

$i$	$t_i$	$u_i$	$u(t_i)$	$\delta u_i, \%$
0	0,0	1,300	1,300	0,000
10	0,5	1,348	1,342	0,481
20	1,0	1,659	1,633	1,556
30	1,5	2,482	2,425	2,345
40	2,0	4,068	3,697	2,542
50	2,5	6,666	6,508	2,417
60	3,0	10,526	10,300	2,197
70	3,5	15,899	15,592	1,974
80	4,0	23,035	22,633	1,775
90	4,5	32,183	31,675	1,604
100	5,0	43,594	42,967	1,459

Анализируя данные, приведенные в табл. 6 и 7, можно заключить, что при увеличении числа шагов значения погрешности  $\delta u_i$  уменьшаются.

### 5.3. Список контрольных вопросов к 5-й главе

1. Что значит – решить задачу Коши для ОДУ первого порядка?
2. В чем заключается графическая интерпретация численного решения ОДУ методом Эйлера?
3. В чем заключается суть метода Эйлера?
4. Какие схемы реализации метода Эйлера Вы знаете? В чем их отличие?
5. Применение каких формул позволяет получить значения искомой функции по методу Эйлера?
6. Как оценить погрешность результатов численного решения ОДУ?
7. Как и почему влияет на точность численного решения ОДУ величина шага  $h$ ?

## **БИБЛИОГРАФИЧЕСКИЙ СПИСОК**

1. **Самарский А.А.** Введение в численные методы: учебное пособие для вузов. М.: Лань, 2005.
2. **Бахвалов Н.С., Жидков Н.П., Кобельков Г.М.** Численные методы. М.: БИНОМ. Лаборатория знаний, 2008.
3. **Самарский А.А., Гулин А.В.** Численные методы. – М.: Наука, 1989.
4. **Физико-математические основы техники и электрофизики высоких напряжений.** Учебное пособие для вузов / В.В. Базуткин, К.П. Кадомская, Е.С. Колечицкий и др. Под ред. К.П. Кадомской. М.: Энергоатомиздат, 1995.
5. **Мирзабекян Г.З.** Численные методы в технике высоких напряжений. М.: Издательство МЭИ, 1983.
6. **Мирзабекян Г.З.** Численные методы в технике высоких напряжений. М.: Издательство МЭИ, 1985.
7. **Федоренко Р.П.** Введение в вычислительную физику. М.: Издательство МФТИ, 1994.
8. **Поршнев С.В., Беленкова И.В.** Численные методы на базе Mathcad. СПб.: БХВ-Петербург, 2005.
9. **Мину М.** Математическое программирование. М.: Наука, 1990.
10. **Грешилов А.А.** Как принять наилучшее решение в реальных условиях. М.: Радио и связь, 1991.
11. **Гремяков А.А.** Методы решения оптимизационных электроэнергетических задач. М.: Издательство МЭИ, 1989.
12. **Белогловский А.А.** Математические основы теории принятия оптимальных решений. М.: Издательство МЭИ, 1999.

## ОГЛАВЛЕНИЕ

Введение	2
1. Численное решение уравнений и их систем	3
1.1. Некоторые сведения из функционального анализа	3
1.2. Решение нелинейного уравнения методом простых итераций	5
1.3. Решение нелинейных уравнений и их систем методом Ньютона	6
1.4. Список контрольных вопросов к 1-й главе	7
2. Основы решения нелинейных задач математического программирования	9
2.1. Постановка задачи математического программирования	9
2.2. Особенности поиска решений в задачах математического программирования	9
2.3. Общий алгоритм решения задачи математического программирования	13
2.4. Метод Лагранжа	14
2.5. Основы численного решения задач безусловной оптимизации	15
2.6. Список контрольных вопросов ко 2-й главе	18
3. Интерполяция функций одной переменной	19
3.1. Понятие интерполяции функции	19
3.2. Кусочно-линейная интерполяция	19
3.3. Сплайн-интерполяция	20
3.4. Сглаживание функции степенным многочленом с использованием метода наименьших квадратов	21
3.5. Список контрольных вопросов к 3-й главе	24
4. Вычисление определенных интегралов	25
4.2. Формула трапеций	25
4.3. Формула Симпсона	26
4.4. Оценка точности численного интегрирования	26
4.5. Вычисление интегралов с особенностями	27
4.6. Вычисление интегралов от быстроосциллирующих функций	28
4.7. Список контрольных вопросов к 4-й главе	29
5. Численное дифференцирование. Приближенное решение обыкновенных дифференциальных уравнений	30
5.1. Численное дифференцирование функций	30

5.2. Численное интегрирование задачи Коши для систем обыкновенных дифференциальных уравнений	31
5.3. Список контрольных вопросов к 5-й главе	34
Библиографический список	35