

**В. А. Гончаров**

# Методы оптимизации

УЧЕБНОЕ ПОСОБИЕ

*Министерство образования и науки Российской Федерации  
Московский государственный институт электронной техники  
(технический университет)*

*Утверждено редакционно-издательским советом института*

**МОСКВА • «ВЫСШЕЕ ОБРАЗОВАНИЕ» • 2008**

УДК 51  
ББК 22.193-018.2\*32.973я73  
Г 57



Настоящая методическая разработка выполнена в рамках инновационной образовательной программы МИЭТ «Современное профессиональное образование для российской инновационной системы в области электроники»

Г 57 **Гончаров, В. А.**  
Методы оптимизации : учебник / В. А. Гончаров. — М. :  
Высшее образование, 2008. — 204 с. — (Основы наук).

ISBN 978-5-9692-0338-9

Пособие посвящено систематическому изложению основ методов оптимизации и имеет прикладную инженерно-техническую направленность. Основное внимание уделено прикладным и вычислительным аспектам оптимизации, связанным с разработкой численных методов решения задач и построением алгоритмов их реализации.

*Для студентов, обучающихся по специальностям «Прикладная математика» и «Программное обеспечение вычислительной техники и автоматизированных систем», однако в силу актуальности рассматриваемых вопросов будет полезным и для студентов, специализирующихся в смежных областях.*

УДК 51  
ББК 22.193-018.2\*32.973я73

По вопросам приобретения обращайтесь в **книготорг «Юрайт»**  
Тел.: (495) 744-00-12. E-mail: sales@urait.ru, www.urait.ru

**Покупайте наши книги:**

—в нашем офисе: 140004, Московская обл., г. Люберцы, 1-й Панковский проезд, д. 1;  
—через интернет-магазин: [www.books.urait.ru](http://www.books.urait.ru); e-mail: [books@books.urait.ru](mailto:books@books.urait.ru).

ISBN 978-5-9692-0338-9 © Гончаров В. А., 2008  
© ООО «Высшее образование», 2008

## Оглавление

|   |           |
|---|-----------|
| <b>Предисловие</b> .....  | <b>6</b>  |
| <b>Глава 1. Задачи оптимизации. Основные определения</b> .....  | <b>8</b>  |
| 1.1. Задачи оптимизации .....   | 8         |
| 1.2. Минимум функции одной переменной .....   | 11        |
| 1.3. Унимодальные функции .....   | 13        |
| 1.4. Выпуклые функции .....   | 15        |
| 1.5. Условие Липшица .....  | 17        |
| 1.6. Классическая минимизация функции одной переменной .....  | 20        |
| <b>Глава 2. Одномерная минимизация функций. Прямые методы</b> .....   | <b>24</b> |
| 2.1. О прямых методах .....   | 24        |
| 2.2. Метод перебора .....   | 25        |
| 2.3. Метод поразрядного поиска .....  | 26        |
| 2.4. Метод дихотомии .....  | 28        |
| 2.5. Метод золотого сечения .....   | 31        |
| 2.6. Сравнение методов перебора, дихотомии и золотого сечения .....   | 34        |
| 2.7. Метод парабол .....  | 35        |
| <b>Глава 3. Одномерная минимизация. Методы, использующие информацию о производных целевой функции</b> ..... | <b>43</b> |
| 3.1. Метод средней точки .....  | 43        |
| 3.2. Метод хорд .....   | 45        |
| 3.3. Метод Ньютона .....  | 47        |
| 3.4. Возможные модификации метода Ньютона .....   | 53        |
| 3.5. Методы минимизации многомодальных функций .....  | 54        |

|   |            |
|---|------------|
| <b>Глава 4. Задача минимизации функции многих переменных. Необходимые и достаточные условия безусловного экстремума .....</b>               | <b>64</b>  |
| 4.1. Постановка задачи и определения.....   | 64         |
| 4.2. Свойства выпуклых множеств и выпуклых функций .....  | 69         |
| 4.3. Необходимые и достаточные условия безусловного экстремума .....  | 71         |
| <br><b>Глава 5. Общие принципы многомерной минимизации. Методы градиентного спуска. Метод сопряженных направлений и метод Ньютона .....</b> | <b>82</b>  |
| 5.1. Выпуклые квадратичные функции .....  | 82         |
| 5.2. Общие принципы многомерной минимизации .....   | 84         |
| 5.3. Метод градиентного спуска.....   | 89         |
| 5.4. Метод наискорейшего спуска.....  | 92         |
| 5.5. Метод сопряженных направлений .....  | 94         |
| 5.6. Метод сопряженных градиентов .....   | 98         |
| 5.7. Метод Ньютона.....   | 101        |
| 5.8. Квазиньютоновские методы.....  | 104        |
| <br><b>Глава 6. Прямые методы безусловной минимизации многомерных задач .....</b>   | <b>110</b> |
| 6.1. Проблема минимизации многомерных задач.....  | 110        |
| 6.2. Минимизация функций по правильному (регулярному) симплексу .....   | 113        |
| 6.3. Минимизация функций при помощи нерегулярного симплекса .....   | 120        |
| 6.4. Метод циклического покоординатного спуска .....  | 122        |
| 6.5. Метод Хука – Дживса .....  | 127        |
| 6.6. Методы случайного поиска.....  | 128        |
| <br><b>Глава 7. Условный экстремум функции многих переменных. Необходимые и достаточные условия условного экстремума .....</b>              | <b>134</b> |
| 7.1. Условный экстремум при ограничениях типа равенств .....  | 134        |
| 7.2. Условный экстремум при ограничениях типа неравенств .....  | 150        |

|  |            |
|--|------------|
| <b>Глава 8. Линейное программирование .....</b>                        | <b>164</b> |
| 8.1. Определения. Примеры задач линейного программирования.....        | 164        |
| 8.2. Общая и каноническая задачи линейного программирования.....       | 167        |
| 8.3. Геометрическое истолкование задач линейного программирования..... | 169        |
| 8.4. Аналитическое решение задач линейного программирования.....       | 174        |
| <br><b>Литература .....</b>  | <b>204</b> |

## Предисловие

Предлагаемое учебное пособие построено на материалах, используемых при чтении одноименного курса лекций, практических занятий и лабораторно-вычислительного практикума, которые проводятся автором в Московском институте электронной техники, и учитывают специфику математической подготовки студентов различных вузов.

Автор попытался создать учебное пособие, которое охватывало бы основную часть предлагаемого курса и не требовало бы частого дополнительного обращения к порой труднодоступной специальной литературе, за исключением случаев, когда необходимо разобраться в доказательствах сходимости отдельных методов. Автор исходил из того, что для будущего инженера важнее понять суть методов и алгоритмов их реализации, условий их применения и получить навыки решения типовых задач оптимизации.

Охарактеризуем кратко особенности содержания книги. В главе 1 вводятся и обсуждаются основные термины и определения. Ядро книги составляют главы 2, 3 и 5, 6, в которых рассматриваются соответственно одномерная и многомерная минимизации с использованием принципиально важных прямых методов, а так же методов, использующих информацию о производной (или градиенте) целевой функции. В главах 4, 7 формулируются необходимые и достаточные условия существования безусловного и условного экстремумов, приводятся примеры аналитического решения экстремальных задач. Заключительная глава посвящена разбору задач линейного программирования на конкретных примерах.

Курс «Методы оптимизации» продолжает традиции математической подготовки студентов, связанной с изучением таких дисциплин, как «Математическое моделирование» и

«Численные методы». В конце глав приводятся вопросы и задания для самопроверки. Кроме того, в заключительной части четырех из восьми глав приведены тексты заданий для численной реализации, которые предполагается выполнить в современной среде программирования MATLAB. Эти задания могут служить основой для лабораторно-вычислительных работ, выполняемых при изучении дисциплины «Методы оптимизации».

## Глава 1

# ЗАДАЧИ ОПТИМИЗАЦИИ. ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

### 1.1. Задачи оптимизации

В любой сфере человеческой деятельности, как на сугубо личном, так и на общегосударственном уровне, явно или неявно, мы встречаемся с оптимизацией. Экономическое планирование, управление, проектирование сложных объектов всегда направлено на поиск наилучшего варианта с точки зрения намеченной цели.

При всем многообразии задач оптимизации дать общие методы их решения может только математика, резкое расширение приложений которой связано с появлением ЭВМ, что привело к математизации не только физики, но и химии, биологии, экономики, психологии, медицины — практически всех наук. Суть математизации состоит в построении математических моделей процессов и явлений и в разработке методов их исследования.

Использование математического аппарата при решении задач оптимизации предполагает формулировку интересующей проблемы на языке математики, придание количественных оценок возможным вариантам вместо слов «лучше», «хуже».

Многие задачи оптимизации сводятся к отысканию наименьшего или наибольшего значения некоторой функции, которую принято называть *целевой функцией*. Такая постановка задачи будет обычной при дальнейшем изложении. В этом случае методы исследования существенно зависят от свойств целевой функции и той информации о ней, которая может считаться доступной до решения задачи и в процессе решения.

Наиболее просты с математической точки зрения случаи, когда целевая функция является дифференцируемой функцией. В этом случае для исследования ее свойств (участки возрастания и убывания, точки локального экстремума) может быть использована производная.

В последние десятилетия в условиях научно-технического прогресса круг задач оптимизации, поставленных практикой, существенно расширился. Во многих из них значения целевой функции могут получаться в результате численных расчетов или браться из эксперимента. Такие задачи являются более сложными, при их решении нельзя исследовать целевую функцию с помощью производной. Это привело к разработке специальных методов, рассчитанных на широкое применение ЭВМ. Следует также иметь в виду, что сложность решения задачи существенно зависит от размерности целевой функции, т.е. от числа ее аргументов.

**Пример 1.1.** Указать наилучший вариант консервной банки фиксированного объема  $V$ , имеющей обычную форму прямого кругового цилиндра.

Рассмотрим два варианта этой задачи.

1. Наилучшая банка должна иметь наименьшую поверхность  $S$  (на изготовление пойдет наименьшее количество жести).

2. Наилучшая банка должна иметь наименьшую длину швов  $l$  (швы нужно сваривать, и эта работа должна быть минимальной).

Запишем формулы объема банки, площади поверхности и длины швов

$$V = \pi r^2 h, \quad S = 2\pi r^2 + 2\pi r h, \quad l = 4\pi r + h.$$

Объем банки задан, это устанавливает связь между радиусом  $r$  и высотой  $h$ . Выразив высоту через радиус  $h = V / \pi r^2$  и подставив полученное выражение в формулы для поверхности и длины швов, получим

$$S(r) = 2\pi r^2 + \frac{2V}{r}, \quad 0 < r < \infty, \quad (1.1)$$

$$l(r) = 4\pi r + \frac{V}{\pi r^2}, \quad 0 < r < \infty. \quad (1.2)$$

Таким образом, с математической точки зрения задача о наилучшей консервной банке сводится к определению значения  $r$ , при котором достигает своего наименьшего значения в одном случае функция  $S(r)$ , в другом —  $l(r)$ . Поскольку  $S(r)$  и  $l(r)$  дифференцируемы, то задача решается просто.

1. Первый вариант постановки задачи: из формулы (1.1) следует

$$S'(r) = 4\pi r - \frac{2V}{r^2} = \frac{2}{r^2}(2\pi r^3 - V).$$

При  $0 < r < r_1 = \sqrt[3]{V/2\pi}$   $S'(r) < 0$ , функция убывает, при  $r_1 < r < \infty$   $S'(r) > 0$ , функция возрастает. Следовательно,  $S = S_{\min}$  при  $r = r_1$ , где

$$r_1 = \sqrt[3]{\frac{V}{2\pi}}, \quad h_1 = 2r_1, \quad (1.3)$$

а  $S(r_1) = 3\sqrt[3]{2\pi V^2} \leq S(r)$ . График функции  $S(r)$  показан на рис. 1.1.

2. Второй вариант постановки задачи: из формулы (1.2) следует

$$l'(r) = 4\pi - \frac{2V}{\pi r^3} = \frac{2}{\pi r^3}(2\pi^2 r^3 - V).$$

При  $0 < r < r_2 = \sqrt[3]{V/2\pi^2}$   $l'(r) < 0$ , функция убывает, при  $r_2 < r < \infty$   $l'(r) > 0$ , функция возрастает. Следовательно,  $l = l_{\min}$  при  $r = r_2$ , где

$$r_2 = \sqrt[3]{\frac{V}{2\pi^2}}, \quad h_2 = 2\pi r_2, \quad (1.4)$$

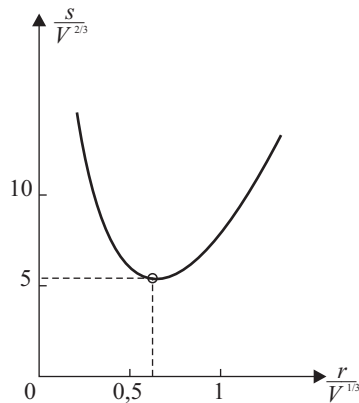


Рис. 1.1. График функции  $S(r)$

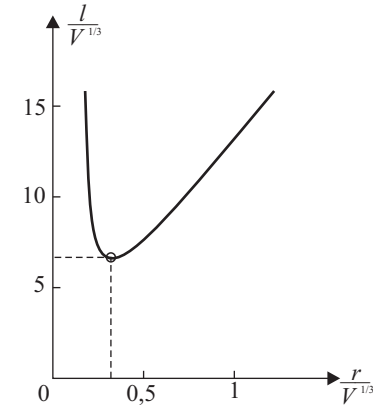


Рис. 1.2. График функции  $l(r)$

при этом  $l(r_2) = 3\sqrt[3]{4\pi V} \leq l(r)$ . График функции  $l(r)$  показан на рис. 1.2.

Очевидно, что при разных критериях оптимизации получаются различные ответы. В первом варианте (1.3) высота наилучшей банки равна ее диаметру, а во втором — (1.4) она в  $\pi$  раз больше него.

Перейдем от конкретного примера к общей постановке задач и определениям.

## 1.2. Минимум функции одной переменной

Рассмотрим математическую модель оптимизации, в которой целевая функция зависит от одной переменной, определенной на множестве  $U$  вещественной оси

$$\begin{aligned} f(x) &\rightarrow \min. \\ x &\in U \end{aligned} \quad (1.5)$$

Максимизация целевой функции ( $f(x) \rightarrow \max$ ) эквивалентна минимизации противоположной величины ( $-f(x) \rightarrow \min$ ),

поэтому, не умаляя общности, будем рассматривать только задачи минимизации.

**Определение.** Число  $x^* \in U$  называется *точкой глобального* (абсолютного) *минимума* или просто *точкой минимума* функции  $f(x)$  на множестве  $U$ , если  $f(x^*) \leq f(x)$  для всех  $x \in U$ . Множество всех точек минимума  $f(x)$  на  $U$  будем в дальнейшем обозначать через  $U^*$ .

**Определение.** Число  $\tilde{x} \in U$  называется *точкой локального минимума* функции  $f(x)$ , если  $f(\tilde{x}) \leq f(x)$  для всех  $x \in U$ , достаточно близких к  $\tilde{x}$ , т.е. если существует  $\varepsilon > 0$  такое, что это неравенство выполняется для любого

$$x \in \{x \in U, |x - \tilde{x}| < \varepsilon\}.$$

**Определение.** Пусть функция  $f(x)$  определена и ограничена снизу на множестве  $U$ , т.е.  $f(x) \geq A > -\infty$  для всех  $x \in U$ . Число  $f_*$  называется *точной нижней гранью* функции  $f(x)$  на множестве  $U$  ( $f_* = \inf_U f(x)$ ), если  $f(x) \geq f_*$  при всех

$x \in U$  и для любого  $\varepsilon > 0$  найдется точка  $x_\varepsilon \in U$  такая, что  $f(x_\varepsilon) < f_* + \varepsilon$  (т.е. среди значений  $f(x)$  на множестве  $U$  найдутся сколь угодно близкие к  $f_*$ ).

Для неограниченных функций  $f(x)$  полагают  $f_* = -\infty$ .

Необходимо отметить следующее.

1. Глобальный минимум  $f(x)$  является и локальным минимумом, а обратное, вообще говоря, неверно.

2. Множество точек минимума  $U^*$  функции  $f(x)$  на множестве  $U$  может быть пустым, состоять из конечного или бесконечного числа точек.

Например:

а) если  $f(x) = \ln x$ ,  $U = (0, 1]$ , то  $U^* = \emptyset$ ;

б) если  $f(x) = x^2$ ,  $U = [-1, 1]$ , то  $U^* = \{0\}$  — конечное множество;

в) если  $f(x) = \sin^2 \pi x$ ,  $U = \mathbb{R}$ , то  $U^* = \mathbb{Z}$  — бесконечное множество.

3. Если  $U^* \neq \emptyset$ , то  $\inf_U f(x) = \min_U f(x)$ . Таким образом,

точная нижняя грань обобщает понятие минимума функции на случай  $U^* = \emptyset$ .

Широкий класс функций, для которых  $U^* \neq \emptyset$ , определяет известная из математического анализа теорема Вейерштрасса, утверждающая, что непрерывная на замкнутом ограниченном множестве функция достигает на нем своих минимального и максимального значений. Таким образом, задача отыскания минимума (1.5) для непрерывных функций всегда имеет решение.

### 1.3. Унимодальные функции

Если функция  $f(x)$ , определенная на множестве  $U$  имеет, кроме глобального, еще и локальные минимумы, отличные от него, то минимизация  $f(x)$ , как правило, сильно затрудняется. Большинство методов поиска минимума  $f(x)$  приспособлено только для тех функций, у которых каждый локальный минимум является одновременно и глобальным. Этим свойством обладают унимодальные функции.

**Определение.** Функция  $f(x)$  называется *унимодальной* на отрезке  $[a, b]$ , если она непрерывна на  $[a, b]$  и существуют числа  $\alpha$  и  $\beta$ ,  $a \leq \alpha \leq \beta \leq b$ , такие, что:

1) если  $a < \alpha$ , то на отрезке  $[a, \alpha]$  функция  $f(x)$  монотонно убывает;

2) если  $\beta < b$ , то на отрезке  $[\beta, b]$  функция  $f(x)$  монотонно возрастает;

3) при  $x \in [\alpha, \beta]$  выполняется  $f(x) = f_* = \min_{[a, b]} f(x)$ .

Отметим, что возможно вырождение в точку одного или двух отрезков из  $[a, \alpha]$ ,  $[\alpha, \beta]$  и  $[\beta, b]$ . Некоторые варианты расположения и вырождения в точку отрезков монотонности и постоянства унимодальной функции показаны на рис.1.3.

Из определения вытекают следующие свойства унимодальных функций.

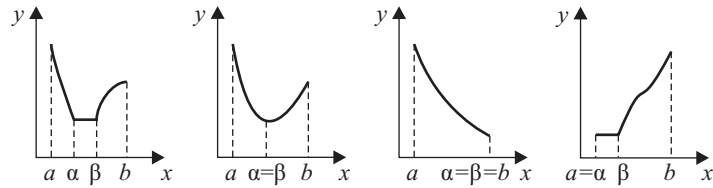


Рис. 1.3. Графики унимодальных функций

1. Любая из точек локального минимума унимодальной функции является и точкой ее глобального минимума на отрезке  $[a, b]$ .

2. Функция, унимодальная на отрезке  $[a, b]$ , унимодальна и на любом меньшем отрезке  $[c, d] \subset [a, b]$ .

3. Пусть  $f(x)$  унимодальна на  $[a, b]$  и  $a \leq x_1 < x_2 \leq b$ . Тогда

если  $f(x_1) \leq f(x_2)$ , то  $x^* \in [a, x_2]$ ,

если  $f(x_1) > f(x_2)$ , то  $x^* \in [x_1, b]$ ,

где  $x^*$  — одна из точек минимума  $f(x)$  на отрезке  $[a, b]$ .

**Пример 1.2.** На какие три части следует разбить отрезок  $[-1, 2]$ , чтобы на каждой из них функция  $f(x) = ||x(x-1)| - 1|$  была унимодальной?

Строя поэтапно график функции  $f(x)$ , получим три отрезка:  $[-1, 0]$ ,  $[0, 1]$ , и  $[1, 2]$ , на каждом из которых функция унимодальна.

**Пример 1.3.** Найти максимальное значение  $b$ , при котором функция  $f(x) = -x^2 + 5x - 6$  унимодальна на отрезке  $[-5, b]$ .

Корни квадратного трехчлена  $x_1 = 2$ ,  $x_2 = 3$ ; вершина параболы  $x_0 = 2,5$ . Поэтому  $b = 2,5$ .

## 1.4. Выпуклые функции

Функция  $f(x)$ , заданная на отрезке  $[a, b]$ , называется *выпуклой* на этом отрезке, если для всех  $x', x'' \in [a, b]$  и произвольного числа  $\alpha \in [0, 1]$  выполняется неравенство

$$f(\alpha x' + (1-\alpha)x'') \leq \alpha f(x') + (1-\alpha)f(x''). \quad (1.6)$$

**Свойство 1.** Если функция  $f(x)$  выпукла на  $[a, b]$ , то на любом  $[x', x''] \subset [a, b]$  ее график расположен не выше хорды, проведенной через точки графика с абсциссами  $x'$  и  $x''$ , рис. 1.4.

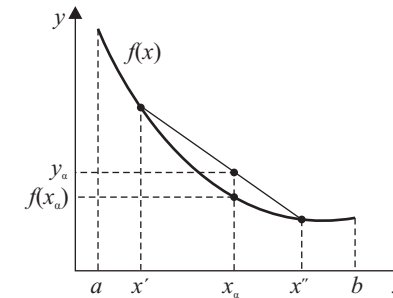


Рис. 1.4. Взаимное расположение графика выпуклой функции и хорды

**Свойство 2.** Можно показать, что всякая выпуклая непрерывная функция на  $[a, b]$  является и унимодальной на этом отрезке. Обратное, вообще говоря, неверно.

Из курса математического анализа известны следующие условия выпуклости функции:

а) для того, чтобы дифференцируемая на отрезке  $[a, b]$  функция  $f(x)$  была выпуклой на этом отрезке, необходимо и достаточно, чтобы ее производная  $f'(x)$  не убывала на

$[a, b]$  (например:  $y = x^2$ ,  $x \in \mathbb{R}$ ;  $y = \cos x$ ,  $x \in [\frac{\pi}{2}, \frac{3\pi}{2}]$ ).

б) для того, чтобы дважды дифференцируемая на отрезке  $[a, b]$  функция  $f(x)$  была выпуклой на этом отрезке, необходимо и достаточно, чтобы при всех  $x \in [a, b]$  выполнялось

неравенство  $f''(x) \geq 0$  ( $y = x^2$ ,  $x \in \mathbb{R}$ ;  $y = \cos x$ ,  $x \in [\frac{\pi}{2}, \frac{3\pi}{2}]$ ).

При исследовании выпуклости функций на практике неравенство (1.6) удастся использовать только в редких случаях.



Поэтому для дифференцируемых достаточно количество раз функций обычно применяют вышеприведенные дифференциальные критерии выпуклости.

Непосредственная проверка унимодальности с помощью определения этого понятия также в большинстве случаев вызывает затруднения. Поэтому для обоснования унимодальности достаточно гладких функций используют те же критерии выпуклости. Если функция оказывается выпуклой, то можно утверждать, что она унимодальна. Однако при отрицательном результате проверки функции на выпуклость нельзя сделать вывод, что она не унимодальна.

**Пример 1.4.** Показать, что функция  $f(x) = x^4 - 10x^3 + 36x^2 + 5x$  унимодальна на отрезке  $[3, 5]$ .

Вторая производная функции  $f(x)$  равна  $f''(x) = 12x^2 - 60x + 72$ . Корни полученного квадратного трехчлена  $x_1 = 2$  и  $x_2 = 3$ . Следовательно,  $f''(x) \geq 0$ , если  $x \geq 3$ , в частности, при  $x \in [3, 5]$ . По дифференциальному критерию «б» получаем, что  $f(x)$  выпукла, а значит, и унимодальна на этом отрезке.

**Пример 1.5.** В следующих задачах убедиться в унимодальности функций  $f(x)$  на указанных отрезках  $[a, b]$ .

1)  $f(x) = x^2 - 3x + x \ln x$ ,  $x \in [1, 2]$ .

$f''(x) = 2 + \frac{1}{x} > 0$  при  $x \in [1, 2]$ . По дифференциальному критерию

«б»  $f(x)$  выпуклая, а значит, и унимодальная при  $x \in [1, 2]$ .

2)  $f(x) = \ln(1+x^2) - \sin x$ ,  $x \in [0, \frac{\pi}{4}]$ .

$f''(x) = \frac{2(1-x^2)}{(1+x^2)^2} + \sin x > 0$  при  $x \in [0, \frac{\pi}{4}]$ . По дифференциальному

критерию «б»  $f(x)$  выпуклая, а значит, и унимодальная при  $x \in [0, \frac{\pi}{4}]$ .

3)  $f(x) = \frac{1}{2}x^2 - \sin x$ ,  $x \in [0, 1]$ .

$f''(x) = 1 + \sin x \geq 0$  при  $x \in [0, 1]$ . По дифференциальному критерию «б»  $f(x)$  выпуклая, а значит, и унимодальная при  $x \in [0, 1]$ .

**Пример 1.6.** Привести примеры функций  $f(x)$ , унимодальных на отрезке  $[a, b]$ , но не выпуклых на нем.

$$f(x) = -\frac{1}{1+x^2}, \quad x \in R \text{ или } x \in [a, b],$$

где  $[a, b]$  — такой отрезок, что  $0 \in (a, b)$ .

Очевидно, что функция унимодальна на  $[a, b]$ , но  $f'(x) = \frac{2x}{(1+x^2)^2}$ ,

вторая производная  $f''(x) = \frac{2(1-3x^2)}{(1+x^2)^3} < 0$  например, при  $|x| > \frac{1}{\sqrt{3}}$ . Функ-

ция  $f(x)$  не является выпуклой по дифференциальному критерию выпуклости «б».

$f(x) = |\ln x|$ ,  $x \in (0, +\infty)$ . Очевидно, что функция унимодальна при  $x \in (0, +\infty)$ , но не является выпуклой при  $x \in (1, +\infty)$ , т.к. на этом ин-

тервале  $f''(x) = -\frac{1}{x^2} < 0$ .

**Пример 1.7.** Будет ли функция  $f(x) = ax^3 - 3x^2 - 10$  унимодальной на отрезке  $[1, 2]$  при  $a > 3$ ?

Первая производная функции  $f'(x) = 3ax^2 - 6x$ , вторая —  $f''(x) = 6a(x - \frac{1}{a})$ . Если  $x \in [1, 2]$ , а величина  $a > 3$ , т.е.  $0 < \frac{1}{a} < \frac{1}{3}$ , то

$f''(x) > 0$  и функция  $f(x)$  унимодальна по дифференциальному критерию «б» выпуклости функций.

**Пример 1.8.** Доказать, что из выпуклости функции  $f(x)$  на отрезке  $[a, b]$  следует ее унимодальность на  $[a, b]$ , если ограничиваться только дифференцируемыми на  $[a, b]$  функциями.

Пусть  $f(x)$  выпукла на  $[a, b]$ , тогда  $f'(x)$  не убывает на этом отрезке. Допустим противное, т.е. что  $f(x)$  не унимодальна на  $[a, b]$ . Тогда существуют  $x_1, x_2, x_3 \in [a, b]$  такие, что  $f(x_1) < f(x_2)$  и  $f(x_3) < f(x_2)$ . А это противоречит дифференциальному критерию «а» выпуклости  $f(x)$  на  $[a, b]$ .

## 1.5. Условие Липшица

Применение некоторых методов одномерной минимизации возможно только в том случае, если скорость изменения целевой функции  $f(x)$  на любых участках отрезка  $[a, b]$

ограничена некоторым числом, одним и тем же для всех участков. В этом случае говорят, что  $f(x)$  удовлетворяет на отрезке  $[a, b]$  условию Липшица. Целевые функции большинства практических задач оптимизации таким свойством обладают.

**Определение.** Функция  $f(x)$  удовлетворяет на отрезке  $[a, b]$  условию Липшица, если существует такое число  $L > 0$  (константа Липшица), что

$$|f(x') - f(x'')| \leq L |x' - x''| \quad (1.7)$$

для всех  $x'$  и  $x''$ , принадлежащих  $[a, b]$ .

Необходимо обратить внимание на следующее.

1. Если неравенство (1.7) выполняется с константой  $L$ , то оно справедливо и для всех  $L' > L$ . Поэтому для функции, удовлетворяющей условию Липшица, существует бесконечное множество констант  $L$  из (1.7). При использовании алгоритмов минимизации, включающих  $L$  как параметр, наилучшие результаты достигаются, как правило, если в качестве  $L$  берется минимальная из констант Липшица.

2. Из условия (1.7) сразу следует непрерывность  $f(x)$  на отрезке  $[a, b]$ . Поэтому, согласно теореме Вейерштрасса, функция  $f(x)$ , удовлетворяющая на отрезке  $[a, b]$  условию Липшица, имеет на нем хотя бы одну точку минимума, хотя не является, вообще говоря, унимодальной.

3. Условие (1.7) означает, что модуль углового коэффициента любой хорды графика  $f(x)$  не превосходит  $L$ . Переходя в (1.7) к пределу при  $|x' - x''| \rightarrow 0$ , убеждаемся, что если в некоторой точке существует касательная к графику  $f(x)$ , то модуль ее углового коэффициента также не может превышать  $L$ . Так, функция  $f(x) = \sqrt{x}$  на отрезке  $[0, 1]$  условию Липшица не удовлетворяет, потому что при  $x \rightarrow +0$  угловой коэффициент касательной к ее графику  $k$  неограниченно возрастает (рис. 1.5).

4. Если функция  $f(x)$  имеет на отрезке  $[a, b]$  непрерывную производную, то она удовлетворяет на этом отрезке условию Липшица с константой  $L = \max_{[a, b]} |f'(x)|$ .

**Пример 1.9.** Найти наименьшую из констант Липшица функции

$$f(x) = \frac{1}{3}x^3 + 2x^2 - 5x + 6 \text{ на отрезках а) } x \in [0, 1], \text{ б) } x \in [0, 10].$$

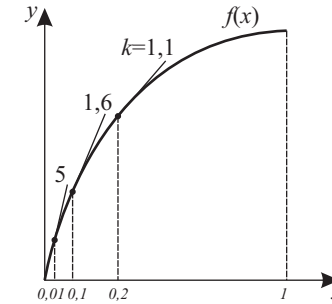


Рис.1.5. График функции  $f(x) = \sqrt{x}, x \in [0, 1]$ , не удовлетворяющей условию Липшица

Производная функции  $f'(x) = x^{-1/2} = 1/\sqrt{x}$ , поэтому в случае «а»  $L = \max_{[0,1]} |f'(x)| = |f'(0)| = 5$ , а в случае «б»  $L = \max_{[0,10]} |f'(x)| = |f'(10)| = 1/\sqrt{10} \approx 0,316$ .

**Пример 1.10.** Показать, что для дифференцируемой на отрезке  $[a, b]$  функции  $f(x)$  величина  $L = \max_{[a, b]} |f'(x)|$  представляет собой минимальную из констант Липшица  $f(x)$  на  $[a, b]$ .

Пусть существует  $L_1 < L = \max_{[a, b]} |f'(x)|$  такая, что  $|f(x') - f(x'')| \leq L_1 |x' - x''|$  для всех  $x', x'' \in [a, b]$ .

Тогда, фиксируя  $x'$  и переходя к пределу при  $x'' \rightarrow x'$ , получим  $|f'(x')| \leq L_1$ , а вследствие произвольности точки  $x' \in [a, b]$  получим  $\max_{[a, b]} |f'(x)| \leq L_1 < L$ . Полученное противоречие доказывает утверждение.

**Пример 1.11.** Найти наименьшую из констант Липшица функции  $f(x) = 4x^3 - 30x^2 + 72x + 12$  на отрезках а)  $x \in [0, 2]$ , б)  $x \in [2, 3]$ .

Производная функции  $f'(x) = 12x^2 - 60x + 72 = 12(x-2)(x-3)$ , поэтому в случае «а»  $L = \max_{[0,2]} |f'(x)| = |f'(0)| = 72$ , а в случае «б»  $L = \max_{[2,3]} |f'(x)| = |f'(2,5)| = 3$ .

**Пример 1.12.** Найти наименьшую из констант Липшица функции  $f(x) = x^3 + 6x^2 - 15x$  на отрезках а)  $x \in [0, 1]$ , б)  $x \in [0, 10]$ .

Производная функции  $f'(x) = 3x^2 + 12x - 15 = 3(x+5)(x-1)$ ,

поэтому в случае «а»  $L = \max_{[0,1]} |f'(x)| = |f'(0)| = 15$ ,

а в случае «б»  $L = \max_{[0,10]} |f'(x)| = |f'(10)| = 405$ .

## 1.6. Классическая минимизация функции одной переменной

Из математического анализа известны следующие условия локального экстремума функции  $f(x)$ , дифференцируемой достаточное количество раз.

1. Если функция  $f(x)$  дифференцируема в точке  $\tilde{x}$  и достигает в ней локального экстремума, то  $f'(\tilde{x}) = 0$  (необходимое условие экстремума).

2. Пусть функция  $f(x)$   $n$  раз дифференцируема в точке  $\tilde{x}$  и в этой точке все производные  $f(x)$  до  $(n-1)$ -го порядка включительно равны нулю, а  $f^{(n)}(\tilde{x}) \neq 0$ . Тогда, если  $n$  — нечетно, то точка  $\tilde{x}$  не является точкой локального экстремума функции  $f(x)$ . Если же  $n$  — четное число, то:

а) при  $f^{(n)}(\tilde{x}) > 0$   $\tilde{x}$  — точка локального минимума  $f(x)$ ;

б) при  $f^{(n)}(\tilde{x}) < 0$   $\tilde{x}$  — точка локального максимума  $f(x)$

(достаточные условия экстремума).

Перечисленные условия позволяют предложить следующий путь решения задачи минимизации (1.5):

1. с помощью условия 1 находим все точки возможного экстремума функции  $f(x)$  на интервале  $(a, b)$ , т.е. корни уравнения

$$f'(x) = 0, \quad (1.8)$$

(стационарные точки функции  $f(x)$ , принадлежащие интервалу  $(a, b)$ );

2. найденные стационарные точки исследуем в соответствии с условием 2, выделяя из них только точки локальных минимумов  $f(x)$ ;

3. значения  $f(x)$  в точках локальных минимумов и на концах отрезка  $[a, b]$  сравниваем между собой. Наименьшему из этих значений соответствует точка глобального минимума  $f(x)$  на  $[a, b]$ .

Применение условия 2 требует вычисления высших производных функции  $f(x)$ , поэтому в большинстве случаев бывает проще сравнить значения  $f(x)$  во всех стационарных точках, не интересуясь их характером. С учетом этого можно предложить следующий алгоритм минимизации  $f(x)$  на отрезке  $[a, b]$  (классический метод, который разберем на примере).

**Пример 1.13.** Решить задачу  $f(x) = x^3 - 3x + 1 \rightarrow \min, x \in [-2, 2]$ .

Шаг 1. Находим корни уравнения  $f'(x) = 3x^2 - 3 = 0$  из интервала  $(-2, 2)$ :  $x_1 = -1, x_2 = 1$ . Полагаем  $x_0 = -2, x_3 = 2$ .

Шаг 2. Вычисляем значения  $f(x)$  в точках  $x_i, i = 0, \dots, 3$ :  $f(x_0) = -17, f(x_1) = 3, f(x_2) = -1, f(x_3) = 1$ .

Шаг 3. Находим  $f^* = \min(-17, 3, -1, 1) = -17 = f(x_0)$ . Поэтому  $x^* = -2, f^* = -17$ .

**Пример 1.14.** Найти  $f(x) = x^3 - 27x + 5 \rightarrow \min, x \in [-4, 4]$ .

Найдем корни уравнения  $f'(x) = 3x^2 - 27 = 0$  из интервала  $x \in (-4, 4)$ :  $x_1 = -3, x_2 = 3$ . Положим  $x_0 = -4, x_3 = 4$ . Так как  $f''(x_1) = f''(-3) = -18 < 0$ , то  $x_1$  — точка локального максимума. Так как  $f''(x_2) = f''(3) = 18 > 0$ ,

то  $x_2$  — точка локального минимума.  $f(3) = -49, f(-4) = 49, f(4) = -39$ .

Производя перебор, получим  $x^* = 3, f^* = -49$ .

**Пример 1.15.** Решить задачу

$$f(x) = \sin(2\pi x) - 2\pi \exp(x-1) \rightarrow \min, x \in \left[\frac{1}{2}, \frac{3}{2}\right].$$

Найдем корни уравнения  $f'(x) = 2\pi \cos(2\pi x) - 2\pi \exp(x-1) = 0$ , из промежутка  $x \in [\frac{1}{2}, \frac{3}{2}]$ . Точка  $x_1 = 1$  является корнем (показать, что других корней на промежутке нет). Далее, полагая  $x_0 = \frac{1}{2}, x_2 = \frac{3}{2}$ , вычислим  $f(x_0) = -2\pi \exp(-\frac{1}{2}), f(x_1) = -2\pi, f(x_2) = -2\pi \exp(\frac{1}{2})$ . Производя перебор, получим  $x^* = x_2 = \frac{3}{2}, f^* = f(x_2) = -2\pi \exp(\frac{1}{2})$ .

**Пример 1.16.** Найти  $f(x) = -\sin(\frac{x}{2}) - \cos(\frac{x}{2}) \rightarrow \min, x \in [\frac{\pi}{3}, \frac{2\pi}{3}]$ .

Найдем корни уравнения

$$f'(x) = -\frac{1}{2}\cos(\frac{x}{2}) + \frac{1}{2}\sin(\frac{x}{2}) = 0, \quad x \in [\frac{\pi}{3}, \frac{2\pi}{3}].$$

Точка  $x_1 = \frac{\pi}{2}$  является корнем (показать, что других корней на промежутке нет). Положим  $x_0 = \frac{\pi}{3}, x_2 = \frac{2\pi}{3}$ . Производя перебор,  $f(x_0) = -\frac{1}{2} - \frac{\sqrt{3}}{2}, f(x_1) = -\sqrt{2}, f(x_2) = -\frac{\sqrt{3}}{2} - \frac{1}{2}$ . Поскольку  $1 + \sqrt{3} < 2\sqrt{2}$ , получим  $x^* = x_1 = \frac{\pi}{2}, f^* = f(x_1) = -\sqrt{2}$ .

### Вопросы и задания для самоконтроля

1. Какая функция называется целевой?
2. Дать определение локального и глобального минимумов функции.
3. Что такое точная нижняя грань функции на множестве?
4. Как соотносятся точная нижняя грань и минимум функции на множестве?
5. Какая функция называется унимодальной на отрезке  $[a, b]$ ?
6. Сформулировать свойства унимодальных функций.

7. Какая функция называется выпуклой на отрезке  $[a, b]$ ?
8. Каков геометрический смысл выпуклости функции?
9. Сформулировать два необходимых и достаточных дифференциальных условий выпуклости функций.
10. Сформулировать условие Липшица для функции  $f(x)$  на отрезке  $[a, b]$ .
11. Всякая ли унимодальная на отрезке  $[a, b]$  функция  $f(x)$  удовлетворяет на нем условию Липшица?
12. Всякая ли функция  $f(x)$ , удовлетворяющая условию Липшица на отрезке  $[a, b]$ , унимодальна на нем?
13. Сформулировать свойства функций, удовлетворяющих на отрезке  $[a, b]$  условию Липшица.
14. В чем заключается классический метод минимизации функций?
15. Для каких целей разработан классический метод минимизации функций?
16. Какова практическая ограниченность применимости этого метода?

## Глава 2

# ОДНОМЕРНАЯ МИНИМИЗАЦИЯ ФУНКЦИЙ. ПРЯМЫЕ МЕТОДЫ

Подробный разбор одномерных задач минимизации функций важен по двум причинам.

*Во-первых*, эти задачи наиболее просты, и на них легче понять постановку проблемы, методы решения и возникающие трудности. Иногда одномерные задачи имеют самостоятельный практический интерес.

*Во-вторых*, и это главное, алгоритмы решения многомерных задач оптимизации часто сводятся к последовательному многократному решению одномерных задач и не могут быть поняты без умения их решать.

Пусть необходимо минимизировать функцию  $f(x)$  на отрезке  $[a, b]$ . Обычно для этого применяются приближенные, т.е. численные методы. Решение задачи находится с необходимой точностью в результате определения конечного числа значений функции  $f(x)$  и ее производных в некоторых точках отрезка  $[a, b]$ .

## 2.1. О прямых методах

Методы, использующие только значения функции и не требующие вычисления ее производных, называются *прямыми методами* минимизации. Большим достоинством прямых методов является то, что от целевой функции не требуется дифференцируемости. Более того, она может быть не задана в аналитическом виде. Единственное, на чем основаны алгоритмы прямых методов минимизации, это возможность определения значений  $f(x)$  в заданных точках.

**Пример 2.1.** Химический завод производит некоторое вещество. Выход интересующего продукта определяется температурой  $y = f(T)$ ,

которую можно варьировать в пределах:  $T_1 \leq T \leq T_2$ . Вид функции  $f(T)$  заранее не известен, он зависит от используемого сырья. Получив очередную партию сырья, нужно найти температуру  $T$ , при которой наиболее выгодно вести производство, т.е. наибольшее значение  $f(T)$ .

**Пример 2.2.** Значения функции  $f(x)$  получаются в результате сложных численных расчетов.

Рассмотрим наиболее распространенные прямые методы поиска точки минимума. Самым слабым требованием на функцию  $f(x)$ , позволяющим использовать эти методы, является ее унимодальность. Поэтому далее будем считать функцию  $f(x)$  унимодальной на отрезке  $[a, b]$ .

Будут подробно рассмотрены следующие четыре прямых метода:

- перебора и поразрядного поиска;
- деления отрезка пополам (метод дихотомии и метод золотого сечения).

Кроме того, будет дано представление о методах полиномиальной аппроксимации (метод парабол).

## 2.2. Метод перебора

Является простейшим из прямых методов и состоит в следующем. Разобьем отрезок  $[a, b]$  на  $n$  равных частей точками деления  $x_i = a + i(b-a)/n$ ,  $i = 0, \dots, n$ . Вычислив значения  $f(x)$  в точках  $x_i$ , путем сравнения найдем точку  $x_m$ ,  $0 \leq m \leq n$ , для которой

$$f(x_m) = \min_{0 \leq i \leq n} f(x_i). \quad (2.1)$$

Далее, положим  $x^* \approx x_m$ ,  $f^* \approx f(x_m)$ .

Необходимо отметить, что погрешность определения точки минимума  $x^*$  функции  $f(x)$  методом перебора не превосходит величины  $\varepsilon_n = (b-a)/n$ .

Предположим, что  $x_m$  из (2.1) является внутренней точкой разбиения отрезка  $[a, b]$ , т.е.  $1 \leq m \leq n-1$  (случаи  $m=0$  и  $m=n$  рассматриваются аналогично). Тогда из соотношения (2.1) и по определению унимодальности функции следует

$$a) \ f(x_{m-1}) \geq f(x_m), \text{ т.е. } x^* \in [x_{m-1}, b];$$

$$б) \ f(x_m) \leq f(x_{m+1}), \text{ т.е. } x^* \in [a, x_{m+1}].$$

Отсюда получаем, что  $x^* \in [x_{m-1}, b] \cap [a, x_{m+1}] = [x_{m-1}, x_{m+1}]$ .

Длина последнего отрезка равна  $2(b-a)/n$ , а точка  $x_m$  является его серединой, поэтому  $|x_m - x^*| \leq (b-a)/n = \varepsilon_n$ .

Таким образом, чтобы обеспечить требуемую точность  $\varepsilon$  определения точки  $x^*$ , число отрезков разбиения  $n$  необходимо выбрать из условия  $\varepsilon_n = (b-a)/n \leq \varepsilon$ , т.е.  $n \geq (b-a)/\varepsilon$ .

**Замечание.** Пусть реализация метода перебора потребовала  $N$  вычислений функции  $f(x)$ . Это означает, что отрезок  $[a, b]$  был разбит на  $n = N - 1$  частей и достигнутая

точность определения  $x^*$  составила  $\varepsilon_n = \varepsilon_{N-1} = \frac{b-a}{N-1}$ . Поэто-

му точность решения  $\varepsilon(N)$ , которую обеспечивает метод перебора в результате  $N$  вычислений  $f(x)$ , будет

$$\varepsilon(N) = \frac{b-a}{N-1}. \quad (2.2)$$

**Пример 2.3.** Методом перебора решить задачу  $f(x) = x^4 + e^{-x} \rightarrow \min, x \in [0, 1]$ , с точностью до  $\varepsilon = 0,1$ .

Функция  $f(x)$  унимодальна на отрезке  $[0, 1]$ . Найдем число  $n$  отрезков разбиения:  $n \geq \frac{1-0}{0,1} = 10$ , т.е. можно взять  $n = 10$ . Вычислим значения  $f(x_i)$ , где  $x_i = 0,1i, i = 0, \dots, 10$  и запишем их в табл. 2.1.

Таблица 2.1.

Значения функции в узловых точках при методе перебора

| $x_i$    | 0,0  | 0,1  | 0,2  | 0,3  | 0,4  | 0,5   | 0,6  | 0,7  | 0,8  | 0,9  | 1,0  |
|----------|------|------|------|------|------|-------|------|------|------|------|------|
| $f(x_i)$ | 1,00 | 0,90 | 0,82 | 0,75 | 0,70 | 0,670 | 0,68 | 0,74 | 0,86 | 1,06 | 1,37 |

В этой таблице минимальное вычисленное значение  $f(x)$  равно 0,67.

Таким образом,  $x^* \approx 0,5, f^* \approx 0,67$ .

## 2.3. Метод поразрядного поиска

Рассмотрим усовершенствованный метод перебора с целью уменьшения количества значений  $f(x)$ , которые необходимо находить в процессе минимизации.

Во-первых, если  $f(x_i) \leq f(x_{i+1})$ , то отпадает необходимость вычислять  $f(x)$  в точках  $x_{i+2}, x_{i+3}$  и т.д., так как из унимодальности функции  $f(x)$  следует, что  $x^* \leq x_{i+1}$ .

Во-вторых, разумно было бы сначала определить отрезок, содержащий точку  $x^*$  с небольшой точностью, а затем искать ее на этом отрезке с меньшим шагом дискретизации, повышая точность.

Такая возможность улучшения метода перебора реализована в методе поразрядного поиска, в котором перебор точек отрезка происходит сначала с шагом  $\Delta = x_{i+1} - x_i > \varepsilon$  до тех пор, пока не выполнится условие  $f(x_i) \leq f(x_{i+1})$  или пока очередная из этих точек не совпадет с концом отрезка. После этого шаг уменьшается (обычно в четыре раза), и перебор точек с новым шагом производится в противоположном направлении до тех пор, пока значения  $f(x)$  не перестанут уменьшаться или очередная точка не совпадет с концом отрезка и т.д. Процесс завершается, когда перебор в данном направлении закончен, а использованный при этом шаг дискретизации не превосходит  $\varepsilon$ .

**Пример 2.4.** Методом поразрядного поиска найти  $f(x) = x^4 + e^{-x} \rightarrow \min, x \in [0, 1]$ , с точностью до  $\varepsilon = 0,1$ .

Начальный шаг  $\Delta = 0,25$ . Вычисляя последовательно значения

$f(x)$  в точках дискретизации с шагом 0,25, получим

| $x$    | $0 \rightarrow 0,25 \rightarrow 0,50 \rightarrow 0,75$ |
|--------|--|
| $f(x)$ | $1,000 > 0,783 > 0,669 < 0,789$                        |

Так как  $f(0,50) < f(0,75)$ , причем  $|\Delta| > \varepsilon$ , то поиск  $x^*$  продолжаем из начальной точки  $x_0 = 0,75$ , изменив его направление и уменьшив шаг в четыре раза

| $x$    | $0,4375 \leftarrow 0,5000 \leftarrow 0,5625 \leftarrow 0,6250 \leftarrow 0,6875 \leftarrow 0,750$ |
|--------|---|
| $f(x)$ | $0,682 > 0,669 < 0,670 < 0,688 < 0,726 < 0,789$   |

Так как  $|\Delta| = 0,0625 < \varepsilon$ , то поиск завершен и  $x^* \approx 0,5, f^* \approx 0,67$  (сравнить с результатом решения примера методом перебора; здесь потребовалось только 8, а не 11 вычисленных значений функции  $f(x)$ ).

Необходимо отметить, что метод поразрядного поиска, несмотря на простоту и меньшую эффективность, чем рассматриваемые ниже методы деления отрезка пополам, имеет значительную область применимости в методах многомерной минимизации. Это связано с тем, что при реализации этого метода не обязательно задавать границы отрезка  $[a, b]$ , роль



которого может играть вся бесконечная прямая или любая ее часть.

В методе перебора точки  $x_i$ , в которых определяются значения  $f(x)$ , выбираются заранее (*прямые методы пассивного поиска*). Если же для выбора очередной точки вычисления  $f(x)$  использовать информацию, содержащуюся в уже найденных ее значениях, то поиск точки минимума можно сделать более эффективным, т.е. сократить число определяемых для этого значений  $f(x)$  (*прямые методы последовательного поиска*). Метод поразрядного поиска является одним из таких методов.

Один из путей более эффективного поиска точки  $x^*$  следует из определения унимодальных функций. Действительно, пусть  $a < x_1 < x_2 < b$ . Сравнив значения  $f(x)$  в пробных точках  $x_1$  и  $x_2$ , можно сократить отрезок поиска точки  $x^*$ , перейдя к отрезку  $[a, x_2]$ , если  $f(x_1) \leq f(x_2)$ , или к отрезку  $[x_1, b]$ , если  $f(x_1) \geq f(x_2)$ . Описанную процедуру можно повторить необходимое число раз, последовательно уменьшая отрезок, содержащий точку минимума.

Чтобы относительное уменьшение отрезка на каждой итерации не зависело от того, какая из его частей исключается из дальнейшего рассмотрения, пробные точки имеет смысл располагать симметрично относительно середины исходного отрезка. В зависимости от способа выбора пробных точек в этом случае получаются различные методы исключения отрезков. Рассмотрим только методы дихотомии и золотого сечения.

## 2.4. Метод дихотомии

В этом методе пробные точки  $x_1$  и  $x_2$  располагаются близко к середине очередного отрезка  $[a, b]$ , т.е.

$$x_1 = \frac{b+a-\delta}{2}, \quad x_2 = \frac{b+a+\delta}{2}, \quad (2.3)$$

где  $\delta$  — малое число. При этом отношение длин нового и исходного отрезков  $\tau = \frac{b-x_1}{b-a} = \frac{x_2-a}{b-a}$  близко к  $\frac{1}{2}$ , этим и объясняется название метода.

Отметим, что для любых точек  $x_1$  и  $x_2$  величина  $\tau > \frac{1}{2}$ ,

поэтому указанный выбор пробных точек объясняется стремлением обеспечить максимально возможное относительное уменьшение отрезка на каждой итерации поиска  $x^*$ .

В конце вычислений методом дихотомии в качестве приближенного значения  $x^*$  берут середину последнего из найденных отрезков  $[a, b]$ , убедившись предварительно, что

$$\text{достигнуто неравенство } \frac{b-a}{2} \leq \varepsilon.$$

Алгоритм метода деления отрезка пополам следующий.

Шаг 1. Определить  $x_1$  и  $x_2$  по формуле (2.3) и вычислить  $f(x_1)$  и  $f(x_2)$ . Перейти к шагу 2.

Шаг 2. Сравнить  $f(x_1)$  и  $f(x_2)$ . Если  $f(x_1) \leq f(x_2)$ , то перейти к отрезку  $[a, x_2]$ , положив  $b = x_2$ , иначе — к отрезку  $[x_1, b]$ , положив  $a = x_1$ . Перейти к шагу 3.

Шаг 3. Найти достигнутую точность  $\varepsilon_n = \frac{b-a}{2}$  (здесь

$n$  — номер итерации). Если  $\varepsilon_n > \varepsilon$ , то перейти к следующей итерации, вернувшись к шагу 1. Если  $\varepsilon_n \leq \varepsilon$ , то завершить поиск  $x^*$ , перейдя к шагу 4.

Шаг 4. Положить  $x^* \approx \bar{x} = \frac{a+b}{2}$ ,  $f^* \approx f(\bar{x})$ .

Необходимо обратить внимание на следующее.

1. Число  $\delta$  из (2.3) выбирается на интервале  $(0, 2\varepsilon)$  с учетом следующих соображений:

а) чем меньше  $\delta$ , тем больше относительное уменьшение длины отрезка на каждой итерации;

б) при чрезмерно малом  $\delta$  сравнение значений функции  $f(x)$  в точках  $x_1$  и  $x_2$ , отличающихся на величину  $\delta$ , затруднительно, поэтому необходимо учитывать количество верных разрядов функции  $f(x)$  при задании аргумента  $x$ .

2. Число  $n$  итераций метода дихотомии, необходимое для определения точки  $x^*$  с точностью  $\varepsilon$ , определяется неравенством

$$n \geq \log_2 \frac{b-a-\delta}{2\varepsilon-\delta} \quad (2.4)$$

Обозначим длину  $[a, b] = \Delta_0$ . Тогда длина отрезка, полученного после первой итерации, равна  $\Delta_1 = \frac{\Delta_0}{2} + \frac{\delta}{2}$ , после второй итерации  $\Delta_2 = \frac{\Delta_1}{2} + \frac{\delta}{2} = \frac{b-a}{4} + \delta(\frac{1}{4} + \frac{1}{2})$ , после третьей  $\Delta_3 = \frac{\Delta_2}{2} + \frac{\delta}{2} = \frac{b-a}{8} + \delta(\frac{1}{8} + \frac{1}{4} + \frac{1}{2})$  и т.д. После  $n$  итераций

длина отрезка поиска точки  $x^*$  станет равна

$$\Delta_n = \frac{b-a}{2^n} + \delta(\frac{1}{2^n} + \frac{1}{2^{n-1}} + \dots + \frac{1}{2}) = \frac{b-a}{2^n} + (1 - \frac{1}{2^n})\delta.$$

При этом будет достигнута точность определения точки минимума  $\epsilon_n = \frac{\Delta_n}{2}$ . Находя  $n$  из условия  $\epsilon_n = \frac{b-a}{2^{n+1}} + (1 - \frac{1}{2^n})\frac{\delta}{2} \leq \epsilon$ , получим неравенство (2.4).

3. Величина  $\delta$  не может быть выбрана слишком большой. Так как  $\epsilon_n = \frac{b-a-\delta}{2^{n+1}} + \frac{\delta}{2} \geq \frac{\delta}{2}$ , но  $\epsilon_n \leq \epsilon$ , необходимо выбирать  $\delta < 2\epsilon$ .

4. При малом значении  $\delta$  величина  $\epsilon_n = \frac{b-a-\delta}{2^{n+1}} + \frac{\delta}{2} \approx \frac{b-a}{2^{n+1}}$ .

На каждой итерации метода дихотомии вычисляются два значения  $f(x)$ . Поэтому после  $N$  вычислений  $f(x)$  производится  $n = N/2$  итераций и достигается точность определения  $x^*$ :

$$\epsilon(N) = \epsilon_{\frac{N}{2}} \approx \frac{b-a}{2^{\frac{N}{2}+1}}. \quad (2.5)$$

**Пример 2.5.** Методом деления отрезка пополам решить задачу  $f(x) = x^4 + e^{-x} \rightarrow \min, x \in [0, 1]$ , с точностью  $\epsilon = 0,1$ . Выберем  $\delta = 0,02$ , ( $\delta < 2\epsilon$ ).

Рассмотрим детально действия при итерации 1.

Шаг 1.  $x_1 = 0,49$ ,  $x_2 = 0,51$ ,  $f(x_1) = 0,688$ ,  $f(x_2) = 0,670$ .

Шаг 2.  $f(x_1) > f(x_2)$ , поэтому полагаем  $a = x_1 = 0,49$ .

Шаг 3.  $\frac{b-a}{2} = 0,255 > 0,1$ , поэтому переходим к следующей итерации.

Исходные данные и результаты вычислений итераций приведены в табл. 2.2.

Таблица 2.2.

Результаты вычислений на итерациях 1 – 4

| Номер итерации | $a$  | $b$   | $(b-a)/2$ | $x_1$                            | $x_2$ | $f(x_1)$ | $f(x_2)$ | Сравнение         |
|----------------|------|-------|-----------|----------------------------------|-------|----------|----------|-------------------|
| 1              | 0    | 1     | 0,5       | 0,49                             | 0,51  | 0,688    | 0,670    | $f(x_1) > f(x_2)$ |
| 2              | 0,49 | 1     | 0,26      | 0,735                            | 0,755 | 0,771    | 0,792    | $f(x_1) < f(x_2)$ |
| 3              | 0,49 | 0,755 | 0,13      | 0,613                            | 0,633 | 0,683    | 0,691    | $f(x_1) < f(x_2)$ |
| 4              | 0,49 | 0,633 | 0,07      | 0,07 < 0,1 — точность достигнута |       |          |          |                   |

Откуда  $x^* \approx \frac{0,49+0,633}{2} \approx 0,56$ ,  $f^* \approx f(0,56) \approx 0,67$  (сравнить с

расчетом примера методами перебора и поразрядного поиска; вычислено 7 значений  $f(x)$ ).

## 2.5. Метод золотого сечения

Рассмотрим такое симметричное расположение точек  $x_1$  и  $x_2$  на отрезке  $[a, b]$ , при котором одна из них становится пробной точкой на новом отрезке, полученном после исключения части исходного отрезка. Использование таких точек позволяет на каждой итерации метода золотого сечения, кроме первой, ограничиться определением только одного значения  $f(x)$ , так как другое значение уже найдено на одной из предыдущих итераций.

Найдем точки  $x_1$  и  $x_2$ , обладающие указанным свойством.

Рассмотрим сначала отрезок  $[0, 1]$  и для определенности предположим, что при его уменьшении исключается правая часть этого отрезка. Пусть  $x_2 = \tau$ , тогда симметрично расположенная точка  $x_1 = 1 - \tau$  (рис.2.1).



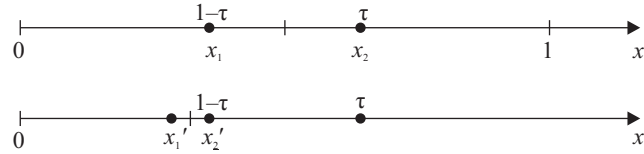


Рис. 2.1. Иллюстрация выбора пробных точек в методе золотого сечения

Пробная точка  $x_1$  отрезка  $[0, 1]$  перейдет в пробную точку  $x_2' = 1 - \tau$  нового отрезка  $[0, \tau]$ . Чтобы точки  $x_2 = \tau$  и  $x_2' = 1 - \tau$  делили отрезки  $[0, 1]$  и  $[0, \tau]$  в одном и том же отношении,

должно выполняться равенство  $\frac{1}{\tau} = \frac{\tau}{1 - \tau}$ , или  $\tau^2 = 1 - \tau$ , отку-

да находим положительное значение  $\tau = \frac{\sqrt{5} - 1}{2} \approx 0,61803 \dots$

Таким образом,

$$x_1 = 1 - \tau = \frac{3 - \sqrt{5}}{2}, \quad x_2 = \tau = \frac{\sqrt{5} - 1}{2}.$$

Для произвольного отрезка  $[a, b]$  выражения для пробных точек примут вид

$$x_1 = a + \frac{3 - \sqrt{5}}{2}(b - a), \quad x_2 = a + \frac{\sqrt{5} - 1}{2}(b - a). \quad (2.6)$$

На каждой итерации отрезок поиска точки минимума уменьшается в одном и том же отношении  $\tau = \frac{\sqrt{5} - 1}{2}$ , поэтому

в результате  $n$  итераций его длина становится  $\Delta_n = \tau^n(b - a)$ . Таким образом, точность определения точки  $x^*$  после  $n$  итераций находится из равенства

$$\epsilon_n = \frac{\Delta_n}{2} = \frac{1}{2} \left( \frac{\sqrt{5} - 1}{2} \right)^n (b - a), \quad (2.7)$$

а условием окончания поиска точки  $x^*$  с точностью  $\epsilon$  служит неравенство  $\epsilon_n \leq \epsilon$ .

Необходимо обратить внимание на следующее.

1. Точки  $x_1$  и  $x_2$  из (2.6) обладают следующим свойством: каждая из них делит отрезок  $[a, b]$  на две неравные части так, что отношение длины всего отрезка к длине его большей части равно отношению длин большей и меньшей части отрезка. Точки с таким свойством называются *точками золотого сечения* отрезка  $[a, b]$ . Это и объясняет название рассматриваемого метода.

2. На каждой итерации исключения отрезков с пробными точками (2.6) одна из них переходит на следующий отрезок и значение  $f(x)$  в ней уже известно.

3. Легко проверить, что  $x_1 = a + b - x_2$  и  $x_2 = a + b - x_1$ . Поэтому на каждой итерации метода золотого сечения недостающую пробную точку нового отрезка можно найти по перешедшей на него пробной точке с помощью сложения и вычитания, не используя (2.6).

4. В конце вычислений по методу золотого сечения в качестве приближенного значения  $x^*$  можно взять середину

последнего из полученных отрезков  $\bar{x} = \frac{a + b}{2}$ .

5. Число итераций, необходимое для достижения заданной точности  $\epsilon$ , можно найти из условия  $\epsilon_n \leq \epsilon$  с учетом соотношения (2.7)

$$n \geq \ln \left( \frac{2\epsilon}{b - a} \right) / \ln \tau = 2,1 \ln \left( \frac{b - a}{2\epsilon} \right).$$

Так как  $N$  вычислений  $f(x)$  позволяют выполнить  $N - 1$  итераций метода золотого сечения, то достигнутая в результате этого точность определения  $x^*$  составляет

$$\epsilon(N) = \epsilon_{N-1} = \frac{1}{2} \left( \frac{\sqrt{5} - 1}{2} \right)^{N-1} (b - a). \quad (2.8)$$

Алгоритм метода золотого сечения следующий.

Шаг 1. Определить  $x_1$  и  $x_2$  по формуле (2.6). Вычислить

$f(x_1)$  и  $f(x_2)$ . Положить  $\tau = \frac{\sqrt{5} - 1}{2}$ ,  $\epsilon_n = \frac{b - a}{2}$ . Перейти к

шагу 2.

Шаг 2. Проверка окончания поиска: если  $\varepsilon_n > \varepsilon$ , то перейти к шагу 3, иначе — к шагу 4.

Шаг 3. Переход к новому отрезку и новым пробным точкам. Если  $f(x_1) \leq f(x_2)$ , то положить

$$b = x_2, \quad x_2 = x_1, \quad f(x_2) = f(x_1), \quad x_1 = b - \tau(b - a)$$

и вычислить  $f(x_1)$ , иначе — положить

$$a = x_1, \quad x_1 = x_2, \quad f(x_1) = f(x_2), \quad x_2 = b - \tau(b - a)$$

и вычислить  $f(x_2)$ . Положив  $\varepsilon_n = \tau\varepsilon_n$ , перейти к шагу 2.

Шаг 4. Окончание поиска: положить

$$x^* \approx \bar{x} = \frac{a+b}{2}, \quad f^* \approx f(\bar{x}).$$

## 2.6. Сравнение методов перебора, дихотомии и золотого сечения

При сравнении прямых методов минимизации обычно учитывают  $N$  — количество вычисленных значений  $f(x)$ , гарантирующее заданную точность определения  $x^*$ . Чем меньше  $N$ , тем более эффективным считается метод. Вспомогательные операции, такие как выбор пробных точек, сравнение значений функции  $f(x)$  и прочее, не учитываются. Во многих практических случаях определение значений целевой функции требует больших затрат (например, времени ЭВМ или средств для проведения экспериментов).

Эффективность методов минимизации можно также сравнивать по гарантированной точности  $\varepsilon(N)$  нахождения точки  $x^*$ , которую они обеспечивают в результате определения  $N$  значений  $f(x)$ .

Из анализа формул для  $\varepsilon(N)$  рассмотренных методов следует, что наиболее эффективным является метод золотого сечения (происходит исключение отрезков и необходимо выбирать только одну пробную точку на итерации). Далее идет метод дихотомии (происходит деление отрезка почти пополам, но необходим выбор двух пробных точек на итерации). Наименее эффективным является метод перебора — прямой метод пассивного поиска, при котором исключения отрезков не применяется вовсе. Эти выводы иллюстрирует табл. 2.3.

Таблица 2.3.

Значения точности  $\varepsilon(N)$  в зависимости от количества  $N$  найденных значений  $f(x)$  на отрезке длины 1 для трех из рассмотренных методов.

| Методы минимизации     | Количество найденных значений $f(x)$ |                     |                     |                      |
|------------------------|--------------------------------------|---------------------|---------------------|----------------------|
|                        | $N = 5$                              | $N = 11$            | $N = 21$            | $N = 51$             |
| Метод золотого сечения | 0,073                                | $4,1 \cdot 10^{-3}$ | $3,3 \cdot 10^{-5}$ | $1,8 \cdot 10^{-11}$ |
| Метод дихотомии        | 0,125                                | $1,6 \cdot 10^{-2}$ | $4,9 \cdot 10^{-4}$ | $1,5 \cdot 10^{-8}$  |
| Метод перебора         | 0,250                                | 0,100               | 0,050               | 0,020                |

Наряду с методами, рассмотренными в таблице, еще раз отметим метод поразрядного поиска — эффективный и не требующий задания фиксированных границ отрезка  $[a, b]$ . Его часто применяют в задачах многомерной минимизации на интервалах неопределенной (например, бесконечной) длины.

**Пример 2.6.** Сравнить необходимые количества вычисленных значений  $N_d$  и  $N_n$  функции  $f(x)$  при поиске ее точки минимума на отрезке длины 1 с точностью  $10^{-5}$  методом деления отрезка пополам и методом перебора.

В методе перебора  $n \geq (b-a)/\varepsilon$ , подставляя числа, получим  $N_n = 10^5$ .

В методе дихотомии  $n \geq \log_2 \frac{b-a-\delta}{2\varepsilon-\delta} \approx \log_2 \frac{b-a}{2\varepsilon}$ . Подставляя числа,

получим  $N_d = 2\log_2(10^5/2) \approx 32$ , откуда  $N_n/N_d = 3125$ .

## 2.7. Метод парабол

Поиск точки минимума методами исключения отрезков основан на сравнении значений функции  $f(x)$  в двух точках, при котором учитывается только знак разности этих значений.

Учесть информацию, содержащуюся в относительных изменениях значений  $f(x)$  в пробных точках, позволяет *метод полиномиальной интерполяции*, идея которого состоит в том,

что для  $f(x)$  строится аппроксимирующий многочлен, и его точка минимума служит приближением к  $x^*$ .

Этот метод можно использовать при условии, что  $f(x)$  является не только унимодальной, но и достаточно гладкой. Метод опирается на доказанную в курсе математического анализа теорему Вейерштрасса об аппроксимации, согласно которой достаточно гладкую на отрезке функцию можно с любой точностью приблизить на нем некоторым полиномом.

Для повышения точности можно, во-первых, увеличивать порядок аппроксимирующего полинома и, во-вторых, уменьшать длину отрезка аппроксимации. Второй путь в нашем случае предпочтительней.

Простейший вариант полиномиальной аппроксимации — метод парабол, использующий полиномы второго порядка. На каждой итерации этого метода строится квадратный трехчлен, график которого (парабола) проходит через три выбранные точки графика функции  $f(x)$  (рис. 2.2).

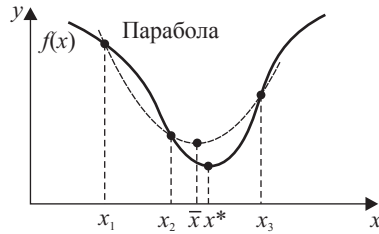


Рис. 2.2. Иллюстрация применения метода парабол

Рассмотрим унимодальную на отрезке  $[a, b]$  функцию  $f(x)$ , достигающую минимума во внутренней точке этого отрезка. Выберем три точки  $x_1, x_2$  и  $x_3$  отрезка  $[a, b]$ , для которых выполняются неравенства

$$x_1 < x_2 < x_3, \quad f(x_1) \geq f(x_2) \leq f(x_3). \quad (2.9)$$

Для определения таких точек, как правило, бывает достаточно нескольких проб. Можно также совершать итерации методом золотого сечения до тех пор, пока для пробных точек

очередного отрезка и одного из его концов не станут выполняться неравенства (2.9).

Из унимодальности функции  $f(x)$  следует, что  $x^* \in [x_1, x_3]$ . Составим квадратный трехчлен

$$q(x) = a_0 + a_1(x - x_1) + a_2(x - x_1)(x - x_2)$$

(полином в форме Ньютона), график которого проходит через точки  $(x_1, f(x_1))$ ,  $(x_2, f(x_2))$ ,  $(x_3, f(x_3))$  графика функции  $f(x)$ . Будем считать, что хотя бы одно из неравенств для  $f(x)$  в (2.9) является строгим. Тогда ветви искомого параболы будут направлены вверх, а точка минимума  $\bar{x}$  трехчлена  $q(x)$  будет принадлежать отрезку  $[x_1, x_3]$ .

Определяя коэффициенты  $a_0, a_1$  и  $a_2$  из системы уравнений

$$q(x_1) = f(x_1) = f_1,$$

$$q(x_2) = f(x_2) = f_2,$$

$$q(x_3) = f(x_3) = f_3,$$

находим

$$a_0 = f_1, \quad a_1 = \frac{f_2 - f_1}{x_2 - x_1}, \quad a_2 = \frac{1}{x_3 - x_2} \left( \frac{f_3 - f_1}{x_3 - x_1} - \frac{f_2 - f_1}{x_2 - x_1} \right). \quad (2.10)$$

Точку минимума  $\bar{x}$  вычислим, приравняв производную квадратного трехчлена к нулю. В результате получим

$$\bar{x} = \frac{1}{2} \left( x_1 + x_2 - \frac{a_1}{a_2} \right), \quad (2.11)$$

где  $a_1$  и  $a_2$  определяются из (2.10). Число  $\bar{x}$  из (2.11) служит очередным приближением метода парабол к  $x^*$ . Далее описанная процедура повторяется для новых точек  $x_1, x_2, x_3$ , удовлетворяющих неравенствам (2.9).

Заметим, что на каждой итерации метода парабол, кроме первой, определяется только одно новое значение  $f(x)$ .

Условием окончания поиска минимума является близость к нулю разности  $\Delta$  чисел  $\bar{x}$ , найденных на данной и предыдущей итерациях, т.е. неравенство  $|\Delta| \leq \varepsilon$ .

**Пример 2.7.** Методом парабол решить задачу  $f(x) = x^4 + e^{-x} \rightarrow \min, \quad x \in [0, 1]$  с точностью  $|\Delta| \leq \varepsilon = 0,0025$ .

Рассмотрим детально действия при выполнении всех итераций.

Итерация 1.

Шаг 1. Выберем точки  $x_1=0,25, x_2=0,5, x_3=0,75$ . Функция принимает в этих точках соответственно значения  $f_1=0,7827, f_2=0,6690, f_3=0,7888$ , удовлетворяющие неравенствам (2.9). Переходим к шагу 2.

Шаг 2. По формулам (2.10) и (2.11) находим  $\bar{x}=0,4968$ . Поскольку на первой итерации вычисление величины  $\Delta$  невозможно, переходим сразу к шагу 4.

Шаг 4. Вычисляем  $f(\bar{x})=0,6694$ . Переходим к шагу 5.

Шаг 5. На данной итерации имеем:  $x_1 < \bar{x} < x_3, f(\bar{x}) \geq f(x_2)$ , следовательно  $x^* \in [\bar{x}, x_3]$ . Поэтому полагаем

$$x_1 = \bar{x} = 0,4968, f(x_1) = f(\bar{x}) = 0,6694,$$

а точки  $x_2, x_3$  и значения  $f(x)$  в них не изменяются. Переходим к следующей итерации.

Итерация 2.

Поскольку точки  $x_1, x_2$  и  $x_3$  уже найдены, то начинаем с шага 2.

Шаг 2. Находим  $\bar{x}=0,5224$ . Переходим к шагу 3.

Шаг 3.  $\Delta = |0,4968 - 0,5224| = 0,026 > 0,025$ , поэтому переходим к шагу 4.

Шаг 4. Вычисляем  $f(\bar{x})=0,6676$ . Переходим к шагу 5.

Шаг 5. Поскольку  $x_1 < x_2 < \bar{x} < x_3, f(x_2) \geq f(\bar{x})$ , следовательно  $x^* \in [x_2, x_3]$ , поэтому полагаем  $x_1 = x_2 = 0,5, f(x_1) = f(x_2) = 0,6690, x_2 = \bar{x} = 0,5224, f(x_2) = f(\bar{x}) = 0,6676$ , а точка  $x_3$  и значение  $f(x_3)$  остаются прежними. Переходим к следующей итерации.

Итерация 3.

Поскольку точки  $x_1, x_2$  и  $x_3$  уже найдены, то начинаем с шага 2.

Шаг 2. Находим  $\bar{x}=0,5248$  и переходим к шагу 3.

Шаг 3. Определяем  $\Delta = |0,5224 - 0,5248| = 0,024 < 0,025$ , т.е. требуемая точность достигнута, поэтому полагаем  $x^* \approx \bar{x} \approx 0,525$ .

Отметим, что в результате пяти вычислений функции  $f(x)$  (три — на итерации 1, и по одному — на итерациях 2 и 3) точка  $x^*$  была найдена с весьма высокой точностью (точное до четвертого знака значение  $x^* = 0,5283$ ).

## Вопросы и задания для самоконтроля

1. Пусть  $f(x)$  — дифференцируемая унимодальная на отрезке  $[a, b]$  функция, причем  $|f'(x)| \leq M$ . Оценить точность  $\Delta(N)$  при определении минимального значения  $f^*$  методом перебора в результате  $N$  вычислений  $f(x)$ .

2. Может ли оценка  $\epsilon(N) = \frac{b-a}{N-1}$  для точности определения  $x^*$

методом перебора нарушаться для функций, не являющихся унимодальными? Ответ пояснить рисунком.

3. Повысится ли эффективность метода поразрядного поиска, если шаг поиска  $\Delta$  последовательно уменьшать не в четыре, а в какое-либо другое число раз?

4. Может ли применение методов исключения отрезков привести к неверному определению  $x^*$ , если функция  $f(x)$  не унимодальна? Ответ пояснить рисунком.

5. Зависит ли точность определения  $x^*$ , которую гарантируют методы дихотомии и золотого сечения в результате  $N$  вычислений  $f(x)$ , от конкретной функции  $f(x)$ ?

6. Требуется найти точку минимума унимодальной функции на отрезке длины 1 с точностью  $\epsilon = 0,02$ . Имеется возможность измерить не более 10 значений  $f(x)$ . Какой из прямых методов минимизации можно использовать для этого?

7. Доказать, что погрешность определения точки минимума  $x^*$  функции  $f(x)$  методом перебора не превосходит величины  $\epsilon_n = (b-a)/n$ .

8. Доказать, что в методе дихотомии число итераций, необходимое для определения точки минимума с точностью  $\epsilon$ , определя-

ется формулой  $n \geq \log_2 \frac{b-a-\delta}{2\epsilon-\delta}$ .

9. Доказать, что число итераций, необходимое для достижения заданной точности  $\epsilon$  на отрезке  $[a, b]$  в методе золотого сечения

определяется формулой  $n \geq \ln \left( \frac{2\epsilon}{b-a} \right) / \ln \tau \approx 2,1 \ln \left( \frac{b-a}{2\epsilon} \right)$ .

10. Сравнить необходимые количества вычисленных значений  $N_d$  и  $N_n$  функции  $f(x)$  при поиске ее точки минимума на отрезке длины 1 с точностью 10-5 методом деления отрезка пополам и методом перебора.

11. Зависит ли точность определения  $x^*$ , которая получается методом парабол в результате  $N$  вычислений функции  $f(x)$ , от конкретной функции  $f(x)$ ?

12. Указать класс функций, для точного определения точек минимума которых достаточно одной итерации метода парабол.

13. В окрестности точки минимума  $x^*$  график  $f_1(x)$  близок к симметричному относительно вертикальной оси, проходящей через точку  $x^*$ , а график  $f_2(x)$  заметно асимметричен. Для какой из этих функций следует ожидать более высокой скорости сходимости, применяя метод парабол?

### Задание для численной реализации в среде программирования MATLAB

1. Написать в среде MATLAB функции, реализующие следующие пять методов: перебора, поразрядного поиска, дихотомии, золотого сечения и парабол.

2. Выбрать для выполнения работы тестовую функцию, номер которой соответствует номеру Вашего компьютера. Например, для компьютера №3 это будет функция 3), для компьютера №13 — функция 4):  $13-9=4$ ; для компьютера №23 это будет функция 5):  $23-9 \times 2=5$ .

1)  $f(x) = x^3 - 3\sin x \rightarrow \min, \quad x \in [0, 1].$

2)  $f(x) = x^4 + x^2 + x + 1 \rightarrow \min, \quad x \in [-1, 0].$

3)  $f(x) = e^x + \frac{1}{x} \rightarrow \min, \quad x \in [0, 5, 1, 5].$

4)  $f(x) = x^2 - 2x + e^{-x} \rightarrow \min, \quad x \in [-1, 1, 5].$

5)  $f(x) = x \sin x + 2 \cos x \rightarrow \min, \quad x \in [-6, -4].$

6)  $f(x) = x + \frac{1}{x^2} \rightarrow \min, \quad x \in [1, 2].$

7)  $f(x) = 10x \ln x - \frac{x^2}{2} \rightarrow \min, \quad x \in [0, 1, 1].$

8)  $f(x) = e^x - \frac{1}{3}x^3 + 2x \rightarrow \min, \quad x \in [-2, 5, -1].$

9)  $f(x) = x^2 - 2x - 2 \cos x \rightarrow \min, \quad x \in [-0, 5, 1].$

3. Для выбранной функции и для каждого реализованного метода изучить зависимость скорости работы (числа вычислений

функции  $N$ ) от заданного значения точности  $\varepsilon$ . Провести сравнение методов друг с другом. *Объяснить полученные результаты.*

4. Вычислить аналитическое значение координаты минимума выбранной функции с точностью до 4 значащих цифр.

5. Определить, сколько вычислений функции потребуется каждому методу для того, чтобы разность между численным и аналитическим решениями была меньше  $\varepsilon = 10^{-4}$ .

6. Найти минимум функции  $f(x) = e^x - 1 - x - \frac{x^2}{2} - \frac{x^3}{6}$  с точностью

$\varepsilon = 10^{-4}$  на отрезке  $[-5, 5]$  методами золотого сечения и парабол. *Объяснить полученные результаты.*

7. Результаты работы сохранить для использования в задании для численной реализации гл. 3.

## Глава 3

# ОДНОМЕРНАЯ МИНИМИЗАЦИЯ. МЕТОДЫ, ИСПОЛЬЗУЮЩИЕ ИНФОРМАЦИЮ О ПРОИЗВОДНЫХ ЦЕЛЕВОЙ ФУНКЦИИ

Рассмотренные в предыдущей главе прямые методы используются при минимальных требованиях к целевой функции  $f(x)$ , которая считается унимодальной, и вычислению подлежат значения только самой функции, но не ее производных. Пусть теперь  $f(x)$  является дифференцируемой или дважды дифференцируемой выпуклой функцией и возможно вычисление производных  $f'(x)$  в произвольно выбранных точках. В этом случае эффективность поиска точки минимума можно существенно повысить.

Рассмотрим три метода минимизации, в которых используются значения производных целевой функции:

- метод средней точки;
- метод хорд;
- метод Ньютона.

Напомним, что для выпуклой дифференцируемой функции равенство  $f'(x) = 0$  является не только необходимым, но и достаточным условием глобального минимума. Поэтому, если известно, что  $x^*$  является внутренней точкой отрезка  $[a, b]$ , то приближенное равенство  $f'(x) \approx 0$  или  $|f'(x)| \leq \varepsilon$  может служить условием остановки вычислений в рассматриваемых трех методах.

### 3.1. Метод средней точки

Если определение производной  $f'(x)$  не представляет затруднений, то в процедуре исключения отрезков методом дихотомии вычисление двух значений  $f(x)$  вблизи середи-

ны очередного отрезка можно заменить вычислением одного значения  $f'(\bar{x})$  в его средней точке  $\bar{x} = \frac{a+b}{2}$ .

В самом деле, если  $f'(\bar{x}) > 0$ , то точка  $\bar{x}$  лежит на участке монотонного возрастания  $f(x)$ , поэтому  $x^* < \bar{x}$ , и точку минимума следует искать на отрезке  $[a, \bar{x}]$ . При  $f'(\bar{x}) < 0$  имеем противоположную ситуацию и переходим к отрезку  $[\bar{x}, b]$ . Равенство  $f'(\bar{x}) = 0$  означает, что точка минимума найдена точно  $x^* = \bar{x}$ , (рис. 3.1).

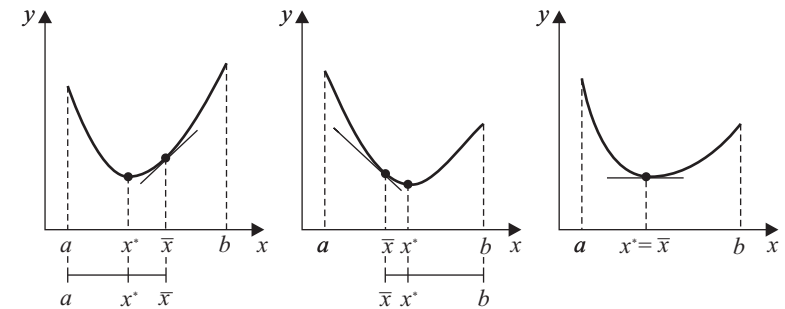


Рис. 3.1. Иллюстрация исключения отрезков методом средней точки

Такое исключение отрезков требует на каждой итерации только одного вычисления  $f'(x)$  и уменьшает отрезок поиска точки  $x^*$  ровно вдвое.

Алгоритм метода средней точки следующий.

Шаг 1. Положить  $\bar{x} = \frac{a+b}{2}$ . Вычислить  $f'(\bar{x})$ . Перейти

к шагу 2.

Шаг 2. Проверка на окончание поиска: если  $|f'(\bar{x})| \leq \varepsilon$ , то положить  $x^* = \bar{x}$ ,  $f^* = f(\bar{x})$  и завершить поиск, иначе перейти к шагу 3.

Шаг 3. Сравнить  $f'(\bar{x})$  с нулем: если  $f'(\bar{x}) > 0$ , то продолжить поиск на отрезке  $[a, \bar{x}]$ , положив  $b = \bar{x}$ , иначе — на отрезке  $[\bar{x}, b]$ , положив  $a = \bar{x}$ . Перейти к шагу 1.



**Пример 3.1.** Методом средней точки решить задачу  $f(x) = x^4 + e^{-x} \rightarrow \min, x \in [0, 1]$ , с точностью до  $|f'(x)| \leq 0,02$ .

Рассмотрим детально действия при выполнении итерации 1.

Шаг 1. Находим  $\bar{x} = 0,5$ ,  $f'(\bar{x}) = -0,107$ , переходим к шагу 2.

Шаг 2.  $|f'(\bar{x})| > 0,02$ , поэтому переходим к шагу 3.

Шаг 3.  $f'(\bar{x}) < 0$ , значит полагаем  $a = \bar{x} = 0,5$ , и переходим к следующей итерации, начиная с шага 1 и т.д. пока не достигнем необходимой точности.

Исходные данные и результаты вычислений при выполнении итераций приведены в табл. 3.1.

Таблица 3.1.

**Результаты вычислений примера методом средней точки**

| Номер итерации | $a$ | $b$   | $x$   | $f'(\bar{x})$ | Знак $f'(\bar{x})$  |
|----------------|-----|-------|-------|---------------|---------------------|
| 1              | 0   | 1     | 0,5   | -0,107        | -                   |
| 2              | 0,5 | 1     | 0,75  | 1,215         | +                   |
| 3              | 0,5 | 0,750 | 0,625 | 0,441         | +                   |
| 4              | 0,5 | 0,625 | 0,563 | 0,142         | +                   |
| 5              | 0,5 | 0,563 | 0,531 | 0,012         | точность достигнута |

Таким образом,  $x^* \approx 0,531$ ,  $f^* \approx 0,668$ .

Метод средней точки напоминает метод дихотомии, но сходится к искомому значению  $x^*$  быстрее, поскольку, в отличие от него, после вычисления  $n$  значений производной, для длины интервала неопределенности минимизируемой

на  $[a, b]$  функции  $f(x)$  получаем  $\Delta_n = \frac{b-a}{2^n}$ .

Таким образом, для одинакового уменьшения значения  $\Delta_n$  при использовании метода средней точки нужно вычислить вдвое меньше значений производной функции по сравнению с числом значений самой функции при методе дихотомии.

## 3.2. Метод хорд

Как уже отмечалось, равенство  $f'(x) = 0$  является необходимым и достаточным условием глобального минимума выпуклой дифференцируемой функции  $f(x)$ . Если на концах отрезка  $[a, b]$  производная  $f'(x)$  имеет разные знаки, т.е.  $f'(a)f'(b) < 0$ , и она непрерывна, то на интервале  $(a, b)$  найдется точка, в которой  $f'(x)$  обращается в нуль. В этом случае поиск точки минимума  $f(x)$  на отрезке  $[a, b]$  эквивалентен решению уравнения

$$f'(x) = 0, \quad x \in (a, b). \quad (3.1)$$

Таким образом, при  $f'(a)f'(b) < 0$  любой приближенный метод решения уравнения (3.1) можно рассматривать как метод минимизации выпуклой непрерывно дифференцируемой функции  $f(x)$  на отрезке  $[a, b]$ .

Сущность приближенного решения уравнения  $F(x) = 0$  на отрезке  $[a, b]$  при  $F(a)F(b) < 0$  методом хорд состоит в исключении отрезков путем определения  $\tilde{x}$  — точки пересечения с осью  $Ox$  хорды графика функции  $F(x)$  на  $[a, b]$ , представленного на рис. 3.2.

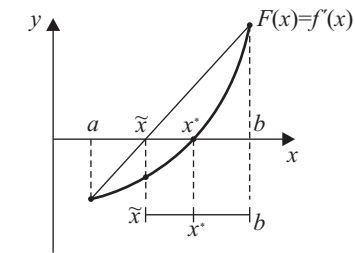


Рис. 3.2. Иллюстрация исключения отрезков методом хорд

Полагая  $F(x) = f'(x)$ , запишем координату точки  $\tilde{x}$

$$\tilde{x} = a - \frac{f'(a)}{f'(a) - f'(b)}(a - b). \quad (3.2)$$

Отрезок дальнейшего поиска точки  $x^*$  ( $[a, \tilde{x}]$  или  $[\tilde{x}, b]$ ) выбирается в зависимости от знака  $f'(\tilde{x})$  так же, как в методе средней точки. На каждой итерации, кроме первой, необходимо вычислять только одно новое значение  $f'(x)$ .

Алгоритма метода хорд следующий.

Шаг 1. Найти  $\tilde{x}$  по формуле (3.2). Вычислить  $f'(\tilde{x})$  и перейти к шагу 2.

Шаг 2. Проверка на окончание поиска: если  $|f'(\tilde{x})| \leq \varepsilon$ , то положить  $x^* = \tilde{x}$ ,  $f^* = f(\tilde{x})$  и завершить поиск, иначе перейти к шагу 3.

Шаг 3. Переход к новому отрезку. Если  $f'(\tilde{x}) > 0$ , то положить  $b = \tilde{x}$ ,  $f'(b) = f'(\tilde{x})$ , иначе положить  $a = \tilde{x}$ ,  $f'(a) = f'(\tilde{x})$ . Перейти к шагу 1.

**Пример 3.2.** Методом хорд решить задачу  $f(x) = x^4 + e^{-x} \rightarrow \min, x \in [0, 1]$ , с точностью до  $|f'(x)| \leq 0,05$ .

Условие  $f'(0)f'(1) < 0$  выполняется.

Рассмотрим детально действия при выполнении итерации 1.

Шаг 1. Находим:  $\tilde{x} = 0,216$ ,  $f'(\tilde{x}) = -0,766$ . Переходим к шагу 2.

Шаг 2.  $|f'(\tilde{x})| > 0,05$ , поэтому переходим к шагу 3.

Шаг 3.  $f'(\tilde{x}) < 0$ , поэтому полагаем  $a = 0,216$ ,  $f'(a) = -0,766$ . Переходим к итерации 2, начиная с шага 1 и т.д. пока не достигнем требуемой точности.

Результаты вычислений при выполнении итераций записаны в табл. 3.2.

Таблица 3.2.

Результаты вычислений примера методом хорд

| Номер итерации | $a$   | $b$ | $\tilde{x}$ | $f'(\tilde{x})$ | Знак $f'(\tilde{x})$ |
|----------------|-------|-----|-------------|-----------------|----------------------|
| 1              | 0     | 1   | 0,216       | -0,766          | —                    |
| 2              | 0,216 | 1   | 0,352       | -0,528          | —                    |
| 3              | 0,352 | 1   | 0,435       | -0,319          | —                    |
| 4              | 0,435 | 1   | 0,480       | -0,175          | —                    |
| 5              | 0,480 | 1   | 0,504       | -0,091          | —                    |
| 6              | 0,504 | 1   | 0,516       | -0,046          | точность достигнута  |

Отсюда  $x^* \approx 0,516$ ,  $f^* \approx 0,668$ .

До сих пор предполагалось, что  $f'(a)f'(b) < 0$ , т.е. производная  $f'(x)$  на концах отрезка имеет разные знаки. При нарушении этого условия точку  $x^*$  можно указать сразу.

Так, если  $f'(a), f'(b) > 0$ , то  $f(x)$  возрастает на  $[a, b]$ , следовательно,  $x^* = a$ , а при  $f'(a), f'(b) < 0$  она убывает и  $x^* = b$ .

В случае  $f'(a)f'(b) = 0$   $x^* = a$  или  $x^* = b$ , в зависимости от того, на каком из концов отрезка  $[a, b]$  производная  $f'(x) = 0$ .

### 3.3. Метод Ньютона

Если выпуклая на отрезке  $[a, b]$  функция  $f(x)$  дважды непрерывно дифференцируема на этом отрезке, то точку  $x^* \in [a, b]$  минимума этой функции можно найти, решая уравнение  $f'(x) = 0$  методом Ньютона (другое название — метод касательных). Пусть  $x_0 \in [a, b]$  — нулевое (начальное) приближение к искомой точке  $x^*$ . Линеаризуем функцию  $F(x) = f'(x)$  в окрестности начальной точки, приближенно заменив дугу графика этой функции касательной в точке  $(x_0, f'(x_0))$

$$F(x) \approx F(x_0) + F'(x_0)(x - x_0). \quad (3.3)$$

Выберем в качестве следующего приближения к  $x^*$  точку  $x_1$  пересечения касательной с осью абсцисс. При-

равнивая к нулю правую часть в (3.3), получим первый элемент  $x_1 = x_0 - \frac{F(x_0)}{F'(x_0)}$  итерационной последовательности

$$\{x_k\}, \quad k = 1, 2, \dots$$

В очередной точке  $x_k$  построим линейную аппроксимирующую функцию для  $F(x)$  и определим точку, в которой эта функция обращается в нуль, используя в качестве следующего приближения  $x_{k+1}$  (рис. 3.3).



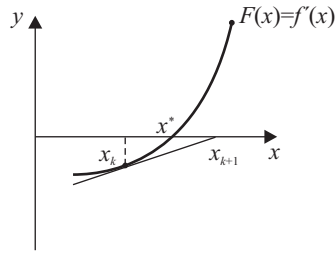


Рис. 3.3. Иллюстрация метода касательных

Уравнение касательной к графику  $F(x)$  в точке  $x = x_k$  имеет вид  $y = F(x_k) + F'(x_k)(x - x_k)$ , поэтому точка  $x = x_{k+1}$ , найденная из условия  $y = 0$ , определяется формулой

$x_{k+1} = x_k - \frac{F(x_k)}{F'(x_k)}$ . Поскольку  $F(x) \equiv f'(x)$ , получим, что для

решения уравнения  $f'(x) = 0$  необходимо построить последовательность

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}, \quad k = 1, 2, \dots \quad (3.4)$$

где  $x_0$  — точка, выбранная в качестве начального приближения. Вычисления по формуле (3.4) производятся до тех пор, пока не выполнится неравенство  $|f'(x_k)| \leq \varepsilon$ , после чего полагают  $x^* \approx x_k$ ,  $f^* \approx f(x_k)$ .

Формулу (3.4) можно получить также из иных соображений. Запишем квадратичную аппроксимацию  $q_k(x)$  функции  $f(x)$  в точке  $x_k$  с помощью формулы Тейлора

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

Функция  $q_k(x)$  — квадратный трехчлен, аппроксимирующий  $f(x)$  в окрестности точки  $x_k$ . Находя точку минимума  $q(x)$  из условия  $q'(x) = 0$  и считая ее следующим приближением  $x_{k+1}$  к  $x^*$ , получаем искомую формулу, (рис. 3.4).

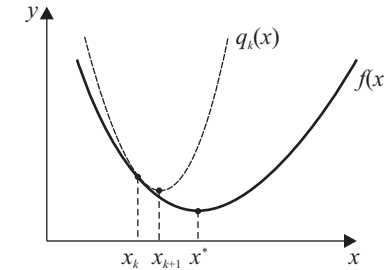


Рис. 3.4. Иллюстрация вывода формулы (3.4)

Для квадратичной функции  $f(x)$  функция  $f'(x)$  линейна, поэтому в (3.4) равенство будет точным, а метод Ньютона будет сходиться за один шаг при любом выборе точки  $x_0$  из области определения этой функции.

Для выяснения достаточных условий сходимости рассмотрим метод Ньютона в случае, если  $f(x)$  — трижды непрерывно дифференцируемая выпуклая функция. Напишем формулу Тейлора для  $f(x^*)$ , где  $x^*$  — искомый корень в окрестности  $k$ -го приближения

$$f'(x^*) = 0 = f'(x_k) + f''(x_k) \cdot (x^* - x_k) + \frac{f'''(\xi)}{2}(x^* - x_k)^2,$$

а точка  $\xi \in [x^*, x_k]$ .

Разделив последнее соотношение на  $f''(x_k)$  и перенеся первые два слагаемых из правой части в левую, получим

$$\left[ x_k - \frac{f'(x_k)}{f''(x_k)} \right] - x^* = \frac{1}{2} \frac{f'''(\xi)}{f''(x_k)} (x_k - x^*)^2,$$

что, учитывая (3.4), переписываем в виде

$$x_{k+1} - x^* = \frac{1}{2} \frac{f'''(\xi)}{f''(x_k)} (x_k - x^*)^2,$$

откуда

$$|x_{k+1} - x^*| = \frac{1}{2} \frac{|f'''(\xi)|}{|f''(x_k)|} |x_k - x^*|^2. \quad (3.5)$$

Из (3.5) следует оценка

$$|x_{k+1} - x^*| = \frac{1}{2} \frac{M_3}{m_2} |x_k - x^*|^2, \quad (3.6)$$

где  $M_3 = \max_{[a,b]} |f'''(x)|$ ,  $m_2 = \min_{[a,b]} |f''(x)|$ .

Ошибка убывает на каждом шаге в том случае, если

$$\frac{1}{2} \frac{M_3}{m_2} |x_0 - x^*| < 1. \quad (3.7)$$

Полученное условие означает, что сходимость зависит от начального приближения. Таким образом, метод Ньютона имеет *локальную сходимость*. Если начальное приближение взято далеко, на сходимость рассчитывать не приходится. С другой стороны, всегда можно добиться выполнения условия (3.7) за счет более точного выбора начального приближения  $x_0$ , например, с помощью нескольких итераций методами золотого сечения или поразрядного поиска.

Оценка (3.6) характеризует скорость убывания погрешности для метода Ньютона: на каждом шаге погрешность пропорциональна квадрату предыдущей (квадратичная скорость сходимости). Это очень высокий темп, например, если в некотором приближении получена одна точная цифра после запятой, то в следующем можно ожидать две точные цифры, затем — четыре и т.д.

Сформулируем теперь *достаточное условие монотонной сходимости* метода Ньютона. Пусть  $x^* \in [a, b]$  и  $f(x)$  трижды непрерывно дифференцируемая и выпуклая на отрезке  $[a, b]$  функция. Ясно, что итерационная последовательность  $\{x_k\}$  будет сходиться к пределу  $x^*$  монотонно, если

$$0 < \frac{x^* - x_{k+1}}{x^* - x_k} < 1.$$

В соответствии с формулой Тейлора с остаточным членом в форме Лагранжа

$$f'(x^*) = 0 = f'(x_k) + f''(x_k)(x^* - x_k) + \frac{f'''(x)}{2}(x^* - x_k)^2,$$

где точка  $x \in [x^*, x_k]$ . Поэтому с учетом основной формулы (3.4) имеем

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

Отсюда следует, что последовательность  $\{x_k\}$  монотонная, если  $\frac{f'''(x)}{f'(x_k)} > 0$ , т.е. достаточным условием монотонной сходимости метода Ньютона является постоянство в диапазоне  $x \in [x^*, x_0]$  знака производной  $f'''(x)$  и совпадение его со знаком  $f'(x_0)$ . При этом квадратичная скорость сходимости не гарантируется. Если кроме того, выполняется условие (3.6), то скорость сходимости метода Ньютона становится квадратичной.

Высокую скорость сходимости метода Ньютона можно объяснить и так: квадратный трехчлен  $q_k(x)$ , построенный с учетом информации как о первой, так и о второй производных  $f(x)$  в точке  $x_k$ , с высокой точностью аппроксимирует выпуклую дважды дифференцируемую функцию  $f(x)$  в достаточно малой окрестности этой точки. Поэтому, если очередное приближение  $x_k$  оказывается достаточно близким к  $x^*$ , то точки минимума  $x_{k+1}$  и  $x^*$  функций  $q_k(x)$  и  $f(x)$  практически совпадают.

**Пример 3.3.** Методом Ньютона найти точку минимума функции

$$f(x) = x \arctg x - \frac{1}{2} \ln(1+x^2) \text{ с точностью } |f'(x)| \leq 10^{-7}.$$

Производные функции равны  $f'(x) = \arctg(x)$ ,  $f''(x) = \frac{1}{1+x^2} > 0$ ,

$$f'''(x) = -\frac{2x}{(1+x^2)^2}. \text{ Видно, что при всех значениях переменной}$$

$f'(x)f'''(x) < 0$ , т.е. достаточные условия монотонной сходимости не выполняются. Выберем начальное приближение  $x_0 = 1$  и построим приближения  $x_k$  по формуле (3.4), результаты записаны в табл. 3.3.

Таблица 3.3.

## Результаты минимизации функции методом Ньютона

| $k$ | $x_k$            | $f'(x_k)$        |
|-----|------------------|------------------|
| 0   | 1                | 0,785            |
| 1   | -0,570           | -0,519           |
| 2   | 0,117            | 0,116            |
| 3   | -1,061 $10^{-3}$ | -1,061 $10^{-3}$ |
| 4   | 9 $10^{-8}$      | 9 $10^{-8}$      |

Откуда  $x^* \approx 9 \cdot 10^{-8} \approx 0$ .

Решение примера демонстрирует высокую скорость сходимости последовательности  $\{x_k\}$ . Можно провести несложные расчеты и убедиться, что оценка (3.6) и условие (3.7) выполняются, ошибка на каждом шаге убывает и имеет место квадратичная скорость сходимости.

Достаточные условия монотонной сходимости не выполняются, и последовательные приближения лежат в данном случае по разные стороны от корня.

Если начальное приближение выбрать не достаточно близким к точке  $x^*$ , то последовательность (3.4) метода Ньютона будет расходиться. Так, положив при решении примера  $x_0 = 3$ , получим последовательность  $x_1 = -9,5$ ,  $x_2 = 124$ ,  $x_3 = -23905$ ,  $x_4 = 8,97 \cdot 10^8$ ,  $x_5 = -1,27 \cdot 10^{18}$ , ... которая, очевидно, расходится.

В подобных случаях необходимо найти лучшее начальное приближение  $x_0$ , например, с помощью нескольких итераций методами золотого сечения или поразрядного поиска.

**Пример 3.4.** Найти  $f(x) = x^2 + \frac{16}{x} \rightarrow \min$ ,  $x \in [1, 4]$ .

Вычислим  $f'(x) = 2x - 16/x^2$ ,  $f''(x) = 2 + 32/x^3$ ,  $f'''(x) = -96/x^4$ . Всюду при  $x \in [1, 4]$  имеем:  $f''(x) > 0$ ,  $f'''(x) < 0$ . Поэтому, если начальную точку  $x_0$  выбрать исходя из условия  $f'(x_0) < 0$ , то построенная итерационная последовательность будет монотонной.

Выберем  $x_0 = 1$  и вычислим  $f'(x_0) = -14$ ,  $f''(x_0) = 34$ , а затем, по формуле (3.4) получим

$$x_1 = x_0 - \frac{f'(x_0)}{f''(x_0)} = 1 - \frac{-14}{34} \approx 1,4118.$$

Далее при помощи формулы (3.4) находим

$$x_2 \approx 1,8010, \quad x_3 \approx 1,9790, \quad x_4 \approx 1,9998.$$

Величина  $|x_4 - x^*| < |x_3 - x_4| \approx 0,0208$ . Действительно, точное значение минимума  $x^* = 2$ .

### 3.4. Возможные модификации метода Ньютона

#### Метод Ньютона-Рафсона

При переходе к очередной итерации новая точка  $x_{k+1}$  рассчитывается по формуле

$$x_{k+1} = x_k - \tau_k \frac{f'(x_k)}{f''(x_k)}, \quad 0 < \tau_k \leq 1.$$

В простейшем варианте метода  $\tau_k = \tau = \text{const}$  ( $\tau = 1$  соответствует исходному методу Ньютона). Оптимальный набор параметров  $\tau_k$  может быть найден из решения задачи минимизации

$$\phi(\tau) = f\left(x_k - \tau \frac{f'(x_k)}{f''(x_k)}\right) \rightarrow \min.$$

На практике для параметров  $\tau_k$  обычно используется приближенное решение последней задачи

$$\tau_k = \frac{(f'(x_k))^2}{(f'(x_k))^2 + (f'(\tilde{x}))^2}, \quad \text{где } \tilde{x} = x_k - \frac{f'(x_k)}{f''(x_k)}.$$

#### Метод Марквардта

При переходе к очередной итерации новая точка  $x_{k+1}$  рассчитывается по формуле

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k) + \mu_k}, \quad \mu_k > 0.$$

Значение параметра  $\mu_0$  выбирается как минимум на порядок больше значения  $f''(x_0)$ . При переходе к оче-

редной итерации новое значение  $\mu_{k+1}$  полагают равным  $\mu_{k+1} = \mu_k / 2$ , если  $f(x_{k+1}) < f(x_k)$ , либо  $\mu_{k+1} = 2\mu_k$  в противном случае.

### 3.5. Методы минимизации многомодальных функций

При использовании описанных до сих пор методов численной минимизации существенную роль играла унимодальность целевой функции  $f(x)$ . Если  $f(x)$  этим свойством не обладает, т.е. является многомодальной, то поиск минимума может оказаться безрезультатным или привести к ошибке (будет найдена точка локального экстремума, отличного от глобального минимума  $f(x)$ ). Поэтому представляют интерес методы минимизации многомодальных функций.

#### Метод перебора

Применение этого метода строго обосновано лишь для унимодальной на  $[a, b]$  функции  $f(x)$ . Однако, если вместо унимодальности потребовать, чтобы  $f(x)$  удовлетворяла на  $[a, b]$  условию Липшица

$$|f(x_1) - f(x_2)| \leq L|x_1 - x_2|, \text{ для всех } x_1, x_2 \in [a, b], \quad (3.8)$$

то можно гарантировать определение минимального значения  $f^*$  методом перебора с любой заданной точностью. Сформулируем утверждение более строго.

Пусть функция  $f(x)$  удовлетворяет на отрезке  $[a, b]$  условию Липшица (3.8) с константой  $L$  и приближенные значения  $x^* \approx x_m$ ,  $f^* \approx f(x_m)$  найдены методом перебора с разбиением отрезка  $[a, b]$  на  $n$  частей. Тогда для погрешности  $\delta_n$  определения минимального значения  $f^*$  справедлива оценка

$$\delta_n = f(x_m) - f^* \leq L \frac{b-a}{2n}. \quad (3.9)$$

Пусть  $x^*$  — точка минимума  $f(x)$  на  $[a, b]$  (существование точки минимума вытекает из теоремы Вейерштрасса

са для непрерывной функции). Среди пробных точек  $x_i = a + \frac{b-a}{n}i$ ,  $i = 0, \dots, n$  метода перебора найдется такая точка  $x_l$ , что  $|x_l - x^*| \leq \frac{b-a}{2n}$ . Поэтому, с учетом условия Липшица (3.8) можно записать

$$0 \leq f(x_m) - f^* \leq f(x_l) - f^* \leq L|x_l - x^*| \leq L \frac{b-a}{2n}.$$

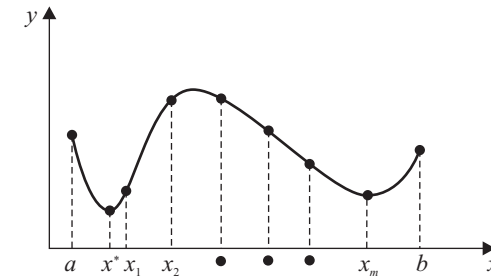


Рис. 3.5. Определение минимума многомодальной функции

Если функция  $f(x)$  многомодальна, то погрешность определения ее точки минимума может быть значительной, несмотря на то, что сам минимум  $f^*$  найден достаточно точно (рис. 3.5).

Однако во многих случаях практический интерес представляют те значения аргумента  $x$  (возможно, далекие от  $x^*$ ), при которых целевая функция принимает значения, достаточно близкие к минимальному. Это позволяет использовать метод перебора для многомодальных функций.

#### Метод ломаных

Этот метод также рассчитан на минимизацию многомодальных функций, удовлетворяющих условию Липшица. В нем используются кусочно-линейные аппроксимации

функции  $f(x)$ , графиками которых являются ломаные, что и объясняет название метода.

**Пример 3.5.** Пусть функция  $f(x)$  удовлетворяет на отрезке  $[a, b]$  условию Липшица с константой  $L$ . Зафиксируем точку  $\bar{x} \in [a, b]$  и введем вспомогательную функцию одной переменной  $g(\bar{x}, x) = f(\bar{x}) - L|\bar{x} - x|$ , график которой показан на рис. 3.6. Эта функция максимальна в точке  $\bar{x}$  и минимальна на концах отрезка  $[a, b]$ .

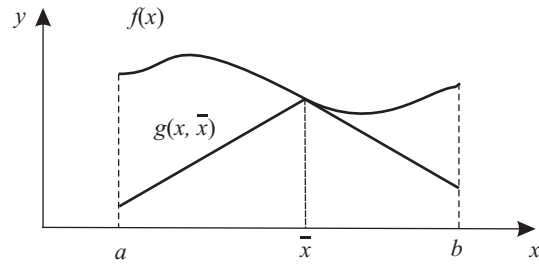


Рис. 3.6. График функции  $g(\bar{x}, x)$

Аппроксимирующие кусочно-линейные функции  $p_k(x)$ ,  $k=0, 1, \dots$  строятся следующим образом. Рассмотрим прямые  $y = f(a) - L(x - a)$  и  $y = f(b) + L(x - b)$ . Они пересекаются в точке  $(x_0, y_0)$  с координатами

$$\begin{aligned} x_0 &= \frac{1}{2L}[f(a) - f(b) + L(a + b)], \\ y_0 &= \frac{1}{2}[f(a) + f(b) + L(a - b)]. \end{aligned} \quad (3.10)$$

Положим

$$p_0(x) = \begin{cases} f(a) - L(x - a), & x \in [a, x_0], \\ f(b) + L(x - b), & x \in [x_0, b]. \end{cases}$$

График функции  $p_0(x)$  показан на рис. 3.7а, ее точка минимума  $x_0^* = x_0$ , а минимальное значение  $p_0^* = y_0$ .

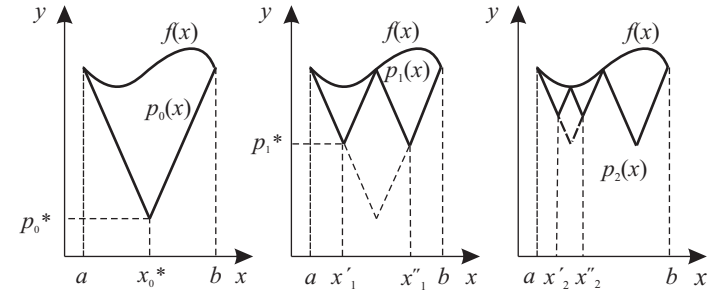


Рис. 3.7. Построение ломаных  $p_0(x)$ ,  $p_1(x)$ ,  $p_2(x)$

Используя вспомогательную функцию  $g(x_0^*, x)$ , определим аппроксимирующую функцию

$$p_1(x) = \max[p_0(x), g(x_0^*, x)],$$

у которой по сравнению с  $p_0(x)$  исчезла точка минимума  $x_0^*$ , но появились две новые точки локального минимума  $x'_1$  и  $x''_1$  (рис. 3.7б):

$$x'_1 = x_0^* - \Delta_1, \quad x''_1 = x_0^* + \Delta_1, \quad \text{где } \Delta_1 = \frac{1}{2L}[f(x_0^*) - p_0^*], \quad (3.11)$$

причем  $p_1(x'_1) = p_1(x''_1) = p_1 = \frac{1}{2}[f(x_0^*) + p_0^*]$ .

Формулы (3.11) получаются исходя из простых геометрических соображений и с учетом того, что тангенсы углов наклона каждого из звеньев ломаной  $p_1(x)$  к горизонтальной оси по модулю равны  $L$ .

Выберем точку глобального минимума  $p_1^*$  функции  $p_1(x)$ , обозначим ее через  $x_1^*$  (в данном случае это  $x'_1$  или  $x''_1$ , пусть, например,  $x_1^* = x'_1$ ) и положим

$$p_2(x) = \max[p_1(x), g(x_1^*, x)].$$

У функции  $p_2(x)$  по сравнению с  $p_1(x)$  вместо  $x_1^*$  появились две новые точки локального минимума  $x_2'$  и  $x_2''$  (рис. 3.7в), которые находятся аналогично (3.11)

$$x_2' = x_1^* - \Delta_2, \quad x_2'' = x_1^* + \Delta_2, \quad \text{где } \Delta_2 = \frac{1}{2L} [f(x_1^*) - p_1^*], \quad (3.12)$$

причем  $p_2(x_2') = p_2(x_2'') = p_2 = \frac{1}{2} [f(x_1^*) + p_1^*]$ .

Пусть теперь функция  $p_{k-1}(x)$  построена. Выбрав точку глобального минимума  $p_{k-1}^*$  функции  $p_{k-1}(x)$  и обозначив ее через  $x_{k-1}^*$ , определим функцию

$$p_k(x) = \max [p_{k-1}(x), g(x_{k-1}^*, x)].$$

Новые точки локального минимума  $x_k'$  и  $x_k''$  функции  $p_k(x)$ , появившиеся взамен  $x_{k-1}^*$ , а также значения  $p_k(x)$

в этих точках находятся по формулам, аналогичным (3.11) и (3.12)

$$x_k' = x_{k-1}^* - \Delta_k, \quad x_k'' = x_{k-1}^* + \Delta_k, \quad \text{где } \Delta_k = \frac{1}{2L} [f(x_{k-1}^*) - p_{k-1}^*],$$

$$p_k(x_k') = p_k(x_k'') = p_k = \frac{1}{2} [f(x_{k-1}^*) + p_{k-1}^*].$$

*Свойства функций  $p_k(x)$ , вытекающие из их построения.*

1.  $p_k(x)$  — непрерывная кусочно-линейная функция, ее графиком является ломаная, состоящая из отрезков с угловыми коэффициентами  $\pm L$ . У функции  $p_k(x)$  существует ровно  $k+1$  локальных минимумов, точками которых являются абсциссы вершин ломаной, а сами минимумы равны их ординатам.

2. Для всех  $x \in [a, b]$  и  $k = 1, 2, \dots$  справедливы неравенства  $p_{k-1}(x) \leq p_k(x) \leq f(x)$ .

Из этих свойств и метода построения функций  $p_k(x)$  следует, что с ростом  $k$  ломаные  $p_k(x)$  приближаются снизу к графику функции  $f(x)$  вблизи точки ее глобального минимума. Идея метода ломаных состоит в том, чтобы искать глобальный минимум не функции  $f(x)$ , а ломаных  $p_k(x)$ , что, вообще говоря, значительно проще, так как точками минимума  $p_k(x)$  могут быть только абсциссы вершин этой

ломаной, определяемые в ходе построения последовательности  $\{p_k(x)\}$ .

**Утверждение.** Пусть функция  $f(x)$  удовлетворяет на отрезке  $[a, b]$  условию Липшица с константой  $L$ . Тогда последовательность  $\{x_k^*\}$  точек минимума функций  $p_k(x)$ , построенных по методу ломаных, обладает следующими свойствами

$$1. \lim_{k \rightarrow \infty} f(x_k^*) = f^*.$$

2. Если точка глобального минимума функции  $f(x)$  на  $[a, b]$  единственна, то  $\lim_{k \rightarrow \infty} x_k^* = x^*$ .

3. Справедлива оценка точности  $\delta_k$  приближения  $f(x_{k-1}^*)$  к минимуму  $f^*$

$$0 \leq \delta_k = f(x_{k-1}^*) - f^* \leq f(x_{k-1}^*) - p_{k-1}^* = 2L\Delta_k, \quad (3.13)$$

которая используется в качестве условия окончания поиска минимума.

Алгоритм метода ломаных, при условии, что минимум  $f^*$  должен быть найден с точностью  $\delta$ , следующий.

Шаг 1. Вычислить  $(x_0, y_0)$  по формулам (3.10) и определить пару  $(x, p)$ , полагая  $x = x_0, p = y_0$ . Перейти к шагу 2.

Шаг 2. Сравнивая значения  $p$  введенных в рассмотрение пар  $(x, p)$ , определить пару  $(x_*, p_*)$ , у которой это значение минимально. Вычислить  $f(x_*)$ . Перейти к шагу 3.

Шаг 3. Проверка на окончание поиска с помощью условия

$$(3.13). \text{ Вычислить } \Delta = \frac{1}{2L} [f(x_*) - p_*] \text{ и положить } \delta_k = 2L\Delta.$$

Если  $\delta_k \leq \delta$ , то поиск завершить, полагая  $x^* \approx x_*, f^* \approx f(x_*)$ , иначе перейти к шагу 4.

Шаг 4. определить новые пары  $(x', p)$  и  $(x'', p)$ , полагая

$$x' = x_* - \Delta, \quad x'' = x_* + \Delta, \quad p = \frac{1}{2} [f(x_*) + p_*] \text{ и ввести их в рас-}$$

смотрение вместо пары  $(x_*, p_*)$ . Перейти к шагу 2.

Необходимо обратить внимание на следующее.

1. На каждой итерации метода ломаных, кроме первой, вычисляется только одно значение  $f(x)$ .



2. На практике часто для оценки минимальной константы Липшица  $L$  вычисляют угловые коэффициенты некоторого числа хорд на графике  $f(x)$ . Максимальный из модулей этих коэффициентов является приближением к  $L$  снизу, причем тем лучшим, чем меньше эти хорды.

3. Использование слишком завышенной оценки для минимальной константы Липшица снижает скорость сходимости метода ломаных, а применение заниженной оценки приводит, вообще говоря, к неверному результату.

\* \* \*

Теоретическое исследование методов одномерной минимизации и численные эксперименты показывают, что для достаточно гладких выпуклых функций методы полиномиальной интерполяции и метод Ньютона являются существенно более эффективными, чем методы исключения отрезков, особенно когда необходима высокая точность определения точки минимума.

Однако при ухудшении свойств функции  $f(x)$  (например, если  $f(x)$  не выпукла, быстро изменяется, имеет разрывные производные), эти алгоритмы минимизации по эффективности и надежности могут уступать методу золотого сечения, который практически не чувствителен к свойствам функции, если она унимодальна, или методу поразрядного поиска.

Поэтому, если особенно важно добиться надежной и устойчивой работы алгоритма, то целесообразно использовать метод золотого сечения или поразрядного поиска, как одни из лучших прямых методов. Можно также применять эти методы совместно с одним из более точных методов, например, квадратичной аппроксимации, переходя к алгоритму золотого сечения, если заданное число итераций более точного метода не приводит к результату.

### Вопросы и задания для самоконтроля

1. Пусть  $f(x)$  — выпуклая дифференцируемая функция и  $|f'(\bar{x})| \leq \epsilon$ . Можно ли указать погрешности определения точки минимума  $x^*$  и минимального значения  $f^*$  по формулам  $x^* = \bar{x}$ ,  $f^* = f(\bar{x})$ ? Ответ пояснить рисунком.

2. Является ли условие  $f'(\bar{x}) = 0$  достаточным для того, чтобы число  $\bar{x}$  было точкой минимума унимодальной, но не выпуклой функции  $f(x)$ ? Ответ сопроводить примером.

3. Указать класс функций, для которых точное определение точки минимума гарантировано в результате всего одной итерации метода Ньютона.

4. Сформулировать достаточные условия сходимости метода Ньютона.

5. Сформулировать достаточные условия монотонной сходимости метода Ньютона. Всегда ли в этом случае скорость сходимости будет квадратичной?

6. Для каких выпуклых дважды дифференцируемых функций метод золотого сечения приводит к цели за меньшее количество итераций, чем метод Ньютона?

7. Минимизировать функцию  $f(x) = (x-1)^8 \rightarrow \min, x \in [0, 2]$  с помощью методов Ньютона и золотого сечения. Сравнить скорость сходимости методов.

8. Сформулировать оценку погрешности определения минимума  $f^*$  многомодальной функции методом перебора.

9. Увеличение используемого значения константы Липшица  $L$  при реализации метода ломаных приводит к замедлению сходимости метода. Объяснить этот факт с помощью геометрической иллюстрации.

10. Показать с помощью рисунка, что если в методе ломаных используется ошибочно заниженное значение константы Липшица  $L$ , то задача минимизации может быть решена неверно.

### Задание для численной реализации в среде программирования MATLAB

1. Написать в среде MATLAB функции, реализующие метод средней точки, метод хорд и метод Ньютона.

2. Выбрать для выполнения работы тестовую функцию, номер которой соответствует номеру Вашего компьютера. Например, для компьютера №3 это будет функция 3), для компьютера №13 — функция 4):  $13-9=4$ ; для компьютера №23 это будет функция 5):  $23-9 \times 2=5$ .

$$1) f(x) = x^3 - 3 \sin x \rightarrow \min, \quad x \in [0, 1].$$

$$2) f(x) = x^4 + x^2 + x + 1 \rightarrow \min, \quad x \in [-1, 0].$$

$$3) f(x) = e^x + \frac{1}{x} \rightarrow \min, \quad x \in [0, 5, 1, 5].$$

$$4) f(x) = x^2 - 2x + e^{-x} \rightarrow \min, \quad x \in [-1, 1, 5].$$

$$5) f(x) = x \sin x + 2 \cos x \rightarrow \min, \quad x \in [-6, -4].$$

$$6) f(x) = x + \frac{1}{x^2} \rightarrow \min, \quad x \in [1, 2].$$

$$7) f(x) = 10x \ln x - \frac{x^2}{2} \rightarrow \min, \quad x \in [0, 1].$$

$$8) f(x) = e^x - \frac{1}{3}x^3 + 2x \rightarrow \min, \quad x \in [-2, 5].$$

$$9) f(x) = x^2 - 2x - 2 \cos x \rightarrow \min, \quad x \in [-0, 5].$$

3. Для выбранной функции и для каждого рассмотренного выше метода изучить зависимость скорости работы (числа вычислений функции  $N$ ) от заданного значения точности  $\epsilon$ . Провести сравнение методов друг с другом. *Объяснить полученные результаты.*

4. Определить, сколько вычислений функции потребуется каждому методу для того, чтобы разность между численным решением и аналитическим решением, найденным в задании для численной реализации гл. 2, была меньше  $\epsilon = 10^{-4}$ .

5. Сравнить полученные результаты с результатами выполнения задания для численной реализации гл. 2. *Сформулировать достоинства и недостатки прямых методов и методов, использующих производную целевой функции.*

6. С помощью метода Ньютона решить задачу минимизации функции

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

Определить диапазон начальных приближений, для которых применим метод Ньютона. *Объяснить полученный результат.*

7. Решить задачу предыдущего пункта с помощью одной из описанных выше модификаций метода Ньютона (метода Марквардта или метода Ньютона-Рафсона). Определить диапазон начальных приближений, для которых применима выбранная модификация метода Ньютона. *Сравнить полученные результаты с результатами выполнения предыдущего пункта.*

8. Составить программу нахождения глобального минимума многомодальных функций методом перебора и методом ломаных. С ее помощью решить один из следующих наборов задач:

$$a) f(x) = \frac{\cos x}{x^2} \rightarrow \min, \quad x \in [1, 12];$$

$$f(x) = \frac{1}{10}x + 2 \sin 4x \rightarrow \min, \quad x \in [0, 4];$$

$$б) f(x) = \frac{\cos(10x)}{e^x} \rightarrow \min, \quad x \in [1, 5];$$

$$f(x) = 0.3 \cos 2x + 2 \sin(4x) \rightarrow \min, \quad x \in [0, 4].$$

Сравнить найденные значения точек минимума с точными значениями, найденными аналитически. Сделать выводы о сравнительных достоинствах и недостатках метода перебора и метода ломаных.



## Глава 4

# ЗАДАЧА МИНИМИЗАЦИИ ФУНКЦИИ МНОГИХ ПЕРЕМЕННЫХ. НЕОБХОДИМЫЕ И ДОСТАТОЧНЫЕ УСЛОВИЯ БЕЗУСЛОВНОГО ЭКСТРЕМУМА

### 4.1. Постановка задачи и определения

Постановка задачи поиска минимума функций.

Найти такой вектор  $x^* = (x_1, \dots, x_n)^T$ , принадлежащий области допустимых значений  $U \subset E_n$  ( $E_n$  — евклидово пространство), целевая функция от которого

$$f(x^*) = \min_{x \in U} f(x). \quad (4.1)$$

Необходимо обратить внимание на следующее.

1. Поиск максимума функции  $f(x)$  сводится к задаче определения минимума путем замены знака перед функцией на противоположный

$$f(x^*) = \max_{x \in U} f(x) = -\min_{x \in U} (-f(x)).$$

2. Задача определения минимума и максимума целевой функции  $f(x)$  называется задачей поиска *экстремума*

$$f(x^*) = \text{extr}_{x \in U} f(x).$$

3. Если множество допустимых решений  $U$  задается ограничениями (условиями), накладываемыми на вектор  $x$ , то решается задача поиска *условного экстремума*. Если  $U = E_n$ , то ограничения (условия) на вектор  $x$  отсутствуют, и решается задача поиска *безусловного экстремума*.

4. Решением задачи поиска экстремума является пара  $(x^*, f(x^*))$ , которая включает точку  $x^*$  и значение целевой функции в ней.

Напомним, что точка  $x^* \in U$  называется точкой глобального минимума, или просто минимума функции  $f(x)$  на множестве  $U$ , если функция достигает в этой точке своего наименьшего значения, т.е.  $f(x^*) \leq f(x)$  для всех  $x \in U$ ;

точка  $x^* \in U$  называется точкой *локального минимума* функции  $f(x)$  на множестве  $U$ , если существует такое  $\varepsilon > 0$ , что если  $x^* \in U$  и  $\|x - x^*\| < \varepsilon$ , то  $f(x^*) \leq f(x)$ . Здесь  $\|x\| = \sqrt{x_1^2 + \dots + x_n^2}$  — евклидова норма вектора  $x$ .

**Определение.** Поверхностью уровня функции  $f(x)$  называется множество точек, в которых функция принимает постоянное значение, т.е.  $f(x) = \text{const}$ . Если  $n = 2$ , то поверхность уровня изображается *линией уровня* на плоскости  $E_2$ .

**Пример 4.1.** Построить линии уровня функций:

а)  $f(x) = x_1^2 + x_2^2$ ; б)  $f(x) = \frac{x_1^2}{4} + x_2^2$ ; в)  $f(x) = x_2^2 - x_1^2$ ; г)  $f(x) = x_2^2$

а)  $f(x) = x_1^2 + x_2^2 = \text{const} = r^2$  — уравнение окружности с центром в точке  $(0, 0)^T$  и радиусом, равным  $r$ ;

б)  $f(x) = \frac{x_1^2}{4} + x_2^2 = \text{const}$  — уравнение эллипса. Если  $\text{const} = 1$ , то  $a = 2, b = 1$  — большая и малая полуоси эллипса;

в)  $f(x) = x_2^2 - x_1^2 = \text{const}$  — уравнение гиперболы;

г)  $f(x) = x_2^2 = \text{const}$  — уравнение двух параллельных оси  $Ox_1$  прямых.

**Определение.** Градиентом  $\nabla f(x)$  непрерывно дифференцируемой функции  $f(x)$  в точке  $x$  называется вектор-столбец, элементами которого являются частные производные первого порядка, вычисленные в данной точке

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \dots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix}.$$

Градиент функции направлен по нормали к поверхности уровня, т.е. перпендикулярно к касательной плоскости, проведенной в точке  $x$ , в сторону наибольшего возрастания функции в данной точке.

**Определение.** Матрицей Гессе  $H(x)$  дважды непрерывно дифференцируемой в точке  $x$  функции  $f(x)$  называется матрица частных производных второго порядка, вычисленных в данной точке

$$H(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \dots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1n} \\ h_{21} & h_{22} & \dots & h_{2n} \\ \dots & \dots & \dots & \dots \\ h_{n1} & h_{n2} & \dots & h_{nn} \end{pmatrix},$$

где  $h_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$ ,  $i, j = 1, \dots, n$ .

Необходимо обратить внимание на следующее.

1. Матрица Гессе является симметрической размера  $n \times n$ .
2. Вместе с градиентом можно определить вектор *анти-градиента*, равный по модулю вектору градиента, но противоположный по направлению. Он указывает в сторону наибольшего убывания функции в данной точке.

3. С помощью градиента и матрицы Гессе, используя разложение в ряд Тейлора, приращение функции  $f(x)$  может быть записано в форме

$$\Delta f(x) = f(x + \Delta x) - f(x) = \nabla f(x)^T \Delta x + \frac{1}{2} \Delta x^T H(x) \Delta x + o(\|\Delta x\|^2), \quad (4.2)$$

где  $o(\|\Delta x\|^2)$  — сумма всех членов разложения, имеющих порядок выше второго,  $\Delta x^T H(x) \Delta x$  — квадратичная форма.

**Пример 4.2.** Для функции  $f(x) = x_1^2 + x_2^2$  вычислить градиент в точках  $x^{(0)} = (0, 1)^T$ ,  $x^{(1)} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T$ ,  $x^{(2)} = (1, 0)^T$ ,  $x^{(3)} = (0, -1)^T$  и найти матрицу Гессе.

По определению градиента и матрицы Гессе имеем

$$\nabla f(x) = (2x_1, 2x_2)^T, \quad \nabla f(x^{(0)}) = (0, 2)^T, \quad \nabla f(x^{(1)}) = (\sqrt{2}, \sqrt{2})^T,$$

$$\nabla f(x^{(2)}) = (2, 0)^T, \quad \nabla f(x^{(3)}) = (0, -2)^T, \quad H(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Заметим, что матрица Гессе квадратичной функции не зависит от  $x$ .

**Пример 4.3.** Для функции  $f(x) = x_1^2 + x_2^4$  вычислить градиент и найти матрицу Гессе в точках  $x^{(0)} = (0, 0)^T$ ,  $x^{(1)} = (1, 1)^T$ .

Согласно определениям имеем

$$\nabla f(x) = (2x_1, 4x_2^3)^T, \quad H(x) = \begin{pmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{pmatrix};$$

$$\nabla f(x^{(0)}) = (0, 0)^T, \quad H(x^{(0)}) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix};$$

$$\nabla f(x^{(1)}) = (2, 4)^T, \quad H(x^{(1)}) = \begin{pmatrix} 2 & 0 \\ 0 & 12 \end{pmatrix}.$$

**Определение.** Квадратичная форма  $\Delta x^T H(x) \Delta x$  (и соответствующая ей матрица Гессе  $H(x)$ ) называется:

*положительно определенной* ( $H(x) > 0$ ), если для любого ненулевого  $\Delta x$  выполняется неравенство  $\Delta x^T H(x) \Delta x > 0$ ;

отрицательно определенной ( $H(x) < 0$ ), если для любого ненулевого  $\Delta x$  выполняется неравенство  $\Delta x^T H(x) \Delta x < 0$ ;

положительно полуопределенной ( $H(x) \geq 0$ ), если для любого  $\Delta x$  выполняется неравенство  $\Delta x^T H(x) \Delta x \geq 0$  и имеется отличный от нуля вектор  $\Delta x$ , для которого  $\Delta x^T H(x) \Delta x = 0$ ;

отрицательно полуопределенной ( $H(x) \leq 0$ ), если для любого  $\Delta x$  выполняется неравенство  $\Delta x^T H(x) \Delta x \leq 0$  и имеется отличный от нуля вектор  $\Delta x$ , для которого  $\Delta x^T H(x) \Delta x = 0$ ;

неопределенной, если существуют такие векторы  $\Delta x$ ,  $\tilde{\Delta x}$ , что выполняются неравенства  $\Delta x^T H(x) \Delta x > 0$ ,  $\tilde{\Delta x}^T H(x) \tilde{\Delta x} < 0$ ;

тождественно равной нулю, ( $H(x) \equiv 0$ ), если для любого  $\Delta x$  выполняется  $\Delta x^T H(x) \Delta x = 0$ .

**Пример 4.4.** Классифицировать квадратичную форму и матрицу Гессе  $H(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ .

Выпишем квадратичную форму

$$\Delta x^T H(x) \Delta x = (\Delta x_1, \Delta x_2) \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} = 2\Delta x_1^2 + 2\Delta x_2^2.$$

Очевидно,  $\Delta x^T H(x) \Delta x > 0$  для любого ненулевого  $\Delta x$ . Согласно определению, квадратичная форма (матрица Гессе) положительно определенная.

**Пример 4.5.** Классифицировать квадратичную форму и матрицу Гессе  $H(x) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ .

Выпишем квадратичную форму

$$\Delta x^T H(x) \Delta x = (\Delta x_1, \Delta x_2) \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} = 2\Delta x_1^2.$$

Очевидно,  $\Delta x^T H(x) \Delta x \geq 0$  для любого вектора  $\Delta x$  и  $\Delta x^T H(x) \Delta x = 0$  для  $\Delta x_1 = 0$  и любых  $\Delta x_2 \neq 0$ . Согласно определению, квадратичная форма (матрица Гессе) положительно полуопределенная.

**Пример 4.6.** Найти матрицу Гессе функции  $f(x) = x_1^2 - x_2^2$  и классифицировать ее.

Следуя определению матрицы Гессе, получаем  $H(x) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$ .

Выпишем соответствующую ей квадратичную форму

$$\Delta x^T H(x) \Delta x = (\Delta x_1, \Delta x_2) \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \end{pmatrix} = 2\Delta x_1^2 - 2\Delta x_2^2.$$

При  $\Delta x_1 = 0$  и любых  $\Delta x_2 \neq 0$  квадратичная форма отрицательна, а при  $\Delta x_1 \neq 0$  и  $\Delta x_2 = 0$  положительна. Квадратичная форма (матрица Гессе) неопределенная.

Решение задач минимизации в  $E_n$ , как правило, сопряжено со значительными трудностями, особенно для многоэкстремальных функций. Некоторые из этих трудностей устраняются, если ограничиться рассмотрением только *выпуклых* целевых функций.

## 4.2. Свойства выпуклых множеств и выпуклых функций

**Определение.** Пусть  $x, y \in E_n$ . Множество точек вида  $\{z\} \subset E_n$ ,  $z = \alpha x + (1 - \alpha)y$ ,  $\alpha \in [0, 1]$  называется *отрезком*, соединяющим точки  $x$  и  $y$ .

В пространстве  $E_n$ ,  $n \leq 3$  это соотношение определяет обычный отрезок, соединяющий точки  $x$  и  $y$ .

**Определение.** Множество  $U \subset E_n$  называется *выпуклым*, если вместе с любыми точками  $x$  и  $y \in U$  оно содержит и весь отрезок  $z = \alpha x + (1 - \alpha)y$ ,  $\alpha \in [0, 1]$ .

Необходимо отметить, что проверка условия выпуклости в большинстве случаев требует громоздких выкладок, поэтому на практике при исследовании выпуклости множества в пространствах  $E_2$  и  $E_3$  часто прибегают к решениям в графическом виде.

Образно говоря, выпуклыми являются множества, которые не содержат «вмятин», «дырок», и состоят из одного «куска». Примерами выпуклых множеств являются само пространство  $E_n$ , отрезок, прямая, шар. Множество  $x^2 + y^2 = 1$  в  $E_2$  не является выпуклым.

**Определение.** Функция  $f(x)$ , заданная на выпуклом множестве  $U \subset E_n$ , называется:

— *выпуклой*, если для любых точек  $x, y \in U$  и любого  $\alpha \in [0, 1]$  выполняется неравенство

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y). \quad (4.3)$$

— *строго выпуклой*, если для всех  $\alpha \in (0, 1)$  неравенство (4.3) выполняется как строгое

$$f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha)f(y). \quad (4.4)$$

— *сильно выпуклой*, с константой  $l > 0$ , если для любых точек  $x, y \in U$  и любого  $\alpha \in [0, 1]$  выполняется неравенство

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y) - \frac{l}{2}\alpha(1-\alpha)\|x-y\|^2. \quad (4.5)$$

Необходимо обратить внимание на следующее.

1. Функцию  $f(x)$  называют выпуклой, если ее график целиком лежит не выше отрезка, соединяющего две ее произвольные точки. Функцию называют строго выпуклой, если ее график целиком лежит ниже отрезка, соединяющего две ее произвольные, но не совпадающие точки.

2. Если функция сильно выпуклая, то она одновременно строго выпуклая и выпуклая. Если функция строго выпуклая, то она одновременно выпуклая.

3. Выпуклость функции можно определить по матрице Гессе:

- если  $H(x) \geq 0$  для любого  $x \in E_n$ , то функция выпуклая,
- если  $H(x) > 0$  для любого  $x \in E_n$ , то функция строго выпуклая,
- если  $H(x) \geq lE$  для любого  $x \in E_n$ , где  $E$  — единичная матрица, то функция сильно выпуклая.

**Пример 4.7.** Исследовать на выпуклость функцию  $f(x) = x^2$ , определенную на множестве  $U = \{x \in [-1, 1]\}$ .

Функция является строго выпуклой, так как ее график целиком лежит ниже отрезка, соединяющего две ее произвольные, но не совпадающие точки. Более того, функция является сильно выпуклой, так как выполняется условие  $H(x) = f''(x) = 2 \geq l$  при  $0 < l \leq 2$ .

**Пример 4.8.** Исследовать на выпуклость функцию  $f(x) = x$ , определенную на множестве  $U = \{x \in [0, 1]\}$ .

Функция является выпуклой, так как ее график целиком лежит не выше отрезка, соединяющего две ее произвольные точки, но не является строго выпуклой и тем более сильно выпуклой.

**Пример 4.9.** Исследовать на выпуклость функцию  $f(x) = x_1^2 + x_2^2$ , определенную на множестве  $E_2$ .

Матрица Гессе функции удовлетворяет условию

$$H(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \geq l \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ при } 0 < l \leq 2.$$

Следовательно, эта функция сильно выпукла. Одновременно она является строго выпуклой и выпуклой.

В дальнейшем часто используются следующие свойства выпуклых функций.

1. Если  $f(x)$  выпуклая функция на выпуклом множестве  $U$ , то всякая точка локального минимума является точкой ее глобального минимума на  $U$ .

2. Если выпуклая функция достигает своего минимума в двух различных точках, то она достигает минимума во всех точках отрезка, соединяющего эти точки.

3. Если  $f(x)$  строго выпуклая функция на выпуклом множестве  $U$ , то она может достигать своего глобального минимума на  $U$  не более чем в одной точке.

### 4.3. Необходимые и достаточные условия безусловного экстремума

**Необходимые условия экстремума первого порядка.** Пусть  $x^* \in E_n$  есть точка локального минимума (максимума) функции  $f(x)$  на множестве  $E_n$  и  $f(x)$  дифференцируема в точке  $x^*$ . Тогда градиент функции  $f(x)$  в точке  $x^*$  равен нулю

$$\nabla f(x^*) = 0, \quad (4.6)$$

или

$$\frac{\partial f(x^*)}{\partial x_i} = 0, \quad i = 1, \dots, n. \quad (4.7)$$

Точки  $x^*$ , удовлетворяющие условию (4.6) или (4.7), называются *стационарными*.

**Необходимые условия экстремума второго порядка.** Пусть  $x^* \in E_n$  есть точка локального минимума (максимума) функции  $f(x)$ , определенной на множестве  $E_n$  и функция  $f(x)$  дважды дифференцируема в этой точке. Тогда матрица Гессе  $H(x^*)$  функции  $f(x)$ , вычисленная в точке  $x^*$ , является

положительно полуопределенной (отрицательно полуопределенной), т.е.  $H(x^*) \geq 0$ ,  $(H(x^*) \leq 0)$ .

**Достаточные условия экстремума.** Пусть функция  $f(x)$  в точке  $x^* \in E_n$  дважды дифференцируема, ее градиент равен нулю, а матрица Гессе является положительно определенной (отрицательно определенной), т.е.

$$\nabla f(x^*) = 0 \text{ и } H(x^*) > 0, (H(x^*) < 0). \quad (4.8)$$

Тогда  $x^*$  — точка локального минимума (максимума) функции  $f(x)$  на  $E_n$ .

**Проверка выполнения условий экстремума.** Рассмотрим матрицу Гессе, составленную для стационарной точки  $x^*$

$$H(x^*) = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1n} \\ h_{21} & h_{22} & \dots & h_{2n} \\ \dots & \dots & \dots & \dots \\ h_{n1} & h_{n2} & \dots & h_{nn} \end{pmatrix}.$$

**Угловыми минорами**  $k$ -го порядка матрицы размера  $n \times n$ , где  $k \leq n$  называются определители, составленные из элементов исходной матрицы, стоящих в ней на пересечении  $k$  верхних строк и  $k$  левых столбцов.

**Главными минорами**  $m$ -го порядка матрицы размера  $n \times n$ , где  $m \leq n$  называются определители, составленные из элементов исходной матрицы, оставшихся после вычеркивания в ней любых  $(n-m)$  строк и  $(n-m)$  столбцов с одинаковыми номерами.

Существуют два основных способа проверки выполнения условий экстремума. *Первый способ* связан с исследованием положительной или отрицательной определенности угловых и главных миноров матрицы Гессе. *Второй способ* основан на анализе собственных значений этой матрицы.

В первом случае, при анализе знаков угловых и главных миноров матрицы Гессе, доказаны и обычно используются на практике два критерия — проверки достаточных и необходимых условий второго порядка. Приведем их без доказательства.

### Критерий Сильвестра проверки достаточных условий экстремума

1. Для того, чтобы матрица Гессе  $H(x^*)$  была положительно определенной ( $H(x) > 0$ ) и стационарная точка  $x^*$  являлась точкой локального минимума, необходимо и достаточно, чтобы знаки угловых миноров были строго положительны

$$(\Delta_1 > 0, \Delta_2 > 0, \dots, \Delta_n > 0).$$

2. Для того, чтобы матрица Гессе  $H(x^*)$  была отрицательно определенной ( $H(x) < 0$ ) и стационарная точка  $x^*$  являлась точкой локального максимума, необходимо и достаточно, чтобы знаки угловых миноров чередовались, начиная с отрицательного

$$(\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \dots, (-1)^n \Delta_n > 0).$$

### Критерий проверки необходимых условий экстремума второго порядка

1. Для того, чтобы матрица Гессе  $H(x^*)$  была положительно полуопределенной ( $H(x) \geq 0$ ) и стационарная точка  $x^*$  может быть являлась точкой локального минимума, необходимо и достаточно, чтобы все главные миноры определителя матрицы Гессе были неотрицательны

$$(\Delta_1 \geq 0, \Delta_2 \geq 0, \dots, \Delta_n \geq 0).$$

2. Для того, чтобы матрица Гессе  $H(x^*)$  была отрицательно полуопределенной ( $H(x) \leq 0$ ) и стационарная точка  $x^*$  может быть являлась точкой локального максимума, необходимо и достаточно, чтобы все главные миноры четного порядка были неотрицательны, а все главные миноры нечетного порядка — неположительны

$$(\Delta_1 \leq 0, \Delta_2 \geq 0, \Delta_3 \leq 0, \dots, (-1)^n \Delta_n \geq 0).$$

*Второй способ* проверки условий экстремума связан с анализом собственных значений матрицы Гессе и применим только в случае, если эти значения удастся вычислить.

**Определение.** Собственные значения  $\lambda_i$ ,  $i = 1, \dots, n$  матрицы  $H(x^*)$  размера  $n \times n$  находятся как корни характеристического уравнения (алгебраического уравнения  $n$ -й степени)  $|H(x^*) - \lambda E| = 0$ .

Необходимо отметить, что собственные значения вещественной симметрической матрицы  $H(x^*)$  вещественны.

Если собственные значения матрицы Гессе вычислены, то дальнейшая проверка условий экстремума не составляет затруднений. Приведем в виде таблицы порядок применения двух способов, позволяющих установить достаточные условия экстремума и необходимые условия второго порядка при соблюдении необходимых условий первого порядка  $\nabla f(x^*) = 0$  (табл. 4.1).



Таблица 4.1.

Критерии проверки достаточных и необходимых условий второго порядка при поиске безусловного экстремума.

| № | $H(x^*)$ | Условия  | Первый способ  | Второй способ                               | Тип стационарной точки $x^*$                                    |
|---|----------|--|--|---|---|
| 1 | $>0$     | Достаточные условия экстремума                 | $\Delta_1 > 0, \Delta_2 > 0, \dots, \Delta_n > 0$                          | $\lambda_1 > 0, \dots, \lambda_n > 0$       | Локальный минимум   |
| 2 | $<0$     | Достаточные условия экстремума                 | $\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \dots, (-1)^n \Delta_n > 0$     | $\lambda_1 < 0, \dots, \lambda_n < 0$       | Локальный максимум  |
| 3 | $\geq 0$ | Необходимые условия экстремума второго порядка | Все главные миноры определителя матрицы $H(x^*)$ неотрицательны            | $\lambda_1 \geq 0, \dots, \lambda_n \geq 0$ | Может быть локальный минимум, требуется дополнит. исследование  |
| 4 | $\leq 0$ | Необходимые условия экстремума второго порядка | Все главные миноры четного порядка $\geq 0$ , нечетного порядка $\leq 0$ . | $\lambda_1 \leq 0, \dots, \lambda_n \leq 0$ | Может быть локальный максимум, требуется дополнит. исследование |
| 5 | $=0$     | Необходимые условия экстремума второго порядка | Матрица Гессе состоит из нулевых элементов                                 | $\lambda_1 = 0, \dots, \lambda_n = 0$       | Требуется дополнит. исследование                                |
| 6 | $>0, <0$ | Необходимые условия экстремума второго порядка | Не выполняются условия 1-5.  | $\lambda_i$ имеют разные знаки              | Нет экстремума  |

Алгоритм решения задачи следующий.

Шаг 1. Записать необходимые условия экстремума первого порядка и найти стационарные точки  $x^*$  в результате решения системы  $n$  в общем случае нелинейных алгебраических уравнений с  $n$  неизвестными. Для численного решения системы могут использоваться методы простой итерации, Зейделя, Ньютона.

Шаг 2. В найденных стационарных точках  $x^*$  проверить выполнение достаточных условий, а если они не выполняются, то необходимыми условий второго порядка с помощью одного из двух способов (табл. 4.1).

Шаг 3. Вычислить значения  $f(x^*)$  в точках экстремума. Необходимо обратить внимание на следующее.

1. Если требуется определить глобальные экстремумы, то они находятся в результате сравнения значений функции в точках локальных экстремумов с учетом ограниченности функции на  $E_n$ .

2. Для функции одной переменной  $f(x)$  можно воспользоваться следующим правилом, заменяющим операции на шаге 2.

Если  $f(x)$  и ее производные непрерывны, то точка  $x^*$  является точкой экстремума тогда и только тогда, когда  $m$  — четное, где  $m$  — порядок первой не обращающейся в нуль в точке  $x^*$  производной. Если  $f^{(m)}(x^*) > 0$ , то в точке  $x^*$  — локальный минимум, а если  $f^{(m)}(x^*) < 0$ , то в точке  $x^*$  — локальный максимум. Если число  $m$  нечетное, то в точке  $x^*$  нет экстремума.

3. Часто на практике при применении численных методов поиска экстремума требуется проверить, выполняются ли необходимые и достаточные условия экстремума в некоторой точке. Такая проверка необходима, так как многие численные методы позволяют найти лишь стационарную точку, тип которой требует уточнения.

**Пример 4.10.** Найти экстремум функции  $f(x) = x_1^2 + x_2^2$ , определенной на множестве  $E_2$ .

1. Запишем необходимые условия экстремума первого порядка

$$\frac{\partial f(x)}{\partial x_1} = 2x_1 = 0; \quad \frac{\partial f(x)}{\partial x_2} = 2x_2 = 0.$$

В результате решения системы получаем стационарную точку  $x^* = (0, 0)^T$ .

2. Проверим выполнение достаточных условий экстремума.

*Первый способ.* Матрица Гессе имеет вид  $H(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ . Так как  $\Delta_1 = h_{11} = 2 > 0$ ,  $\Delta_2 = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4 > 0$ , то в точке  $x^*$  локальный минимум (строка 1 табл. 4.1).

*Второй способ.* Найдем собственные значения матрицы Гессе

$$\begin{vmatrix} 2-\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} = 0. \text{ Отсюда } (2-\lambda)^2 = 0, \lambda_1 = \lambda_2 = 2 > 0.$$

Так как все собственные значения положительны, то в точке  $x^*$  локальный минимум (строка 1 табл. 4.1). Так как  $H(x^*) > 0$ , то функция строго выпукла на множестве  $E_2$ . Поэтому точка локального минимума является и точкой глобального минимума.

3. Вычислим значение функции в точке глобального минимума  $f(x^*) = 0$ .

**Пример 4.11.** Найти экстремум функции  $f(x) = x_1^2 - x_2^2$  на множестве  $E_2$ .

1. Запишем необходимые условия экстремума первого порядка

$$\frac{\partial f(x)}{\partial x_1} = 2x_1 = 0; \quad \frac{\partial f(x)}{\partial x_2} = -2x_2 = 0.$$

В результате решения системы получаем стационарную точку  $x^* = (0, 0)^T$ .

2. Проверим выполнение достаточных условий экстремума и необходимых условий второго порядка.

*Первый способ.* Матрица Гессе имеет вид  $H(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$ . Так как  $\Delta_1 = h_{11} = 2 > 0$ ,  $\Delta_2 = \begin{vmatrix} 2 & 0 \\ 0 & -2 \end{vmatrix} = -4 < 0$ , то достаточные условия экстремума

не выполняются (строки 1 и 2 табл. 4.1). Проверим выполнение необходимых условий второго порядка. Главные миноры первого порядка ( $m=1$ ) получаются из  $\Delta_2$  в результате вычеркивания  $n-m=2-1=1$  строк и столбцов с одинаковыми номерами, а их значения равны соответственно  $-2$  и  $2$ . Главный минор второго порядка получается из  $\Delta_2$  в результате вычеркивания  $n-m=0$  строк и столбцов и совпадает с  $\Delta_2 = -4$ . Отсюда следует, что необходимые условия экстремума второго порядка не выполняются (строки 3 и 4 в табл. 4.1). Так как матрица Гессе не является нулевой, то делаем вывод, что в точке  $x^*$  экстремума нет (строка 6 в табл. 4.1).

*Второй способ.* Найдем собственные значения матрицы Гессе:

$$\begin{vmatrix} 2-\lambda & 0 \\ 0 & -2-\lambda \end{vmatrix} = 0. \text{ Отсюда } (2-\lambda)(-2-\lambda) = 0, \lambda_1 = -2 < 0, \lambda_2 = 2 > 0,$$

т.е. собственные значения имеют разные знаки. Поэтому точка  $x^*$  не является точкой минимума или максимума (строка 6 в табл. 4.1).

3. Так как экстремума нет ни в одной точке, значение  $f(x^*)$  не вычисляется.

**Пример 4.12.** Найти экстремум функции  $f(x) = x_1^2 + x_2^4$ , определенной на множестве  $E_2$ .

1. Запишем необходимые условия экстремума первого порядка

$$\frac{\partial f(x)}{\partial x_1} = 2x_1 = 0; \quad \frac{\partial f(x)}{\partial x_2} = 4x_2^3 = 0.$$

В результате решения системы получаем стационарную точку  $x^* = (0, 0)^T$ .

2. Проверим выполнение достаточных условий второго порядка.

Матрица Гессе имеет вид  $H(x^*) = \begin{pmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ . Так как

$$\Delta_1 = h_{11} = 2 > 0, \quad \Delta_2 = \begin{vmatrix} 2 & 0 \\ 0 & 0 \end{vmatrix} = 0,$$

то достаточные условия экстремума не выполняются. Проверим выполнение необходимых условий экстремума второго порядка. Аналогично решению предыдущего примера получаем значения главных миноров первого порядка, соответственно 2 и 0 и главный минор второго порядка 0. Так как все главные миноры неотрицательные, то в точке  $x^*$  может быть минимум и требуется дополнительное исследование.

3. Вычислим значение целевой функции в стационарной точке  $f(x^*) = 0$  и рассмотрим  $\epsilon$ -окрестность точки  $x^*$ , а также поведение функции  $f(x)$  на множестве  $E_2$ . При любых  $x \in E_2$  имеем  $f(x) \geq f(x^*) = 0$ . Поэтому точка  $x^*$  является точкой глобального минимума.

**Пример 4.13.** Найти экстремум функции

$$f(x) = -x_1^2 - x_2^2 - x_3^2 - x_1 + x_1x_2 + 2x_3 \text{ на } E_3.$$

1. Запишем необходимые условия экстремума первого порядка

$$\frac{\partial f(x)}{\partial x_1} = -2x_1 - 1 + x_2 = 0, \quad \frac{\partial f(x)}{\partial x_2} = -2x_2 + x_1 = 0, \quad \frac{\partial f(x)}{\partial x_3} = -2x_3 + 2 = 0.$$

В результате решения системы получаем стационарную точку

$$x^* = \left(-\frac{2}{3}, -\frac{1}{3}, 1\right)^T.$$

2. Проверим выполнение достаточных условий экстремума.

*Первый способ.* Матрица Гессе имеет вид  $H(x^*) = \begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$ .

Так как  $\Delta_1 = -2 < 0$ ,  $\Delta_2 = \begin{vmatrix} -2 & 1 \\ 1 & -2 \end{vmatrix} = 4 - 1 = 3 > 0$ ,  $\Delta_3 = (-2) \cdot 3 = -6 < 0$ ,

т.е. знаки угловых миноров чередуются, начиная с отрицательного, то  $x^*$  — точка локального максимума.

*Второй способ.* Найдем собственные значения матрицы Гессе

$$\det(H - \lambda E) = \begin{vmatrix} -2-\lambda & 1 & 0 \\ 1 & -2-\lambda & 0 \\ 0 & 0 & -2-\lambda \end{vmatrix} = 0.$$

Отсюда  $(-2-\lambda)[(-2-\lambda)^2 - 1] = 0$  и  $\lambda_1 = -2 < 0$ ,  $\lambda_2 = -1 < 0$ ,  $\lambda_3 = -3 < 0$ . Так как все собственные значения матрицы Гессе отрицательны, то в точке  $x^*$  локальный максимум.

3. Вычислим значение функции в точке локального максимума

$$f(x^*) = \frac{4}{3}.$$

**Пример 4.14.** Найти экстремум функции  $f(x) = x_1^3 + x_2^3 - 3x_1x_2$  на  $E_2$ .

1. Запишем необходимые условия экстремума первого порядка

$$\frac{\partial f(x)}{\partial x_1} = 3x_1^2 - 3x_2 = 0; \quad \frac{\partial f(x)}{\partial x_2} = 3x_2^2 - 3x_1 = 0.$$

Решая систему, найдем стационарные точки  $x_1^* = (0, 0)^T$ ;  $x_2^* = (1, 1)^T$ .

2. Проверим выполнение достаточных условий второго порядка.

*Первый способ.* Матрица Гессе имеет вид

$$H(x_1^*) = \begin{pmatrix} 6x_1 & -3 \\ -3 & 6x_2 \end{pmatrix} = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix} \text{ при } x = x_1^*.$$

Так как  $\Delta_1 = h_{11} = 0$ ,  $\Delta_2 = \begin{vmatrix} 0 & -3 \\ -3 & 0 \end{vmatrix} = -9$ , то достаточные условия экстремума не выполняются. Проверим выполнение необходимых условий экстремума второго порядка. Главные миноры первого

порядка равны 0, главный минор второго порядка равен  $-9$ . Необходимое условие второго порядка не выполняется, следовательно, в точке  $x_1^* = (0, 0)^T$  экстремума быть не может. В точке  $x_2^* = (1, 1)^T$  матрица Гессе имеет вид

$$H(x_2^*) = \begin{pmatrix} 6x_1 & -3 \\ -3 & 6x_2 \end{pmatrix} = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}.$$

Так как  $\Delta_1 = h_{11} = 6 > 0$ ,  $\Delta_2 = \begin{vmatrix} 6 & -3 \\ -3 & 6 \end{vmatrix} = 27 > 0$ , то выполняются достаточные условия, следовательно,  $x_2^* = (1, 1)^T$  — точка локального минимума.

*Второй способ.* Найдем собственные значения матрицы Гессе в точке

$$x_1^* = (0, 0)^T : \det(H - \lambda E) = \begin{vmatrix} -\lambda & -3 \\ -3 & -\lambda \end{vmatrix} = 0, \text{ откуда } \lambda_1 = -3, \lambda_2 = 3. \text{ Так как}$$

собственные значения имеют разные знаки, экстремума нет.

Найдем собственные значения матрицы Гессе в точке  $x_2^* = (1, 1)^T$ :

$$\det(H - \lambda E) = \begin{vmatrix} 6-\lambda & -3 \\ -3 & 6-\lambda \end{vmatrix} = 0, \text{ откуда } \lambda_1 = 3, \lambda_2 = 9. \text{ Так как все собс-}$$

твенные значения положительны, следовательно,  $x_2^* = (1, 1)^T$  — точка локального минимума.

3. Вычислим значение функции в точке локального максимума:  $f(x_2^*) = -1$ .

**Пример 4.15.** Найти экстремум функции

$$f(x) = 5x^6 - 36x^5 + \frac{165}{2}x^4 - 60x^3 + 36 \text{ на } E_1.$$

1. Выпишем необходимое условие экстремума первого порядка

$$\frac{df(x)}{dx} = 30x^5 - 180x^4 + 330x^3 - 180x^2 = 30x^2(x-1)(x-2)(x-3) = 0.$$

Отсюда получаем стационарные точки  $x_1^* = 0$ ,  $x_2^* = 1$ ,  $x_3^* = 2$ ,  $x_4^* = 3$ .

2. Проверим выполнение достаточных условий экстремума

$$\frac{d^2f(x)}{dx^2} = 150x^4 - 720x^3 + 990x^2 - 360x,$$

$$f''(x_1^*) = 0, f''(x_2^*) = 60 > 0, f''(x_3^*) = -120 < 0, f''(x_4^*) = 540 > 0,$$

поэтому в точках  $x_2^*$ ,  $x_4^*$  — локальный минимум, а в точке  $x_3^*$  — локальный максимум. В точке  $x_1^*$  достаточные условия не выполняются, поэтому вычислим третью производную

$$f'''(x) = 600x^3 - 2160x^2 + 1980x - 360, \quad f'''(x_1^*) = -360.$$



Так как эта производная отлична от нуля и имеет нечетный порядок, то в точке  $x_1^*$  экстремума нет.

3. Вычислим значения экстремумов функции

$$f(1) = \frac{55}{2}; f(2) = -108; f(3) = \frac{531}{2}.$$

### Вопросы и задания для самоконтроля

1. Что такое градиент и антиградиент функции многих переменных и каков их геометрический смысл?
2. Что такое матрица Гессе функции многих переменных?
3. Записать приращение функции  $f(x) \in C^2(E_n)$  в точке  $x$  через градиент и матрицу Гессе.
4. Вычислить и нарисовать градиенты, а также вычислить матрицу Гессе функции  $f(x) = x_1^2 - x_2^2$  в точках  $x_1 = (1, 1)^T$ ;  $x_2 = (1, -1)^T$ .
5. Классифицировать квадратичные формы и соответствующие им матрицы Гессе

$$H(x) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}; H(x) = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}; H(x) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}; H(x) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

6. Почему в этой главе и далее в книге рассматриваются, в основном, выпуклые функции?
7. Дать определение выпуклого множества. Сформулировать его геометрический смысл.
8. Дать определение выпуклой функции. Сформулировать геометрическую интерпретацию этого свойства, свойство матрицы Гессе выпуклой функции.
9. Дать определение строго выпуклой функции. Сформулировать геометрическую интерпретацию этого свойства, свойство матрицы Гессе строго выпуклой функции.
10. Дать определение сильно выпуклой функции. Сформулировать свойство матрицы Гессе сильно выпуклой функции.
11. Исследовать на выпуклость

$$f(x) = x^2, x \in [-1, 1]; f(x) = x, x \in [0, 1]; f(x) = x_1^2 + x_2^2, x \in E_2.$$

12. Какие свойства выпуклых функций Вы знаете?
13. Сформулировать необходимое условие первого порядка для безусловного экстремума функции многих переменных.

14. Сформулировать необходимое условие второго порядка для безусловного экстремума функции многих переменных.
15. Сформулировать достаточное условие безусловного экстремума функции многих переменных.
16. Что такое угловые миноры и главные миноры квадратной матрицы?
17. Сформулировать критерий Сильвестра проверки достаточных условий безусловного экстремума функции многих переменных.
18. Сформулировать критерий проверки необходимых условий экстремума второго порядка функции многих переменных.
19. Сформулировать второй способ проверки условий экстремума функции многих переменных (через собственные значения матрицы Гессе).
20. Сформулировать необходимые и достаточные условия безусловного экстремума функции  $f(x)$  одной переменной.

## Глава 5

# ОБЩИЕ ПРИНЦИПЫ МНОГОМЕРНОЙ МИНИМИЗАЦИИ. МЕТОДЫ ГРАДИЕНТНОГО СПУСКА. МЕТОД СОПРЯЖЕННЫХ НАПРАВЛЕНИЙ И МЕТОД НЬЮТОНА

Решение задач минимизации в многомерных пространствах сопряжено, как правило, со значительными трудностями, особенно для многоэкстремальных функций. Многие из этих трудностей устраняются, если ограничиться рассмотрением только выпуклых квадратичных функций.

### 5.1. Выпуклые квадратичные функции

Во многих задачах оптимизации рассматриваются квадратичные функции, которые в  $n$ -мерном случае являются обобщением квадратного трехчлена одной переменной

$$f(x) = \frac{1}{2}ax^2 + bx + c.$$

**Определение.** Функция вида

$$f(x) = \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} x_i x_j + \sum_{j=1}^n b_j x_j + c \quad (5.1)$$

называется *квадратичной функцией  $n$  переменных*.

Положив  $a_{ij} = \alpha_{ij} + \alpha_{ji}$ , получим симметрическую матрицу  $A = (a_{ij})$ , с помощью которой выражение (5.1) можно записать в другой форме

$$f(x) = \frac{1}{2}(Ax, x) + (b, x) + c, \quad (5.2)$$

где  $b = (b_1, \dots, b_n)^T \in E_n$  — вектор коэффициентов,  $x = (x_1, \dots, x_n)^T$ ;  $(x, y)$  — скалярное произведение векторов  $x, y \in E_n$ .

**Пример 5.1.** Функция

$$f(x) = 2x_1^2 - 2x_1x_2 + 3x_1x_3 + x_2^2 - 2x_2x_3 + 4x_3^2 + x_1 - x_2 + 3x_3 + 5$$

является квадратичной. Запишем ее матрицу  $A$ , вектор  $b$  и коэффициент  $c$  из (5.2)

$$A = \begin{pmatrix} 4 & -2 & 3 \\ -2 & 2 & -2 \\ 3 & -2 & 8 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}, \quad c = 5.$$

Перечислим наиболее важные свойства квадратичных функций.

1. Для градиента квадратичной функции (5.2) справедлива формула

$$\nabla f(x) = Ax + b.$$

Запишем  $k$ -ю координату вектора  $\nabla f(x)$

$$\frac{\partial f}{\partial x_k} = \frac{\partial}{\partial x_k} \left( \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + \sum_{j=1}^n b_j x_j \right) = \frac{1}{2} \sum_{i=1}^n (a_{ik} + a_{ki}) x_i + b_k = \sum_{i=1}^n a_{ki} x_i + b_k.$$

2. Гессиан квадратичной функции (5.2) совпадает с матрицей  $A$

$$H(x) = A.$$

Вычислим элемент матрицы Гессе

$$\frac{\partial^2 f}{\partial x_l \partial x_k} = \frac{\partial}{\partial x_l} \left( \frac{\partial f}{\partial x_k} \right) = \frac{\partial}{\partial x_l} \left( \sum_{i=1}^n a_{ki} x_i + b_k \right) = a_{kl}.$$

3. Квадратичная функция (5.2) с положительно определенной матрицей  $A$  сильно выпукла.

Так как матрица  $H(x) = A$  симметрична и положительно определена, то все ее собственные значения  $\lambda_i$  положительны и существует ортонормированный базис из собственных векторов этой матрицы.

$$A = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}, \quad A - IE = \begin{pmatrix} \lambda_1 - l & 0 & \dots & 0 \\ 0 & \lambda_2 - l & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n - l \end{pmatrix}.$$

В этом базисе все угловые миноры матрицы  $A$  и матрицы  $A - lE$  положительны при достаточно малом  $0 < l < \lambda_{\min}$ , а это означает, что функция  $f(x)$  сильно выпукла.

**Пример 5.2.** Показать, что квадратичная функция из предыдущего примера  $f(x) = 2x_1^2 - 2x_1x_2 + 3x_1x_3 + x_2^2 - 2x_2x_3 + 4x_3^2 + x_1 - x_2 + 3x_3 + 5$  сильно выпукла.

Матрица  $H(x) = A$  — положительно определена, так как

$$\Delta_1 = 4 > 0; \quad \Delta_2 = \begin{vmatrix} 4 & -2 \\ -2 & 2 \end{vmatrix} = 4 > 0; \quad \Delta_3 = \begin{vmatrix} 4 & -2 & 3 \\ -2 & 2 & -2 \\ 3 & -2 & 8 \end{vmatrix} = 22 > 0.$$

Следовательно,  $f(x)$  сильно выпукла по свойству 3 квадратичных функций.

Выпуклые квадратичные функции играют важную роль в теории многомерной оптимизации. Некоторые алгоритмы, разработанные с учетом свойств таких функций, позволяют найти их точку минимума за конечное число итераций. Во многих случаях эти алгоритмы оказываются эффективными и для неквадратичных выпуклых функций, так как в достаточно малой окрестности точки минимума  $x^*$  дважды дифференцируемая функция  $f(x)$  с положительно определенной матрицей Гессе  $H(x)$  хорошо аппроксимируется сильно выпуклой квадратичной функцией.

## 5.2. Общие принципы многомерной минимизации

Для численного решения задач безусловной минимизации  $f(x) \rightarrow \min, x \in E_n$  разработано много алгоритмов, использующих итерационные процедуры вида

$$x^{k+1} = \Phi(x^k, x^{k-1}, \dots, x^0), \quad x^0 \in E_n. \quad (5.3)$$

Эти алгоритмы позволяют при определенных условиях построить последовательность  $\{x^k\}$  такую, что

$$\lim_{k \rightarrow \infty} f(x^k) = f^* = \min_{E_n} f(x), \quad \text{если } U^* \neq \emptyset \quad (5.4a)$$

$$\lim_{k \rightarrow \infty} f(x^k) = f^* = \inf_{E_n} f(x), \quad \text{если } U^* = \emptyset \quad (5.4б)$$

где  $U^*$  — множество точек глобального минимума функции  $f(x)$ . Последовательность  $\{x^k\}$ , удовлетворяющая требова-

нию (5.4), называется *минимизирующей* для функции  $f(x)$ . Если, кроме того, для случая  $U^* \neq \emptyset$  дополнительно выполняется условие

$$\lim_{k \rightarrow \infty} \rho(x^k, U^*) = 0, \quad (5.5)$$

то говорят, что минимизирующая последовательность *сходится к множеству*  $U^*$ . Если множество  $U^*$  состоит из единственной точки  $x^*$ , то для сходящейся к  $U^*$  минимизирующей последовательности  $\lim_{k \rightarrow \infty} x^k = x^*$ .

Необходимо отметить, что

1.  $\rho(x, U) = \inf_{y \in U} \rho(x, y)$  есть расстояние от точки  $x$  до множества  $U$ .

2. Минимизирующая последовательность может и не сходиться к точке минимума. Например, для функции

$$f(x) = \frac{x^2}{1+x^4}, \quad x \in E_1, \quad \text{последовательность } x^k = k \text{ является}$$

минимизирующей, но не сходится к единственной точке минимума  $x^* = 0$ . Напротив, минимизирующая последовательность  $x^k = 1/k$  сходится к точке минимума  $x^* = 0$ .

Вопрос о существовании точки минимума обычно решается с помощью теоремы Вейерштрасса: если функция  $f(x)$  непрерывна в  $E_n$  и множество  $U^\alpha = \{x: f(x) \leq \alpha\}$  для некоторого  $\alpha$  непусто и ограничено, то  $f(x)$  достигает глобального минимума в  $E_n$ .

Важной характеристикой сходящихся минимизирующих последовательностей является *скорость сходимости*. Последовательность  $\{x^k\}$  сходится к точке  $x^*$  *линейно* (со скоростью геометрической прогрессии), если существует такое число  $q \in (0, 1)$ , что выполняется неравенство  $\rho(x^k, x^*) \leq q \rho(x^{k-1}, x^*)$ , т.е.  $\rho(x^k, x^*) \leq q^k \rho(x^0, x^*)$ .

Сходимость называется *сверхлинейной*, если

$$\rho(x^k, x^*) \leq q_k \rho(x^{k-1}, x^*), \quad \text{и } q_k \rightarrow +0 \text{ при } k \rightarrow \infty.$$

Наконец, термин *квадратичная* сходимость используется, если  $\rho(x^k, x^*) \leq [c \rho(x^{k-1}, x^*)]^2, \quad c > 0$ .

Установление сходимости последовательности  $\{x^k\}$  из (5.5) и оценка скорости сходимости дают существенную информацию об итерационном процессе. Конкретный вычислительный алгоритм на основе (5.3), в котором может

получаться, вообще говоря, бесконечная последовательность  $\{x^k\}$ , необходимо дополнять критерием окончания итерационного процесса. На практике используются следующие условия прекращения итераций

$$\begin{aligned} \rho(x^{k+1}, x^k) &< \varepsilon_1 \\ |f(x^{k+1}) - f(x^k)| &< \varepsilon_2 \\ \|\nabla f(x^k)\| &< \varepsilon_3 \end{aligned} \quad (5.6)$$

где  $\varepsilon_i$  — заранее заданные параметры точности.

Далее будут рассмотрено несколько двухслойных вычислительных алгоритмов, основанных на рекуррентных формулах вида

$$x^{k+1} = x^k + \alpha_k p^k, \quad k = 0, 1, \dots, \quad (5.7)$$

где вектор  $p^k$  — направление поиска из точки  $x^k$  в точку  $x^{k+1}$ , а число  $\alpha_k$  — величина шага, которая выбирается так, чтобы выполнялось условие  $f(x^{k+1}) < f(x^k)$ . Эти алгоритмы будут различаться способом построения вектора  $p^k$  и выбора шага  $\alpha_k$ .

**Определение.** В итерационном процессе (5.7) производится *исчерпывающий спуск*, если величина шага  $\alpha_k$  находится из решения одномерной задачи минимизации

$$\Phi_k(\alpha) \rightarrow \min_{\alpha}, \quad \Phi_k(\alpha) = f(x^k + \alpha p^k). \quad (5.8)$$

Таким образом, при исчерпывающем спуске на каждом шаге полностью реализуется возможность уменьшить значение целевой функции  $f(x)$  при перемещении из точки  $x^k$  в направлении, коллинеарном вектору  $p^k$ . Величина шага  $\alpha_k$  может быть найдена с помощью рассмотренных ранее методов одномерной минимизации. В дальнейшем будет использоваться следующее важное свойство исчерпывающего спуска.

**Теорема.** Если функция  $f(x)$  дифференцируема в  $E_n$ , то в итерационном процессе (5.7) с выбором шага с исчерпывающим спуском для любого  $k \geq 1$  выполняется условие

$$(\nabla f(x^{k+1}), p^k) = 0. \quad (5.9)$$

Запишем необходимое условие минимума функции одной переменной  $\Phi_k(\alpha)$  из (5.8), используя правило дифференцирования сложной функции

$$\frac{d\Phi_k(\alpha)}{d\alpha} = \sum_{j=1}^n \frac{\partial f(x^{k+1})}{\partial x_j} \frac{dx_j^{k+1}}{d\alpha} = 0.$$

Учитывая, что  $x_j^{k+1} = x_j^k + \alpha p_j^k$ , поэтому  $\frac{dx_j^{k+1}}{d\alpha} = p_j^k$ . Отсюда

получаем условие (5.9).

Геометрическая иллюстрация соотношения (5.9) в  $E_2$  выглядит так. При перемещении из точки  $x^k$  вдоль прямой, задаваемой вектором  $p^k$  в направлении убывания функции, происходит пересечение линий уровня функции  $f(x)$ . Это происходит до тех пор, пока либо не будет достигнута стационарная точка ( $\nabla f(x^{k+1}) = 0$ ), либо прямая не коснется в точке  $x^{k+1}$  некоторой линии уровня функции  $f(x)$ . Равенство (5.9) и есть условие касания (рис.5.1).

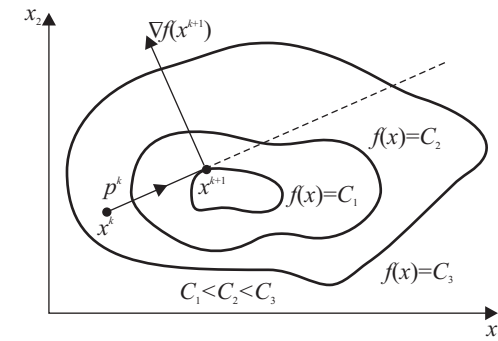


Рис. 5.1. Ортогональность направления  $p^k$  градиенту  $\nabla f(x^{k+1})$  при исчерпывающем спуске

Свойство  $(\nabla f(x^{k+1}), p^k) = 0$  позволяет в явном виде найти величину  $\alpha_k$  для квадратичной функции.

**Теорема.** Для квадратичной функции

$$f(x) = \frac{1}{2}(Ax, x) + (b, x) + c$$

величина  $\alpha_k$  исчерпывающего спуска в итерационном процессе  $x^{k+1} = x^k + \alpha_k p^k$ ,  $k = 0, 1, \dots$  равна

$$\alpha_k = -\frac{(\nabla f(x^k), p^k)}{(Ap^k, p^k)} = -\frac{(Ax^k + b, p^k)}{(Ap^k, p^k)}. \quad (5.10)$$

Действительно, умножив равенство  $x^{k+1} = x^k + \alpha_k p^k$  слева на матрицу  $A$  квадратичной функции  $f(x)$  и прибавив к обеим частям вектор  $b$ , получим  $Ax^{k+1} + b = Ax^k + b + \alpha_k Ap^k$ . Учитывая, что градиент квадратичной функции равен  $\nabla f(x) = Ax + b$ , имеем  $\nabla f(x^{k+1}) = \nabla f(x^k) + \alpha_k Ap^k$ . Подставляя выражение для  $\nabla f(x^{k+1})$  в равенство  $(\nabla f(x^{k+1}), p^k) = 0$ , получаем формулу (5.10).

**Определение.** Направление вектора  $p^k$  называется *направлением убывания* функции  $f(x)$  в точке  $x^k$ , если при всех достаточно малых положительных  $\alpha$  выполняется неравенство  $f(x^k + \alpha p^k) < f(x^k)$ .

В итерационном процессе используются, как правило, направления убывания. Сформулируем признак направления убывания.

**Теорема** (достаточное условие направления убывания). Пусть функция  $f(x)$  дифференцируема в точке  $x^k$ . Если вектор  $p^k$  удовлетворяет условию  $(\nabla f(x^k), p^k) < 0$ , то направление вектора  $p^k$  является направлением убывания.

Из свойства дифференцируемости функции и условия теоремы следует, что  $f(x^k + \alpha p^k) - f(x^k) = (\nabla f(x^k), \alpha p^k) + o(\alpha) = \alpha((\nabla f(x^k), p^k) + o(\alpha)/\alpha) < 0$  при всех достаточно малых  $\alpha > 0$ , т.е. вектор  $p^k$  задает направление убывания функции  $f(x)$  в точке  $x^k$ .

Геометрически условие теоремы  $(\nabla f(x^k), p^k) < 0$  означает, что вектор  $p^k$  составляет тупой угол с градиентом  $\nabla f(x^k)$ .

Итак, пусть функция  $\|\nabla f(x)\| < \varepsilon$  дифференцируема в  $E_n$ . Рассмотрим итерационную процедуру минимизации

$$x^{k+1} = x^k + \alpha_k p^k, \quad k = 0, 1, \dots, \quad (5.11)$$

где направление убывания  $y = x - \alpha \nabla f(x)$  определяется с учетом информации о частных производных функции  $f(x)$ , а величина шага  $\alpha_k > 0$  такова, что

$$f(x^{k+1}) < f(x^k), \quad k = 0, 1, \dots \quad (5.12)$$

Так как функция предполагается дифференцируемой, то в качестве прекращения итераций можно выбрать условие  $\|\nabla f(x^k)\| < \varepsilon$ .

### 5.3. Метод градиентного спуска

Положим в (5.11) на каждом шаге  $p^k = -\nabla f(x^k)$ . Если  $\nabla f(x^k) \neq 0$ , то условие  $(\nabla f(x^k), p^k) < 0$ , очевидно, выполнено. Следовательно, направление вектора  $p^k$  является направлением убывания функции  $f(x)$ , причем в малой окрестности точки  $x^k$  направление  $p^k$  обеспечивает наискорейшее убывание этой функции. Поэтому можно найти такое  $\alpha_k > 0$ , что будет выполняться условие (5.12).

Алгоритм метода градиентного спуска следующий.

Шаг 1. Задать параметр точности  $\varepsilon > 0$ , начальный шаг  $\alpha > 0$ , выбрать  $x \in E_n$  и вычислить  $f(x)$ . Перейти к шагу 2.

Шаг 2. Вычислить  $\nabla f(x)$  и проверить условие достижения точности  $\|\nabla f(x)\| < \varepsilon$ . Если оно выполнено, то вычисления завершить, полагая  $x^* = x$ ,  $f^* = f(x)$ . Иначе перейти к шагу 3.

Шаг 3. Найти  $y = x - \alpha \nabla f(x)$  и  $f(y)$ . Если  $f(y) < f(x)$ , то положить  $x = y$ ,  $f(x) = f(y)$  и перейти к шагу 2, иначе — к шагу 4.

Шаг 4. Положить  $\alpha = \alpha / 2$  и перейти к шагу 3.

Необходимо отметить, что вблизи стационарной точки функции  $f(x)$  величина  $\|\nabla f(x)\|$  становится малой. Это часто приводит к замедлению сходимости последовательности  $\{x^k\}$ . Поэтому в основной формуле (5.11) иногда полагают

$$p^k = -\nabla f(x^k) / \|\nabla f(x^k)\|,$$

используя вместо антиградиента вектор единичной длины в этом же направлении.

Сформулируем теоретическое обоснование сходимости градиентного метода с постоянным шагом  $\alpha_k = \alpha$ , т.е.  $x^{k+1} = x^k - \alpha \nabla f(x^k)$  и оценку скорости сходимости при минимизации выпуклой квадратичной функции

$$f(x) = \frac{1}{2}(Ax, x) + (b, x) + c.$$

**Теорема.** Пусть симметрическая матрица  $A$  квадратичной функции  $f(x)$  положительно определена,  $l$  и  $L$  — ее наименьшее и наибольшее собственные значения ( $0 < l \leq L$ ). Тогда при любых  $\alpha \in (0, 2/L)$  и  $x^0 \in E_n$  итерационная последовательность  $x^{k+1} = x^k - \alpha \nabla f(x^k)$  сходится к единственной точке глобального минимума  $x^*$  функции  $f(x)$  линейно (т.е. со скоростью геометрической прогрессии), а именно

$$\rho(x^k, x^*) \leq q^k \rho(x^0, x^*),$$

где  $q = \max\{|1 - \alpha l|, |1 - \alpha L|\}$ .

Так как матрица  $A$  положительно определена, то функция  $f(x)$  сильно выпукла. Следовательно, точка  $x^*$  существует и единственна. Градиент  $\nabla f(x^*)$  в этой точке обращается в нуль, поэтому  $\nabla f(x^k) = Ax^k + b = Ax^k + b - Ax^* - b = A(x^k - x^*)$ .

Оценим норму разности

$$\begin{aligned} \|x^k - x^*\| &= \|x^{k-1} - \alpha \nabla f(x^{k-1}) - x^*\| = \|x^{k-1} - x^* - \alpha A(x^{k-1} - x^*)\| = \\ &= \|(E - \alpha A)(x^{k-1} - x^*)\| \end{aligned}$$

Последняя величина может быть оценена сверху

$$\|x^k - x^*\| \leq \|E - \alpha A\| \cdot \|x^{k-1} - x^*\| \leq q \|x^{k-1} - x^*\| \leq q^k \|x^0 - x^*\|.$$

Здесь величина  $q$  — оценка нормы матрицы через величину ее собственных значений:  $\|E - \alpha A\| \leq q = \max\{|1 - \alpha l|, |1 - \alpha L|\}$ .

Зависимость  $q(\alpha)$  представлена на рис. 5.2, из которого видно, что при  $\alpha \in (0; 2/L)$  величина  $q < 1$ .

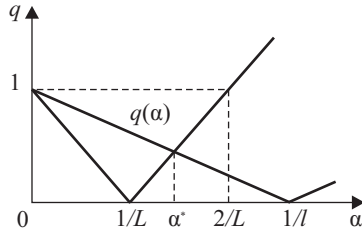


Рис. 5.2. Зависимость оценки знаменателя геометрической прогрессии от величины шага  $\alpha$  градиентного метода

Из формулы для величины  $q$  следует, что она принимает

минимальное значение  $q^* = \frac{L-l}{L+l}$  при  $\alpha = \alpha^* = \frac{2}{L+l}$  (так как

$l < L$ , то  $1 - \alpha l = -(1 - \alpha L)$ ). Поэтому от соотношения между  $L$  и  $l$  существенно зависит число итераций градиентного метода при минимизации выпуклой квадратичной функции. Рассмотрим это на примере квадратичной функции двух переменных.

При  $L = l > 0$  точка минимума  $f(x)$  находится за один шаг.

**Пример. 5.3.** Решить задачу  $f(x) = x_1^2 + x_2^2 \rightarrow \min$  градиентным методом из начальной точки  $x^0 = (1, 1)^T$ , положив  $\alpha = \alpha^*$ .

В данном случае матрица квадратичной функции  $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ ,

отсюда  $l = L = 2$ . Поэтому  $\alpha^* = \frac{2}{L+l} = \frac{1}{2}$  и  $x^1 = x^0 - \frac{1}{2} \nabla f(x^0) = (0, 0)^T$ .

Следовательно,  $x^1 = x^*$ .

При  $l = L$  линии уровня целевой функции  $f(x)$  — это концентрические окружности, поэтому направление антиградиента указывает на их центр, т.е. на точку глобального минимума  $f(x)$ .

Если  $L \gg l > 0$ , то линиями уровня являются эллипсы, полуоси которых сильно различаются. Поэтому направление антиградиента в некоторой точке может сильно отличаться от направления к точке глобального минимума.

**Пример. 5.4.** Из начальной точки  $x^0 = (1, 1)^T$  выполнить несколько итераций поиска точки минимума функции  $f(x) = x_1^2 + 100x_2^2$  градиентным методом, полагая  $\alpha = \alpha^*$ .

Собственные значения матрицы  $A$  этой квадратичной функции:  $l = 2, L = 200$ . Линиями уровня  $f(x)$  являются эллипсы, сильно вытянутые вдоль оси  $Ox_1$ . Поэтому в точке  $x^0$  направление вектора  $-\nabla f(x^0) = (-2, -200)^T$  сильно отличается от направления к точке глобального минимума  $x^* - x^0 = (-1, -1)^T$ . Положив в формуле

$x^{k+1} = x^k - \alpha \nabla f(x^k)$  величину  $\alpha = \alpha^* = \frac{2}{L+l} = \frac{1}{101}$ , получим с учетом

выражения для градиента  $\nabla f(x) = (2x_1, 200x_2)^T$  закон изменения ко-

ординат точек минимизирующей последовательности:  $x_1^{k+1} = \frac{99}{101} x_1^k$ ,

$x_2^{k+1} = -\frac{99}{101} x_2^k$ . Отсюда видно, что последовательность  $\{x^k\}$  сходится

к точке глобального экстремума медленно и траектория сходимости имеет ярко выраженный зигзагообразный характер.

В курсах линейной алгебры и численных методов для симметрической положительно определенной матрицы вводится *число обусловленности*, которое равно  $\mu = L/l$  — отношению наибольшего и наименьшего собственных значений матрицы. В задаче минимизации сильно выпуклой функции  $f(x)$  эта величина для матрицы ее вторых производных (гессиана) характеризует степень вытянутости линий уровня  $f(x) = C$ . Если  $\mu$  велико, то линии уровня сильно вытянуты и говорят,



что функция имеет *овражный характер*, т.е. резко меняется по одним направлениям и слабо — по другим. В этих случаях задачу минимизации называют *плохо обусловленной*. Если же число  $\mu$  близко к единице, то линии уровня близки к окружностям и задача минимизации является *хорошо обусловленной*.

Вычисление градиента на каждом шаге, позволяющее каждый раз двигаться в направлении все более быстрого убывания целевой функции, может в то же время замедлить вычислительный процесс. Дело в том, что подсчет градиента — обычно гораздо более сложная операция, чем подсчет самой функции. Поэтому часто пользуются модификацией градиентного метода, получившей название *метода наискорейшего спуска*.

Согласно этому методу после вычисления в начальной точке градиента функции делают в направлении антиградиента не маленький шаг, а передвигаются до тех пор, пока функция убывает. Достигнув точки минимума на выбранном направлении, снова вычисляют градиент функции и повторяют описанную процедуру. При этом градиент вычисляется гораздо реже, поскольку это происходит только при смене направлений движения.

## 5.4. Метод наискорейшего спуска

Данный метод является вариантом градиентного спуска. Здесь также полагают  $p^k = -\nabla f(x^k)$ , но величина шага  $\alpha_k$  из (5.11) находится в результате решения задачи одномерной оптимизации

$$\Phi_k(\alpha) \rightarrow \min, \quad \Phi_k(\alpha) = f(x^k - \alpha \nabla f(x^k)), \quad \alpha > 0, \quad (5.13)$$

т.е. на каждой итерации в направлении антиградиента  $-\nabla f(x^k)$  совершается исчерпывающий спуск.

Алгоритм метода следующий.

Шаг 1. Задать параметр точности  $\varepsilon > 0$ , выбрать  $x \in E_n$ . Вычислить  $f(x)$ . Перейти к шагу 2.

Шаг 2. Вычислить  $\nabla f(x)$  и проверить условие достижения точности:  $\|\nabla f(x)\| < \varepsilon$ . Если оно выполнено, вычисления завершить, полагая  $x^* = x$ ,  $f^* = f(x)$ . Иначе — перейти к шагу 3.

Шаг 3. Решить задачу одномерной оптимизации (5.13) для  $x^k = x$ , т.е. найти  $\alpha^*$ . Положить  $x = x - \alpha^* \nabla f(x)$  и перейти к шагу 2.

Следует отметить, что решение задачи (5.13) можно получить одним из рассмотренных для одномерной минимизации способов. Если, кроме того, функция  $f(x)$  квадратична, то для величины шага исчерпывающего спуска можно применить

$$\text{формулу } \alpha_k = -\frac{(\nabla f(x^k), p^k)}{(Ap^k, p^k)} = -\frac{(Ax^k + b, p^k)}{(Ap^k, p^k)}, \text{ полученную выше, положив } p^k = -\nabla f(x^k).$$

**Пример 5.5.** Рассмотрим функцию  $f(x) = x_1^2 + 100x_2^2$  и используем метод наискорейшего спуска для решения задачи ее минимизации из начальной точки  $x^0 = (1, 1)$ .

$$\text{Найдем компоненты градиента } \frac{\partial f}{\partial x_1} = 2x_1, \quad \frac{\partial f}{\partial x_2} = 200x_2 \text{ и гессиан.}$$

Величину шага исчерпывающего спуска  $\alpha_k$  найдем по формуле для квадратичной функции. Результаты вычислений приведены в табл. 5.1.

Таблица 5.1

Результаты вычислений задачи методом наискорейшего спуска

| Номер итерации | $x_1^k$             | $x_2^k$              | $f(x^k)$            | $\ \nabla f(x^k)\ $  |
|----------------|---------------------|----------------------|---------------------|----------------------|
| 0              | 1                   | 1                    | 101                 | 200                  |
| 1              | $9,9 \cdot 10^{-1}$ | $-9,9 \cdot 10^{-5}$ | $9,8 \cdot 10^{-1}$ | 1,98                 |
| 2              | $9,7 \cdot 10^{-3}$ | $9,7 \cdot 10^{-3}$  | $9,5 \cdot 10^{-3}$ | 1,94                 |
| 3              | $9,6 \cdot 10^{-3}$ | $-9,6 \cdot 10^{-7}$ | $9,2 \cdot 10^{-5}$ | $1,92 \cdot 10^{-2}$ |
| ...            | ...                 | ...                  | ...                 | ...                  |

Полученные результаты можно сравнить с решением этой же задачи методом градиентного спуска.

Отметим главный недостаток градиентных методов, который заключается в неприемлемо низкой скорости сходимости в случаях, когда поверхности уровня целевой функции сильно «вытянуты», т.е. имеют *овражный характер*. Геометрически происходит следующее: генерируемые методом приближения быстро спускаются почти на «дно оврага», а затем начинают «прыгать» с одного его «склона» на другой, причем длина осуществляемых шагов все время уменьшается. Особенно сильный замедляющий эффект это явление оказывает в том случае, когда «дно оврага» изогнуто.



В следующих разделах главы будут рассмотрены методы, обладающие высокой скоростью сходимости и значительно менее чувствительные к овражному характеру целевой функции задачи.

Рассмотрим один важный класс методов, который не требует информации о вторых производных целевой функции, как метод Ньютона, однако обладает более высокой скоростью сходимости, чем градиентные методы. Повышение скорости сходимости здесь не связано с существенным ростом трудоемкости итерации, а обусловлено идеей, лежащей в основе метода.

### 5.5. Метод сопряженных направлений

Этот итерационный метод, предназначенный для отыскания минимума выпуклой квадратичной функции, приводит к результату за конечное число шагов. От такого метода разумно ожидать высокой эффективности и в случае выпуклой неквадратичной функции.

Разберем идею метода на примере поиска минимума в случае функции двух переменных. Линиями уровня сильно выпуклой квадратичной функции являются эллипсы. Пусть  $p^1, p^2$  — направления главных осей этих эллипсов (они могут быть найдены как ортонормированный базис из собственных векторов матрицы  $A$  квадратичной функции). Если выбрать произвольную точку  $x^0 \in E_2$  и выполнить итерационную процедуру  $x^{k+1} = x^k + \alpha_k p^k$ ,  $k=0, 1, \dots$ , где величина  $\alpha_k$  находится из условия исчерпывающего спуска, то, очевидно, потребуется не более двух шагов для отыскания точки  $x^*$ .

Такого же результата можно достичь и другим способом. Выберем некоторое направление  $p^1$  и две точки  $x^0$  и  $y^0$  такие, что векторы  $x^0 - y^0$  и  $p^1$  были неколлинеарны. Выполнив исчерпывающий спуск из точек  $x^0$  и  $y^0$  в направлении  $p^1$ , получим точки  $x^1$  и  $y^1$ . По свойству исчерпывающего спуска в точках  $x^1$  и  $y^1$  имеет место касание соответствующих прямых (направлений убывания) и эллипсов (линий уровня целевой функции). По свойству эллипсов, точки  $x^*$ ,  $x^1$  и  $y^1$  расположены на одной прямой. Поэтому, полагая  $p^2 = x^1 - y^1$ , и решая задачу  $f(x^1 + \alpha p^2) \rightarrow \min$ , мы находим точку  $x^*$ . Таким образом, и в этом случае решение задачи минимизации сильно выпуклой квадратичной функции будет получено за конечное число шагов.

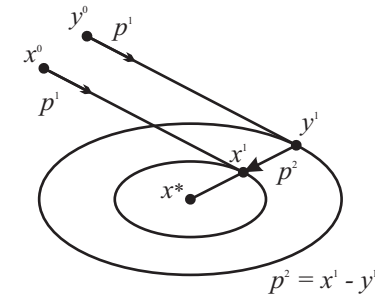


Рис. 5.3. Определение направления  $p^2$  и положение точки  $x^*$  в  $E_2$

Рассмотренному на примере квадратичной функции двух переменных методу может соответствовать такой алгоритм.

Шаг 1. Выбрать начальную точку  $x^0 \in E_2$ . Перейти к шагу 2.

Шаг 2. Положить  $p^1 = e^1$  и найти точку  $x^1$  с помощью исчерпывающего спуска из точки  $x^0$  по направлению  $p^1$

$$f(x^1) = \min_{\alpha \in E} f(x^0 + \alpha p^1). \text{ Перейти к шагу 3.}$$

Шаг 3. а) положить  $y^0 = x^1 + e^2$ ;

б) найти точку  $y^1$  из условия исчерпывающего спуска из точки  $y^0$  по направлению  $p^1$

$$f(y^1) = \min_{\alpha \in E} f(y^0 + \alpha p^1);$$

в) положить  $p^2 = x^1 - y^1$ , найти точку  $x^2$  из условия  $f(x^2) = \min_{\alpha \in E} f(x^1 + \alpha p^2)$ , вычисления закончить, положив  $x^* = x^2$ .

Графическая иллюстрация работы алгоритма представлена на рис. 5.4. Поиск точки минимума проводится по так называемым сопряженным направлениям.

**Определение.** Ненулевые векторы  $p^1, \dots, p^k$  называются сопряженными относительно матрицы  $A$  размера  $(n \times n)$  или  $A$  — ортогональными, если

$$(Ap^i, p^j) = 0, \quad i \neq j, \quad i, j = 1, \dots, k. \quad (5.14)$$

**Пример 5.6.** Направления  $p^1, p^2$ , использованные в описанном выше алгоритме минимизации квадратичной функции двух переменных, являются  $A$  — ортогональными.

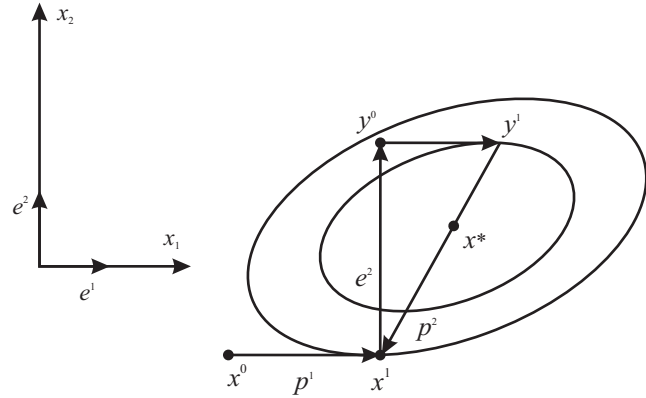


Рис. 5.4. Минимизация функции с помощью сопряженных направлений в  $E_2$

Рассмотрим скалярное произведение

$$(Ap^2, p^1) = (A(x^1 - y^1), p^1) = (\nabla f(x^1), p^1) - (\nabla f(y^1), p^1).$$

Так как точки  $x^1$  и  $y^1$  получены в результате исчерпывающего спуска по направлению  $p^1$ , то оба скалярных произведения  $(\nabla f(x^1), p^1)$  и  $(\nabla f(y^1), p^1)$  равны нулю, поэтому и  $(Ap^2, p^1) = 0$ .

**Утверждение.** Система из  $n$  векторов  $p^1, \dots, p^n$ , сопряженных относительно положительно определенной матрицы  $A$ , линейно независима.

Предположим противное, т.е. что существует линейная комбинация, равная нулю

$$\sum_{i=1}^n \gamma_i p^i = 0, \quad (5.15)$$

где не все  $\gamma_i = 0$ , например,  $\gamma_k \neq 0$ . Умножим обе части равенства (5.15) скалярно на вектор  $Ap^k$ . Тогда, с учетом свойства  $A$  — ортогональности (5.14), получим  $\gamma_k (Ap^k, p^k) = 0$ . В силу положительной определенности матрицы  $A$  для ненулевого вектора  $p^k$  последняя квадратичная форма положительна, следовательно,  $\gamma_k = 0$ . Полученное противоречие доказывает утверждение.

Таким образом,  $n$  ненулевых  $A$  — ортогональных векторов образуют базис в  $E_n$ . Рассмотрим минимизацию в  $E_n$

квадратичной функции  $f(x) = \frac{1}{2}(Ax, x) + (b, x) + c$  с положительно определенной матрицей  $A$ , используя итерационный процесс

$$x^k = x^{k-1} + \alpha_k p^k, \quad k = 1, 2, \dots, \quad (5.16)$$

где векторы  $p^k$   $A$  — ортогональны.

**Утверждение.** Если в итерационном процессе (5.16) на каждом шаге используется исчерпывающий спуск, то величина шага  $\alpha_k$  будет

$$\alpha_k = -\frac{(\nabla f(x^0), p^k)}{(Ap^k, p^k)}, \quad k = 1, 2, \dots \quad (5.17)$$

Раскрывая рекуррентную формулу (5.16), получаем

$$x^k = x^0 + \sum_{i=1}^k \alpha_i p^i. \quad (5.18)$$

Из формулы (5.18), учитывая выражение для градиента квадратичной функции  $\nabla f(x) = Ax + b$ , находим

$$\nabla f(x^k) = \nabla f(x^0) + \sum_{i=1}^k \alpha_i Ap^i.$$

Умножая обе части этого равенства скалярно на  $p^k$  и учитывая условие исчерпывающего спуска по направлению  $p^k$ :  $(\nabla f(x^k), p^k) = 0$  и  $A$  — ортогональность векторов, получаем

$$(\nabla f(x^0), p^k) + \alpha_k (Ap^k, p^k) = 0.$$

Так как матрица  $A$  положительно определена, квадратичная форма  $(Ap^k, p^k) > 0$  и для величины шага  $\alpha_k$  получаем выражение (5.17).

**Теорема.** Последовательный исчерпывающий спуск по  $A$  — ортогональным направлениям (5.16) приводит к точке минимума квадратичной формы не более чем за  $n$  шагов.

Доказать самостоятельно. Предположить, что существуют  $u_k \neq \alpha_k$ , и получить, что они совпадают.

Вопрос о нахождении базиса из  $A$  — ортогональных векторов в пространстве  $E_n$  решается неоднозначно. В качестве такого базиса можно, например, взять ортогональный базис из собственных векторов матрицы  $A$ . Однако их поиск при  $n > 2$  представляет собой самостоятельную и довольно сложную задачу.

Итерационный процесс  $x^k = x^{k-1} + \alpha_k p^k$ ,  $k=1, 2, \dots$  последовательной одномерной минимизации по сопряженным направлениям  $p^k$  можно организовать и без предварительного построения векторов  $p^1, \dots, p^n$ , последовательно находя их в процессе минимизации, как это было сделано выше в примере с минимизацией функции двух переменных. И в этом случае для квадратичной функции с положительно определенной матрицей  $A$  для нахождения минимума достаточно конечное число шагов. Если  $f(x)$  не является квадратичной функцией или вспомогательные задачи одномерной минимизации решаются приближенно, потребуются дополнительные вычисления.

Метод сопряженных направлений, рассмотренный выше, относится к числу наиболее эффективных методов минимизации выпуклых квадратичных функций. Его недостатком является необходимость решать довольно большое количество задач одномерной минимизации.

## 5.6. Метод сопряженных градиентов

При использовании методов градиентного и наискорейшего спуска в итерационной процедуре

$$x^{k+1} = x^k + \alpha_k p^k, \quad k=0, 1, \dots$$

в качестве направления убывания функции  $f(x)$  использовалось направление антиградиента:  $p^k = -\nabla f(x^k)$ . Однако такой выбор направления убывания не всегда бывает удачным. В частности, для плохо обусловленных задач минимизации направление антиградиента в точке  $x^k$  может значительно отличаться от направления к точке минимума  $x^*$ . В результате траектория приближения к точке минимума имеет зигзагообразный характер. Воспользуемся другим подходом, идея которого была изложена при построении метода сопряженных направлений. Будем определять направления спуска  $p^k$  не только через вектор антиградиента  $-\nabla f(x^k)$ , но и через направление спуска  $p^{k-1}$  на предыдущем шаге. Это позволит более полно учесть особенности функции  $f(x)$  при нахождении ее минимума.

Используется итерационный процесс

$$x^{k+1} = x^k + \alpha_k p^k, \quad k=0, 1, \dots; \quad x^0 \in E_n, \quad p^0 = -\nabla f(x^0), \quad (5.19)$$

в котором величина шага  $\alpha_k$  находится из условия исчерпывающего спуска по направлению  $p^k$ . Далее, после вычисления очередной точки  $x^{k+1}$ ,  $k=0, 1, \dots$ , новое направление поиска  $p^{k+1}$  находится по формуле, отличной от антиградиента:

$$p^{k+1} = -\nabla f(x^{k+1}) + \beta_k p^k, \quad k=0, 1, \dots, \quad (5.20)$$

где коэффициенты  $\beta_k$  выбираются так, чтобы при минимизации квадратичной функции  $f(x)$  с положительно определенной матрицей  $A$  получалась последовательность  $A$ —ортогональных векторов  $p^0, p^1, \dots$ . Из условия  $(Ap^{k+1}, p^k) = 0$  имеем:

$$\beta_k = \frac{(A\nabla f(x^{k+1}), p^k)}{(Ap^k, p^k)}. \quad (5.21)$$

Ранее, при обсуждении метода сопряженных направлений было показано, что для квадратичной функции шаг исчерпывающего спуска по направлению  $p^k$  равен

$$\alpha_k = -\frac{(\nabla f(x^k), p^k)}{(Ap^k, p^k)}. \quad (5.22)$$

**Утверждение.** Итерационный процесс (5.19)—(5.22) минимизации квадратичной функции с положительно определенной симметрической матрицей  $A$  дает точки  $x^0, \dots, x^k$  и векторы  $p^0, \dots, p^k$  такие, что если  $\nabla f(x^i) \neq 0$  при  $0 \leq i < k \leq n-1$ , то векторы  $p^0, \dots, p^k$   $A$ —ортогональны, а градиенты  $\nabla f(x^0), \dots, \nabla f(x^i)$  взаимно ортогональны.

Так как направления  $p^k$  в (5.20) являются  $A$ —ортогональными, то метод гарантирует нахождение точки минимума сильно выпуклой квадратичной функции не более чем за  $n$  шагов.

С учетом взаимной ортогональности градиентов  $\nabla f(x^i)$  и условий исчерпывающего спуска по направлениям  $p^k$  можно упростить выражения (5.21) и (5.22) для  $\alpha_k$  и  $\beta_k$ . В результате получим, что итерационный процесс метода сопряженных градиентов описывается соотношениями

$$x^{k+1} = x^k + \alpha_k p^k, \quad k=0, 1, \dots; \quad x^0 \in E_n, \quad p^0 = -\nabla f(x^0), \quad (5.23)$$

$$f(x^k + \alpha_k p^k) = \min_{\alpha > 0} f(x^k + \alpha p^k), \quad k=0, 1, \dots, \quad (5.24)$$

$$p^{k+1} = -\nabla f(x^{k+1}) + \beta_k p^k, \quad k=0, 1, \dots, \quad (5.25)$$

$$\beta_k = \frac{\|\nabla f(x^{k+1})\|^2}{\|\nabla f(x^k)\|^2}, \quad k=1, 2, \dots \quad (5.26)$$

Следует отметить, что выражение для коэффициента  $\beta_k$  не содержит в явном виде матрицу  $A$  квадратичной формы. Поэтому метод сопряженных градиентов может применяться для минимизации неквадратичных функций.

Итерационный процесс (5.23)–(5.26) может не приводить к точке минимума неквадратичной функции  $f(x)$  за конечное число итераций. Более того, точное определение  $\alpha_k$  из условия (5.22) возможно лишь в редких случаях, а вектора  $p^k$  на образуют, вообще говоря,  $A$  — ортогональную систему относительно какой-либо матрицы  $A$ . Поэтому реализация каждой итерации метода будет сопровождаться неизбежными погрешностями. Эти погрешности, накапливаясь, могут привести к тому, что векторы  $p^k$  перестанут указывать направление убывания функции и сходимость метода может нарушаться. Поэтому в методе сопряженных градиентов применяется практический прием — через каждые  $N$  шагов производят обновление метода, полагая  $\beta_{mN} = 0$ ,  $m=1, 2, \dots$ . Номера  $mN$  называют моментами обновления метода, или *рестарта*. Часто полагают  $N=n$  — размерности пространства  $E_n$ . Если  $N=1$ , то получается частный случай метода сопряженных градиентов — метод наискорейшего спуска.

Вблизи точки минимума дважды дифференцируемая функция с положительно определенной матрицей Гессе  $H(x^*)$ , как правило, достаточно хорошо аппроксимируется квадратичной функцией. Поэтому можно надеяться на хороший результат применения метода сопряженных градиентов для функций такого вида.

**Пример 5.7.** Методом сопряженных градиентов найти точку минимума функции  $f(x) = 4x_1^2 + 3x_2^2 - 4x_1x_2 + x_1$  из начальной точки  $x^0 = (0, 0)^T$ .

Итерация 1.

Шаг 1. Положим  $\varepsilon = 0,01$ ,  $x^0 = (0, 0)^T$ , и найдем  $\nabla f(x^0) = (1, 0)^T$ . Перейдем к шагу 2.

Шаг 2. Положим  $k=0$ ,  $p^0 = -\nabla f(x^0) = (-1, 0)^T$ . Перейдем к шагу 3.

Шаг 3. Решим задачу одномерной минимизации  $f(x^0 + \alpha p^0) \rightarrow \min$ . Получим  $\alpha_0 = 1/8$ . — Здесь применили формулу

$$\alpha_0 = -\frac{(\nabla f(x^0), p^0)}{(Ap^0, p^0)} = -\frac{(Ax^0 + b, p^0)}{(Ap^0, p^0)}.$$

Перейдем к шагу 4.

Шаг 4. Найдем  $x^1 = x^0 + \alpha_0 p^0 = (-1/8, 0)^T$  и  $\nabla f(x^1) = (0, 1/2)^T$ . Точность не достигнута, перейдем к шагу 5.

Шаг 5. Условие  $k+1=n$  не выполняется (нет рестарта), перейдем к шагу 6.

Шаг 6. Найдем коэффициент  $\beta_0 = 1/4$  и новое направление спуска  $p^1 = -\nabla f(x^1) + \beta_0 p^0 = (-1/4, -1/2)^T$ . Перейдем к следующей итерации.

Поскольку  $x^1, \nabla f(x^1)$  и  $p^1 = -\nabla f(x^1) + \beta_0 p^0$  уже вычислены на итерации 1, то итерацию 2 начинаем с шага 3.

Итерация 2.

Шаг 3. Решим задачу одномерной минимизации  $f(x^1 + \alpha p^1) \rightarrow \min$ . Получим  $\alpha = 1/4$ . Перейдем к шагу 4.

Шаг 4. Найдем  $x^2 = x^1 + \alpha_1 p^1 = (-3/16, -1/8)^T$  и  $\nabla f(x^2) = (0, 0)^T$  — задача решена точно.

Таким образом,  $x^* = x^2 = (-3/16, -1/8)^T$ . Решение получено в результате двух итераций метода сопряженных градиентов, поскольку целевая функция квадратична в  $E_2$  и одномерные задачи минимизации на шаге 2 алгоритма решены точно.

*Проверка решения примера.* Решим задачу аналитическим методом.  $f(x) = 4x_1^2 + 3x_2^2 - 4x_1x_2 + x_1$ ;  $\frac{\partial f}{\partial x_1} = 8x_1 - 4x_2 + 1 = 0$ ,  $\frac{\partial f}{\partial x_2} = 6x_2 - 4x_1 = 0$ .

Решая систему этих двух уравнений для нахождения стационарной точки, получим  $x_1 = -3/16, x_2 = -1/8$ .

## 5.7. Метод Ньютона

Пусть функция  $f(x)$  дважды дифференцируема в  $E_n$ . Тогда, с помощью градиента и матрицы Гессе, для нее можно записать разложение в ряд по формуле Тейлора в окрестности точки  $x^k$

$$f(x) = f(x^k) + \nabla f(x^k)^T \Delta x + \frac{1}{2} \Delta x^T H(x^k) \Delta x + o(\|\Delta x\|^2), \quad (5.27)$$

где  $o(\|\Delta x\|^2)$  — сумма всех членов разложения, имеющих порядок выше второго;  $\Delta x^T H(x^k) \Delta x$  — квадратичная форма.

Отсюда следует, что поведение функции  $f(x)$  с точностью до величины порядка  $o(\|\Delta x\|^2)$  может быть описано квадратичной функцией

$$\Phi_k(x) = f(x^k) + \nabla f(x^k)^T \Delta x + \frac{1}{2} \Delta x^T H(x^k) \Delta x. \quad (5.28)$$

Минимизируем функцию  $\Phi_k(x)$  вместо  $f(x)$ . Найдем ее точку минимума  $x^{k+1}$  из условия стационарности точки  $\nabla \Phi_k(x) = 0$

$$\nabla \Phi_k(x) = H(x^k)(x - x^k) + \nabla f(x^k) = 0. \quad (5.29)$$

Пусть матрица Гессе  $H(x)$  положительно определена при всех  $x \in E_n$ , следовательно, она невырождена ( $\det H(x) > 0$ ). Тогда существует обратная матрица  $H^{-1}(x)$ . Отметим, что квадратичная функция (5.28) с положительно определенной матрицей  $H(x^k)$  сильно выпукла и уравнение (5.29) определяет единственную точку глобального минимума функции  $\Phi_k(x)$ . Умножим слева обе части равенства (5.29) на матрицу  $H^{-1}(x^k)$  и найдем точку минимума  $x^{k+1}$  квадратичной функции  $\Phi_k(x)$ , аппроксимирующей  $f(x)$  в окрестности точки  $x = x^k$

$$x^{k+1} = x^k - H^{-1}(x^k) \nabla f(x^k), \quad k = 0, 1, \dots \quad (5.30)$$

Итерационный процесс минимизации функции многих переменных (5.30), начатый из произвольной точки  $x^0 \in E_n$ , называется *методом Ньютона* и является обобщением метода Ньютона в одномерном случае.

Очевидно, для квадратичной функции с положительно определенной матрицей  $A$  применение метода Ньютона обеспечивает получение точки глобального минимума ровно за один шаг из любой точки  $x^0 \in E_n$ .

Для выпуклой функции, отличной от квадратичной, применение этого метода обеспечивает, как правило, быструю сходимость. Дело в том, что на каждом шаге итерационного процесса используется информация о поведении функции  $f(x)$  в окрестности точки  $x^k$ , содержащаяся в значениях не только первых, но и вторых ее частных производных. Поэтому при прочих равных условиях следует ожидать

более быструю сходимость метода Ньютона по сравнению с градиентными методами.

При выборе достаточно хорошего начального приближения  $x^0 \in E_n$  минимизирующая последовательность  $\{x^k\}$  для сильно выпуклой дважды дифференцируемой функции  $f(x)$  сходится к точке минимума с квадратичной скоростью  $\rho(x^k, x^*) \leq [\rho(x^0, x^*)]^2$ . Если точка  $x^0$  выбрана недостаточно

близко к точке  $x^*$ , то последовательность (5.30) может расходиться.

Отметим, что даже сходящаяся последовательность  $\{x^k\}$  метода Ньютона не всегда обеспечивает монотонное убывание  $f(x)$ , т.е. неравенство  $f(x^{k+1}) < f(x^k)$  для некоторых  $k = 0, 1, \dots$  может нарушаться. Этот недостаток устранен в *обобщенном методе Ньютона*

$$x^{k+1} = x^k - \alpha_k H^{-1}(x^k) \nabla f(x^k),$$

где величина  $\alpha_k > 0$  находится на каждом шаге из условия исчерпывающего спуска по направлению  $p^k = -H^{-1}(x^k) \nabla f(x^k)$ . Можно показать, что если целевая функция является сильно выпуклой и ее матрица Гессе  $H(x)$  для любых точек  $x, y \in E_n$  удовлетворяет неравенству  $\|H(x) - H(y)\| \leq L \|x - y\|$ , то при произвольном выборе начальной точки обобщенный метод Ньютона обладает квадратичной скоростью сходимости.

Основным недостатком метода Ньютона можно считать необходимость вычисления и обращения матрицы Гессе на каждой итерации. Итерация этим методом существенно более трудоемка, чем, например, итерации методов градиентного или наискорейшего спусков. Обсуждение модификаций метода Ньютона, позволяющих в определенном смысле избавиться от этого недостатка и при этом сохранить высокую скорость сходимости, приводится в п. 5.8.

**Пример 5.8.** Найти точку минимума функции

$$f(x) = 4x_1^2 + 3x_2^2 - 4x_1x_2 + x_1$$

методом Ньютона из начальной точки  $x^0 = (0, 0)^T$ .

Градиент  $\nabla f(x^0) = (1, 0)^T$ , матрица Гессе  $A = \begin{pmatrix} 8 & -4 \\ -4 & 6 \end{pmatrix}$ . Найдем

обратную матрицу  $A^{-1}(x^0) = \frac{1}{32} \begin{pmatrix} 6 & 4 \\ 4 & 8 \end{pmatrix}$ . Далее



$$x^1 = x^0 - A^{-1}(x^0) \nabla f(x^0) = (-3/16, -1/8)^T. \text{ Так как } \nabla f(x^1) = (0, 0)^T,$$

то задача решена,  $x^* = x^1$ . Целевая функция квадратична, поэтому решение получено за одну итерацию.

**Пример 5.9.** Рассмотрим задачу поиска минимума функции  $f(x) = (x_1^2 - x_2)^2 + (x_1 - 1)^2$ . Зададим параметр точности поиска  $|\nabla f(x)| < \varepsilon, \varepsilon = 10^{-3}$ ; начальную точку  $x^0 = (-1, -2)^T, f(x^0) = 13$ . В результате решения задачи методом Ньютона понадобилось  $N = 5$  итераций, методом наискорейшего спуска —  $N = 96$  итераций, методом сопряженных градиентов —  $N = 8 + 13$  итераций, в зависимости от формулы, по которой вычислялся коэффициент для сопряженных направлений. Обобщенный метод Ньютона справляется с решением задачи за  $N = 5$  итераций, как и классический метод Ньютона.

## 5.8. Квазиньютоновские методы

Среди алгоритмов многомерной минимизации можно выделить группу алгоритмов, которые объединяют достоинства метода наискорейшего спуска и метода Ньютона. Такие алгоритмы называют *квазиньютоновскими методами*. Особенность этих алгоритмов состоит в том, что при их применении нет необходимости обращать матрицу Гессе целевой функции  $f(x)$  и в то же время удается сохранить высокую скорость сходимости, присущую методу Ньютона и его модификациям.

Элементы релаксационной последовательности  $\{x^k\}$  в алгоритмах квазиньютоновских методов минимизации непрерывно дифференцируемой в  $E_n$  целевой функции  $f(x)$  строят в соответствии с рекуррентным соотношением  $x^{k+1} = x^k + \alpha_k p^k$ ,  $k = 0, 1, \dots$ , но направление спуска на каждой  $k$ -й итерации задают в виде

$$p^k = -A_k \nabla f(x_k) = A_k w^k, \quad (5.31)$$

где  $w^k = -\nabla f(x^k)$  — антиградиент целевой функции, а  $A_k$  — положительно определенная матрица порядка  $n$ , обновляемая на каждой итерации. Отметим, что направление, задаваемое на каждой  $k$ -й итерации вектором  $p^k$ , является направлением спуска, так как с учетом (5.31)

$$(\nabla f(x^k), p^k) = -(w^k, A_k w^k) < 0.$$

Нетрудно заметить, что при  $A_k = H^{-1}(x^k)$  метод переходит в классический метод Ньютона, а при  $A_k = \alpha_k$  — в один

из градиентных методов. Матрицы  $\{A_k\}$  определяют таким образом, чтобы последовательность  $\{A_k\}$  при  $k \rightarrow \infty$  имела предел, равный  $H^{-1}(x^*)$ , где  $H(x^*)$  — матрица Гессе целевой функции, вычисленная в точке минимума  $x^*$ .

Известно, что если целевая функция является сильно выпуклой, то алгоритмы метода Ньютона обладают квадратичной скоростью сходимости. Поэтому можно ожидать, что алгоритмы квазиньютоновских методов будут иметь достаточно высокую скорость сходимости, если на каждой  $k$ -й итерации матрица  $A_k$  выбрана близкой к матрице  $H^{-1}(x^k)$ . Используя при конструировании  $A_k$  аппроксимацию матрицы  $H^{-1}(x^k)$  с учетом информации о градиенте целевой функции в точке  $x^k$ , можно существенно упростить процедуру нахождения направления спуска на  $k$ -й итерации. Именно эти соображения лежат в основе построения алгоритмов квазиньютоновских методов.

Квазиньютоновские методы чрезвычайно популярны среди пользователей методов оптимизации в связи с тем, что они сочетают высокую скорость сходимости с невысокой трудоемкостью итерации. Практическое значение этих методов трудно переоценить, хотя с теоретической точки зрения любой квазиньютоновский метод есть не более чем умная реализация фундаментальной идеи метода Ньютона.

**Пример 5.10.** Рассмотрим задачу поиска минимума функции  $f(x) = (x_1^2 - x_2)^2 + (x_1 - 1)^2$ . Зададим параметр точности поиска  $|\nabla f(x)| < \varepsilon, \varepsilon = 10^{-3}$ , начальную точку  $x^0 = (-1, -2)^T, f(x^0) = 13$ . В результате решения задачи методом Ньютона понадобилось  $N = 5$  итераций, методом наискорейшего спуска —  $N = 96$  итераций. Квазиньютоновский метод Давидона-Флетчера-Пауэлла, или ДФП метод, справляется с решением этой задачи за  $N = 7$  итераций.

\* \* \*

Необходимо обратить внимание на то, что алгоритмы безусловной минимизации функций многих переменных можно сравнивать и исследовать как теоретически, так и экспериментально.

Первый подход может быть реализован полностью только для ограниченного класса задач, например, для сильно выпуклых квадратичных функций в  $E_n$ . Мощным инструментом теоретического исследования алгоритмов являются теоремы о сходимости методов. Однако формулировки таких теорем,

использующих при доказательстве аппарат современного функционального анализа, обычно абстрактны. Поэтому этот материал интересен в первую очередь профессионалам-математикам. Кроме того, бывает непросто установить связь полученных результатов с практикой вычислений. Поэтому на практике сравнение алгоритмов часто проводят с помощью вычислительных экспериментов при решении так называемых специальных тестовых задач. Эти задачи могут быть как с малым, так и с большим числом переменных, иметь различный вид нелинейности. Они могут быть составлены специально или возникнуть из практических приложений.

Приведем несколько общепринятых тестов, на которых испытано большинство известных методов минимизации.

#### 1. Тест Розенброка

$$f(x_1, x_2) = 100(x_1 - x_2)^2 + (x_1 - 1)^2 \rightarrow \min, \quad x^0 = (1, 1)^T \quad \varepsilon = 10^{-3}.$$

Здесь некоторые алгоритмы не позволяют получить точку минимума  $(1, 1)^T$  вследствие преждевременного окончания процесса поиска.

#### 2. Многомодальная функция Химмельблау

$$f(x_1, x_2) = 100(x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2.$$

Тест состоит в поиске минимума при точности  $\varepsilon = 10^{-3}$  из различных начальных точек а)  $(5, 5)^T$ , б)  $(5, -5)^T$ , в)  $(0, 0)^T$ , г)  $(-5, -5)^T$ , д)  $(5, 0)^T$ .

#### 3. Функция Пауэлла

$$f(x_1, x_2, x_3, x_4) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4.$$

Функция минимизируется из начальной точки  $x^0 = (3, -1, 0, 1)^T$  при параметре точности  $\varepsilon = 10^{-3}$ . Тест состоит в получении точки минимума  $x^* = (0, 0, 0, 0)^T$ .

### Вопросы и задания для самопроверки

1. Функции какого вида называются квадратичными функциями  $n$  переменных?
2. Чему равны градиент и гессиан квадратичной функции?
3. Каким свойством обладает квадратичная функция с положительно определенной матрицей  $A$ ?
4. При каких  $a, b, c$  функция  $f(x) = ax_1^2 + bx_1x_2 + cx_2^2$  будет выпуклой?

#### 5. Выписать матрицу $A$ квадратичной функции

$$f(x) = x_1^2 + 3x_2^2 + 2x_1x_2 - x_2x_3 + 2x_2 + x_3.$$

#### 6. Какая последовательность $\{x^k\}, k=0, 1, 2, \dots$ называется минимизирующей?

7. Привести пример минимизирующей последовательности, не сходящейся к точке минимума.

8. Что такое скорость сходимости минимизирующей последовательности? Какие скорости сходимости Вы знаете?

9. Когда говорят, что в итерационном процессе  $x^{k+1} = x^k + \alpha_k p^k, k=0, 1, \dots$  производится исчерпывающий спуск?

10. Какие направления дифференцируемой в точке  $x^k$  функции  $f(x)$  называются направлениями убывания? Каков геометрический смысл направления убывания?

11. Какова скорость сходимости метода градиентного спуска для квадратичной функции  $f(x)$  с положительно определенной симметрической матрицей  $A$ , где  $0 < l < L$  — ее наименьшее и наибольшее собственные значения?

12. Когда говорят, что сильно выпуклая функция  $f(x)$  имеет овражный характер? Какие задачи минимизации называются хорошо обусловленными, а какие — плохо обусловленными?

13. В чем состоят преимущества и недостатки метода наискорейшего спуска по сравнению с методом градиентного спуска?

14. Каков главный недостаток градиентных методов?

15. В чем состоит идея метода сопряженных градиентов? Чем этот метод отличается от методов градиентного и наискорейшего спуска?

16. Какова скорость сходимости метода Ньютона для дважды дифференцируемой выпуклой функции  $f(x)$  многих переменных? Какова трудоемкость этого метода?

17. Чем отличаются классический и обобщенный методы Ньютона для сильновыпуклой дважды дифференцируемой функции многих переменных?

18. Сформулировать общий принцип построения квазиньютоновских методов. Какую скорость сходимости следует ожидать от квазиньютоновских методов? Оценить их трудоемкость.

### Задание для численной реализации в среде программирования MATLAB

1. Реализовать в среде MATLAB метод градиентного спуска, метод наискорейшего спуска и метод сопряженных градиентов.

В методе наискорейшего спуска и в методе сопряженных градиентов наряду с аналитическим выражением для величины шага



исчерпывающего спуска для квадратичной функции реализовать решение задач одномерной минимизации методом поразрядного поиска.

2. Протестировать работу реализованных методов на примере овражной функции

$$f(x) = x_1^2 + ax_2^2,$$

при  $a = 1, 100, 500, 1000$ . При  $\epsilon = 10^{-3}$  и  $\epsilon = 10^{-5}$  сравнить скорость работы методов при различных значениях параметра  $a$  по числу итераций и по числу вызовов совокупности значений функций и производных.

3. Выбрать для выполнения работы тестовую функцию, номер которой соответствует номеру Вашего компьютера. Например, для компьютера №3 это будет функция 3), для компьютера №13 – функция 4):  $13-9=4$ ; для компьютера №23 это будет функция 5):  $23-9 \times 2=5$ .

$$1) f(x) = 64x_1^2 + 126x_1x_2 + 64x_2^2 - 10x_1 + 30x_2 + 13$$

$$2) f(x) = 129x_1^2 - 256x_1x_2 + 129x_2^2 - 51x_1 - 149x_2 - 27$$

$$3) f(x) = 254x_1^2 + 506x_1x_2 + 254x_2^2 + 50x_1 + 130x_2 - 111$$

$$4) f(x) = 151x_1^2 - 300x_1x_2 + 151x_2^2 + 33x_1 + 99x_2 + 48$$

$$5) f(x) = 85x_1^2 + 168x_1x_2 + 85x_2^2 + 29x_1 - 51x_2 + 83$$

$$6) f(x) = 211x_1^2 - 420x_1x_2 + 211x_2^2 - 192x_1 + 50x_2 - 25$$

$$7) f(x) = 194x_1^2 + 376x_1x_2 + 194x_2^2 + 31x_1 - 229x_2 + 4$$

$$8) f(x) = 45x_1^2 - 88x_1x_2 + 45x_2^2 + 102x_1 + 268x_2 - 21$$

$$9) f(x) = 99x_1^2 + 196x_1x_2 + 99x_2^2 - 95x_1 - 9x_2 + 91$$

4. Графически отобразить линии уровня выбранной функции. Сравнить эффективность методов градиентного, наискорейшего спуска, а так же метода сопряженных градиентов для задачи п.2 при  $a=250$  и тестовой функции п.3 по числу итераций. *Объяснить полученные результаты.*

5. Минимизировать функцию Розенброка

$$f(x) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2$$

с точностью  $\epsilon = 10^{-3}$  и  $\epsilon = 10^{-5}$ , выбрав начальную точку  $x^0 = (-1, 1)^T$ .

Для решения задачи использовать метод Ньютона, один из квазиньютоновских методов (воспользоваться одним из следующих квазиньютоновских алгоритмов из дополнительной литературы: метод Давидона – Флетчера – Пауэлла (ДФП); метод Бroyдена – Флетчера – Шенно (БФШ); метод Мак-Кормика) и метод сопряженных градиентов. Определить, сколько итераций потребуется каждому методу для того, чтобы разность между численным и точ-

ным решением  $x^* = (1, 1)^T$  была меньше  $\epsilon$ . *Сравнить эффективность методов.*

6. Исследовать работу метода сопряженных градиентов в зависимости от частоты обновлений  $N$  на примере функции Розенброка. *Какое значение можно назвать оптимальным?*

7. На примере функции Химмельблау

$$f(x) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2 - 7)^2$$

рассмотреть особенности применения градиентных методов для минимизации многомодальных функций. В качестве начального приближения взять точки  $(0, 0)$  и  $(-5, 0)$ . Как зависит работа рассматриваемых алгоритмов от выбора начального приближения?

8. Результаты работы сохранить для использования в задании для численной реализации гл. 6.

## Глава 6

# ПРЯМЫЕ МЕТОДЫ БЕЗУСЛОВНОЙ МИНИМИЗАЦИИ МНОГОМЕРНЫХ ЗАДАЧ

### 6.1. Проблема минимизации многомерных задач

Подавляющее число реальных задач оптимизации, представляющих практический интерес, являются многомерными, причем число аргументов целевой функции иногда может быть весьма большим.

Например, в задаче о химическом производстве (см. гл. 2), где целевая функция зависит от температуры, при определенном ее выборе производительность (выход интересующего нас продукта) оказывается максимальной. Однако, наряду с температурой, производительность зависит также от давления, соотношения между концентрациями вводимого сырья, катализаторов и ряда других факторов. Таким образом, задача выбора наилучших условий химического производства — это типичная многомерная задача оптимизации.

Как и в одномерном случае, характер задачи и возможные методы решения существенно зависят от информации о целевой функции, которая доступна в процессе ее исследования. В одних случаях, рассмотренных выше в градиентных методах минимизации, функция задается аналитически, являясь при этом дифференцируемой функцией. Тогда можно вычислить ее частные производные, получить явное выражение для градиента, определяющего в каждой точке направления возрастания и убывания функции, и использовать эту информацию для решения задачи. В других случаях никакого аналитического выражения для целевой функции нет, а имеется лишь возможность определить ее значение в любой точке рассматриваемой области (с помощью расчетов, в результате эксперимента и т. д.). В таких

задачах в процессе решения можно найти значение целевой функции лишь в конечном числе точек, и по этой информации необходимо приближенно установить ее наименьшее значение для всей области.

В многомерном случае задачи становятся более сложными и трудоемкими, причем трудности при их решении обычно возрастают при увеличении размерности. Для того чтобы лучше почувствовать это, возьмем самый простой по своей идее приближенный метод поиска наименьшего значения функции, который уже обсуждался для одномерных задач. Покроем рассматриваемую область сеткой с шагом  $h$  и определим значения функции в ее узлах. Сравнивая полученные числа между собой, найдем среди них наименьшее и примем его приближенно за наименьшее значение функции для всей области.

Для одномерных задач этот метод носит название *метода перебора*. Иногда он применяется также для решения двумерных, реже трехмерных задач. Однако для задач большей размерности он практически непригоден ввиду резкого увеличения временных затрат, необходимых для проведения расчетов и вычисления значений целевой функции в большом числе узлов.

Действительно, предположим, что целевая функция зависит от семи переменных, а область определения является семимерным кубом, каждая сторона которого при построении сетки делится на 40 частей. Тогда общее число узлов сетки будет равно  $41^7 \approx 10^{11}$ . Пусть вычисление значений функции в одной точке требует 1000 арифметических операций (это немного для функции семи переменных). В этом случае общее число операций составит  $10^{14}$ . Если в распоряжении имеется компьютер с быстродействием 3,2 ГГц, то для решения задачи с помощью данного метода потребуется  $3 \cdot 10^4$  с, что составляет сутки непрерывной работы. Добавление еще одной независимой переменной увеличит это время в 40 раз.

Проведенная оценка показывает, что для больших задач оптимизации метод перебора непригоден. Иногда его заменяют случайным поиском, при котором точки сетки просматриваются не подряд, а в случайном порядке. В результате поиск наименьшего значения целевой функции существенно ускоряется, но теряет свою надежность.

Рассмотрим прямые методы минимизации, позволяющие вести поиск наименьшего значения функции многих

переменных целенаправленно. Необходимо отметить, что алгоритмы прямых методов решения задачи безусловной минимизации опираются только на вычисление значений функции  $f(x)$ . Сначала рассмотрим один из наиболее простых алгоритмов минимизации функции  $f(x)$ , определенной в  $E_n$ . Выберем точку  $x^0 \in E_n$ , называемую обычно *базовой*. Вычислим в ней значение  $f(x^0)$ , а затем построим  $n$ -мерный куб с центром в этой точке и ребрами длиной  $l$ . Множество вершин этого куба вместе с точкой  $x^0$  составляют так называемый *шаблон*. Вычислив значения функции в вершинах куба, выберем в качестве новой базовой точки ту из вершин, в которой значение функции меньше  $f(x^0)$ , и повторим процедуру выбора следующей базовой точки и построения шаблона. Если такой вершины не оказалось, то оставим прежнюю базовую точку  $x^0$  и построим шаблон с уменьшенной (например, вдвое) длиной ребер куба. Процесс поиска закончим тогда, когда длина ребра куба станет меньше заданного числа  $\varepsilon > 0$ . Геометрическая иллюстрация такого алгоритма, называемая обычно *поиском при помощи шаблона*, для случая двух переменных представлена на рис. 6.1.

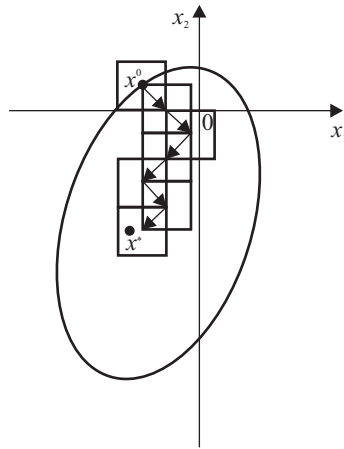


Рис. 6.1. Иллюстрация поиска точки минимума при помощи шаблона

Основной недостаток поиска при помощи шаблона состоит в резком росте количества вычислений значений целевой функции при увеличении размерности пространства  $n$  (порядка  $2^n$ ), а главное преимущество — простота алгоритма. Эффективность этого алгоритма можно повысить. Геометрически это соответствует процедуре построения вокруг базовой точки в качестве шаблона не  $n$ -мерного куба, а конфигурации с меньшим количеством вершин.

Реализация такой идеи привела к разработке методов *симплексного поиска*, получивших к началу 1970-х г. широкое распространение. *Симплекс* (от лат. simplex — простой) — это простейший выпуклый многогранник в  $E_n$  с  $n+1$  вершинами (в  $E_2$  симплекс — совокупность вершин треугольника, в  $E_3$  — тетраэдра). Каждую вершину многогранника называют *вершиной симплекса*, отрезок, соединяющий две вершины — *ребром симплекса*.

В первых разработанных вариантах симплексного поиска был использован *правильный симплекс*, у которого все вершины равноудалены друг от друга (в  $E_2$  правильный симплекс — совокупность вершин равностороннего треугольника, в  $E_3$  — правильного тетраэдра). Позднее был рассмотрен произвольный (деформируемый) симплекс и предложен способ расширения или сжатия симплекса в зависимости от того, удачно ли выбран шаг поиска в направлении убывания функции. Изменение формы симплекса позволяет улучшить процесс адаптации используемой конфигурации к рельефу графика минимизируемой функции.

## 6.2. Минимизация функций по правильному (регулярному) симплексу

Рассмотрим простейший алгоритм симплексного поиска минимума унимодальной целевой функции  $f(x)$ ,  $x \in E_n$ , с использованием регулярного симплекса постоянных размеров. На первом шаге поиска строим регулярный симплекс с вершинами  $x^{1:i}$ ,  $i = 1, \dots, n+1$ , который обозначим  $S_1$ . Построение симплекса можно провести двумя простыми способами.

*Первый способ.* Если  $x^{1:1} \in S_1$  — заданная базовая точка, то координаты  $x_j^{1:i}$ ,  $i = 2, \dots, n+1$ ,  $j = 1, \dots, n$  остальных  $n$  вершин

$x^{1:i} \in S_1$  регулярного симплекса  $S_1$ , имеющего ребра длиной  $l$ , можно вычислить по формулам

$$x_j^{1:i} = \begin{cases} x_j^{1:1} + \frac{\sqrt{n+1}-1}{n\sqrt{2}}l, & i = j+1 \\ x_j^{1:1} + \frac{\sqrt{n+1}+n-1}{n\sqrt{2}}l, & i \neq j+1 \end{cases} \quad (6.1)$$

где  $x_j^{1:1}$ ,  $j=1, \dots, n$  — координаты вершины  $x^{1:1} \in S_1$ . Например, если в  $E_2$  выбрана базовая точка  $x^{1:1} = (0, 0)^T$ , то при длине ребра  $l=2$  остальные две вершины  $x^{1:2}$  и  $x^{1:3}$  имеют координаты  $x_1^{1:2}=0,518$ ,  $x_2^{1:2}=1,932$ ,  $x_1^{1:3}=1,932$ ,  $x_2^{1:3}=0,518$ . На рис. 6.2 приведен построенный симплекс.

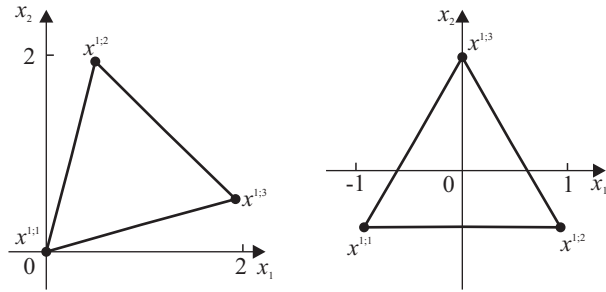


Рис. 6.2. Построение регулярного симплекса двумя способами

*Второй способ.* Если  $x^0 \in E_n$  — заданная базовая точка, определяемая как центр регулярного симплекса  $S_1$ , имеющего ребра длиной  $l$ , то координаты  $x_j^{1:i}$ ,  $i=1, \dots, n$ ,  $j=1, \dots, n$  всех вершин  $x^{1:i} \in S_1$  находятся по формулам

$$x_j^{1:i} = \begin{cases} x_j^0, & j < i-1; \\ x_j^0 + \sqrt{\frac{j}{2(j+1)}}l, & j = i-1; \\ x_j^0 - \frac{1}{\sqrt{2j(j+1)}}l, & j > i-1, \end{cases} \quad (6.2)$$

где  $x_j^0$ ,  $j=1, \dots, n$  — координаты точки  $x^0$ . Например, если в  $E_2$  задана точка  $x^0 = (0, 0)^T$  и  $l=2$ , то вершинами правильного симплекса будут точки  $x^{1:1} = (-1, 000, -0,578)^T$ ,  $x^{1:2} = (1,000, -0,578)^T$ ,  $x^{1:3} = (0,000, 1,156)^T$ . На рис. 6.2 приведен построенный симплекс.

После построения симплекса  $S_1$  на первом шаге поиска, в вершинах  $x^{1:i} \in S_1$ ,  $i=1, \dots, n+1$  вычисляют значения минимизируемой функции и выбирают вершину, в которой значение функции является наибольшим. Пусть для определенности это будет вершина  $x^{1:n+1}$ , если таких вершин несколько, то может быть взята любая из них. Далее по формуле

$$x^{2;n+1} = 2x_c^1 - x^{1:n+1} = \frac{2}{n} \sum_{i=1}^n x^{1:i} - x^{1:n+1} \quad (6.3)$$

находят точку  $x^{2;n+1}$ , симметричную вершине  $x^{1:n+1}$  относительно гиперплоскости, в которой лежат остальные вершины симплекса  $S_1$ . Точка  $x_c^1$  равноудалена от вершин, т.е. является центром регулярного симплекса с  $n$  вершинами, или, иначе говоря, центром масс системы материальных точек, расположенных в вершинах симплекса и имеющих одинаковую массу.

Нахождение точки  $x^{2;n+1}$  называется *отражением вершины*. В результате получают новый регулярный симплекс  $S_2$ , образованный новой вершиной  $x^{2;n+1}$  и  $n$  вершинами  $x^{1:i}$ ,  $i=1, \dots, n$ , принадлежавшими симплексу  $S_1$ . На рис. 6.3 представлена процедура построения нового регулярного симплекса в случае  $E_2$ .

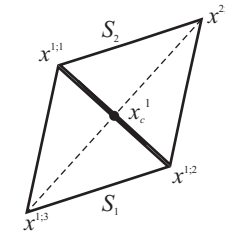


Рис. 6.3. Построение нового симплекса отражением вершины в пространстве  $E_2$

На втором шаге поиска в новой вершине  $x^{2;n+1} \in S_2$  симплекса  $S_2$  вычисляют значение  $f(x^{2;n+1})$  и, если  $f(x^{2;n+1}) < f(x^{1;n+1})$ , то процедуру поиска повторяют далее. При реализации алгоритма удобно на каждом  $k$ -м шаге поиска отражаемой вершине симплекса  $S_k$  присваивать номер  $n+1$ , т.е. обозначать ее  $x^{k;n+1} \in S_k$ . Нумерацию вершин симплекса  $S_k$  назовем *правильной*, если выполнены неравенства

$$f(x^{k;1}) \leq \dots \leq f(x^{k;i}) \leq \dots \leq f(x^{k;n+1}). \quad (6.4)$$

Если на  $k$ -м шаге после отражения вершины  $x^{k;n+1} \in S_k$  в точку  $x^{k+1;n+1}$  и вычисления значения  $f(x^{k+1;n+1})$  окажется, что  $f(x^{k+1;n+1}) \geq f(x^{k;n+1})$ , то следует вернуться к симплексу  $S_k$ , считая отражение вершины  $x^{k;n+1}$  неудачным. После этого, предполагая нумерацию вершин симплекса  $S_k$  правильной, проводят отражение вершины  $x^{k;n} \in S_k$ , получают точку  $x^{k+1;n}$  и сравнивают значения  $f(x^{k+1;n})$  и  $f(x^{k;n})$ . Если  $f(x^{k+1;n}) \geq f(x^{k;n})$ , то отражение и этой вершины считают неудачным и проводят отражение вершины  $x^{k;n-1} \in S_k$  и т.д.

При фиксированной длине ребра симплекса поиск прекращают на шаге с номером  $K$ , если отражения всех вершин симплекса  $S_K$  оказались неудачными. Тогда в качестве оценки искомого наименьшего значения  $f^*$  минимизирующей функции  $f(x)$  можно взять наименьшее из вычисленных значений  $f(x^{K;i})$ ,  $x^{K;i} \in S_K$ ,  $i = 1, \dots, n+1$ . В процессе поиска построена последовательность симплексов  $S_k$ ,  $k = 1, \dots, K$ . Ей соответствует невозрастающая последовательность минимизируемой функции  $\{f(x^{k;1})\}_K$ .

Исследовать свойства последовательности  $\{f(x^{k;1})\}_K$  сложно. Поэтому на практике вместо нее изучают свойства последовательности  $\{f_k\}$  значений минимизируемой функции  $f_k = f(x^k)$ , каждое из которых вычислено в центре  $x^k$  симплекса  $S_k$ ,  $k = 1, \dots, K$ . Процедуру поиска точки  $x^* \in E_n$ , в котором функция  $f(x)$  достигает наименьшего значения, в этом случае проводят в соответствии с рекуррентным соотношением

$$x^{k+1} = x^k + \alpha_k p^k, \quad k = 1, \dots, K,$$

где  $p^k$  — единичный  $n$ -мерный вектор, определяющий направление смещения центра симплекса на  $k$ -м шаге;  $\alpha_k > 0$  — смещение центра симплекса в направлении вектора  $p^k$ . Геометрическая иллюстрация процедуры такого поиска в случае  $E_2$  представлена на рис.6.4.

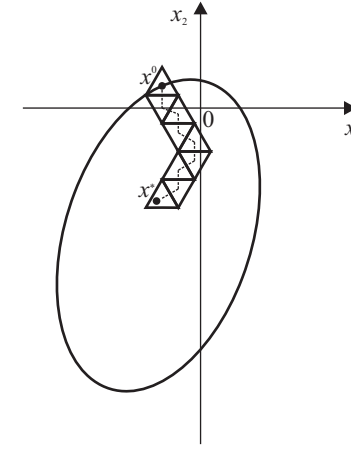


Рис. 6.4. Поиск минимума в  $E_2$  при помощи правильного симплекса

Рассмотрим путь построения эффективных алгоритмов, связанный с идеей изменения размера регулярного симплекса в процессе поиска (алгоритмы *управляемого прямого поиска*). Это важно, когда наименьшее значение функции  $f(x)$  необходимо найти с высокой точностью. Действительно, чем меньше размер симплекса, тем точнее можно локализовать точку  $x^*$ , в которой эта функция достигает наименьшего значения. При этом малый размер симплекса приводит к его уменьшенному смещению в направлении точки  $x^*$ , т.е. к росту числа шагов поиска, связанному с увеличением вычислительных затрат. Большой размер симплекса, наоборот, позволяет за каждый шаг поиска осуществить большее смещение центра симплекса, но обеспечивает лишь грубую локализацию точки  $x^*$ . Поэтому изменение размера симплекса наделяет алгоритм поиска адаптивными свойствами. Рассмотрим подход, при котором размер симплекса изменяется при выполнении определенного условия, проверяемого на каждом шаге поиска.

Пусть  $S_k \in E_n$  — регулярный симплекс на  $k$ -м шаге поиска, имеющий правильную нумерацию вершин. Процедура отыскания вершины  $x^{k+1;n+1} \in S_{k+1}$  нового симплекса  $S_{k+1}$ , включает в себя два этапа — отражение вершины  $x^{k;n+1} \in S_k$  и, если необходимо, уменьшение размера  $S_k$ , называемое *редукцией симплекса*.

*Первый этап.* Отражение вершины  $x^{k;n+1} \in S_k$  осуществля-  
ют в соответствии с формулой, аналогичной (6.3):

$$x^{k+1;n+1} = 2x_c^k - x^{k;n+1} = \frac{2}{n} \sum_{i=1}^n x^{k;i} - x^{k;n+1}, \quad (6.5)$$

где  $x_c^k$  — равноудалена от вершин  $x^{k;i}$ ,  $i=1, \dots, n$ . Затем вы-  
числяют  $f(x^{k+1;n+1})$ .

*Второй этап.* Редукцию симплекса  $S_k$  проводят только при  
выполнении неравенства  $f(x^{k+1;n+1}) \geq f(x^{k;n+1})$ . При этом дли-  
ну всех ребер симплекса уменьшают в  $1/\delta$  раз, где  $\delta \in (0, 1)$  —  
заданный коэффициент редукции, и находят вершины нового  
симплекса  $S_{k+1}$  по формуле

$$x^{k+1;i} = x^{k;1} + \delta(x^{k;i} - x^{k;1}), i = 2, \dots, n+1,$$

сжимая симплекс в  $1/\delta$  раз к вершине  $x^{k;1}$ , в которой значе-  
ние целевой функции меньше, чем в других вершинах симп-  
лекса. Затем осуществляют переход к первому этапу, полагая  
 $k = k+1$ .

Условие прекращения поиска зависит от конкретной  
задачи минимизации. Например, поиск можно вести до  
тех пор, пока длина ребер симплекса не станет меньше  
заранее выбранного значения. Или пока не выполнится  
неравенство

$$\left( \frac{1}{n+1} \sum_{i=1}^{n+1} (f(x^{k;i}) - f(x^k))^2 \right)^{\frac{1}{2}} < \varepsilon,$$

где  $\varepsilon > 0$  — заданное достаточно малое число.

**Пример 6.1.** Используя метод регулярного симплекса, найти реше-  
ние задачи минимизации функции

$$f(x_1, x_2) = 6x_1^2 - 4x_1x_2 + 3x_2^2 + 4\sqrt{5}(x_1 + 2x_2) + 22.$$

Функция квадратичная, имеет минимум  $f^* = -28$  при  $x = (-\sqrt{5}, -2\sqrt{5})^T$ .  
Выберем базовую точку  $x^0 = (-2, 1)^T$  и положим  $\varepsilon = 0,01$ . На рис. 6.5.  
дана графическая иллюстрация процесса поиска при коэффициенте  
редукции  $\delta = 1/2$  и трех значениях начальной длины ребра симплекса,  
указанных в табл. 6.1.

Таблица 6.1.

Поиск минимума функции  
при помощи регулярного симплекса.

| $l$ | $x^*$              | $f(x^*)$ | $N$ |
|-----|--------------------|----------|-----|
| 0,5 | $(-2,233; -4,470)$ | -28,0    | 32  |
| 1,0 | $(-2,233; -4,470)$ | -28,0    | 19  |
| 2,0 | $(-2,233; -4,470)$ | -28,0    | 18  |

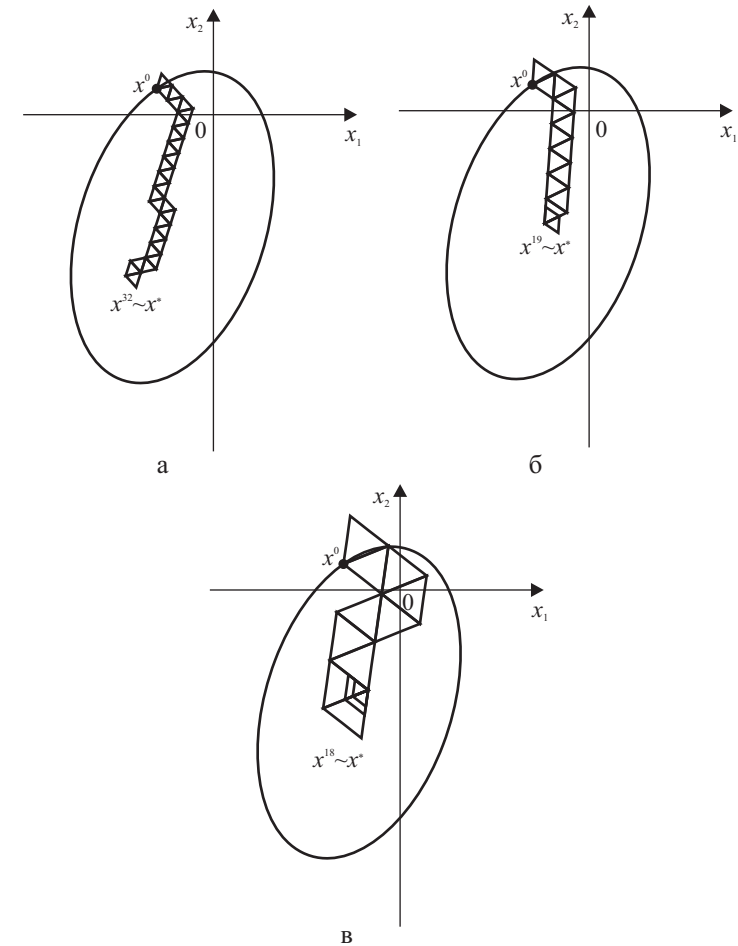


Рис.6.5. Иллюстрация к примеру 6.1



### 6.3. Минимизация функций при помощи нерегулярного симплекса

Симплексные процедуры безусловной минимизации, в которых используются регулярные симплексы, эффективны лишь в тех случаях, когда топография поверхностей уровня целевой функции достаточно проста. В противном случае эффективность применения таких процедур значительно снижается. В частности, это характерно для функций, скорость убывания которых по одному или нескольким направлениям значительно больше, чем по остальным. В таком случае говорят, что график функции имеет овражную структуру.

Примером функции, график которой обладает овражной структурой, является функция Розенброка

$$f(x_1, x_2) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2$$

с минимумом в точке  $x^* = (1, 1)^T$ . Препятствием для успешного поиска минимума  $f(x)$  является в этом случае форма симплекса: регулярный симплекс нельзя изменить так, чтобы он «вытянулся» вдоль «оврага», а это не позволяет продолжить поиск минимума. В этих случаях естественной выглядит идея деформирования симплекса в процессе поиска, т.е. изменения его формы и размера.

Рассмотрим простейшую схему алгоритма прямого поиска по деформируемому симплексу — алгоритм Нелдера—Мида. Этот алгоритм допускает возможность отражения только одной вершины симплекса на каждом шаге поиска. К процедуре отражения при построении нового симплекса добавляются процедуры *сжатия* и *растяжения*. А именно, положение новой вершины  $x^{k+1;n+1}$  вместо вершины  $x^{k;n+1}$ , соответствующей наибольшему значению функции, находится сравнением и выбором наименьшего среди значений целевой функции в точках

$$\begin{aligned} z^1 &= x_c^k - \alpha(x_c^k - x^{k;n+1}) = \frac{1-\alpha}{n} \sum_{i=1}^n x^{k;i} + \alpha x^{k;n+1}, \quad 0 < \alpha < 1; \\ z^2 &= x_c^k + \alpha(x_c^k - x^{k;n+1}) = \frac{1+\alpha}{n} \sum_{i=1}^n x^{k;i} - \alpha x^{k;n+1}, \quad 0 < \alpha < 1; \\ z^3 &= x_c^k + \beta(x_c^k - x^{k;n+1}) = \frac{1+\beta}{n} \sum_{i=1}^n x^{k;i} - \beta x^{k;n+1}, \quad \beta \approx 1; \\ z^4 &= x_c^k + \gamma(x_c^k - x^{k;n+1}) = \frac{1+\gamma}{n} \sum_{i=1}^n x^{k;i} - \gamma x^{k;n+1}, \quad \gamma > 1. \end{aligned} \quad (6.6)$$

Геометрическая иллюстрация этих процедур для пространства  $E_2$  приведена на рис. 6.6 и 6.7. Так как величина  $\alpha \in (0, 1)$ , то выбор точек  $z^1$  и  $z^2$  соответствуют сжатию симплекса;  $\beta \approx 1$ , поэтому выбор точки  $z^3$  соответствует отражению, а  $\gamma > 1$  и выбор точки  $z^4$  приводит к растяжению симплекса. Численные эксперименты показывают, что этот алгоритм хорошо работает в пространстве  $E_n$  для  $n \leq 6$ .

Необходимо отметить, что при деформациях утрачивается свойство правильности исходного симплекса. Поэтому, не стремясь к правильности начального симплекса, его строят из произвольной базовой точки  $x^0 \in E_n$  по формулам

$$x^i = x^0 + a e^i, \quad (6.7)$$

где  $e^i$ ,  $i = 1, \dots, n$  — базисные вектора,  $a$  — параметр симплекса.

На практике хорошо зарекомендовал себя следующий набор параметров  $\alpha, \beta$  и  $\gamma$  для выбора пробных точек  $z^i$  в формулах (6.6):  $\alpha = 1/2$ ,  $\beta = 1$ ,  $\gamma = 2$ .

Для того чтобы избежать сильной деформации симплекса, алгоритм дополняют процедурой обновления. Например, после  $N$  шагов алгоритма из точки  $x^0$  снова строят симплекс по формулам (6.1), (6.2) или (6.7), полагая  $a = \|x^0 - x^n\|$ .

С теоретической точки зрения симплексные методы минимизации слабо исследованы, но практика показывает их работоспособность.

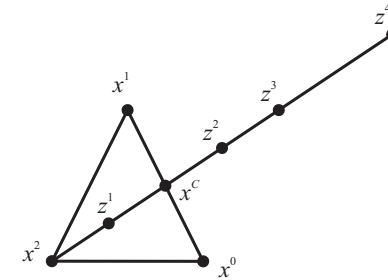


Рис. 6.6. Расположение пробных точек  $z^1, z^2, z^3, z^4$  для перехода к новому симплексу



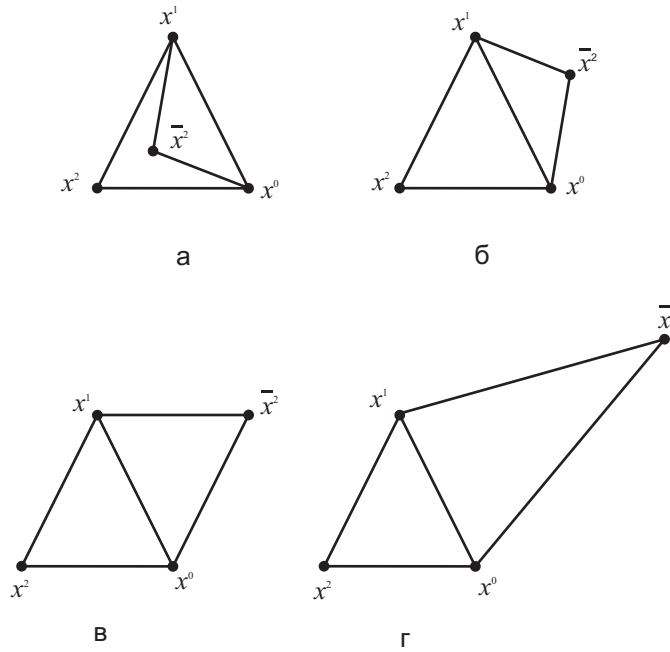


Рис. 6.7. Новые симплексы, полученные в результате процедур сжатия (а, б); отражения (в); растяжения (г)

Вернемся теперь к описанию некоторых вычислительных процедур вида  $x^{k+1} = x^k + \alpha_k p^k$ . В зависимости от способа выбора направления спуска  $p^k$  и шага  $\alpha_k$  получаются различные алгоритмы минимизации.

#### 6.4. Метод циклического покоординатного спуска

Этот метод заключается в последовательной минимизации целевой функции  $f(x)$  с исчерпывающим спуском сначала по направлению первого базисного вектора  $e^1$ , затем второго —  $e^2$  и так далее. После окончания минимизации по направлению последнего базисного вектора  $e^n$  цикл повторяется.

Алгоритм метода следующий.

Шаг 1. Выбрать  $x \in E_n$ , критерий достижения точности  $\rho(x, x^n) < \epsilon_1$  или  $|f(x) - f(x^n)| < \epsilon_2$ , величины  $\epsilon_1, \epsilon_2$ . Найти  $f(x)$ , положить  $j=1$ . Перейти к шагу 2.

Шаг 2. Решить задачу одномерной минимизации  $\Phi(\alpha) = f(x + \alpha e^j) \rightarrow \min, \alpha \in E$ , т.е. найти  $\alpha^*$ . Положить  $\bar{x} = x + \alpha^* e^j$ , вычислить  $f(\bar{x})$ . Перейти к шагу 3.

Шаг 3. Если  $j < n$ , то положить  $x = \bar{x}, j = j+1$  и перейти к шагу 2, иначе перейти к шагу 4.

Шаг 4. Проверить условие достижения точности  $\rho(x, \bar{x}) < \epsilon_1$  или  $|f(x) - f(\bar{x})| < \epsilon_2$ . Если оно выполняется, то положить  $x^* = \bar{x}, f(x^*) = f(\bar{x})$  и закончить поиск. Иначе — положить  $x = \bar{x}, f(x) = f(\bar{x}), j=1$  и перейти к шагу 2.

Для приближенного решения вспомогательной задачи одномерной минимизации на шаге 2 на практике удобно использовать метод поразрядного поиска. Алгоритм этого одномерного метода не требует обязательного задания границ интервала минимизации.

Алгоритм метода циклического покоординатного спуска прост, но его эффективность заметно зависит от свойств целевой функции. Метод эффективен в случае, если минимизируемая функция является *сепарабельной*, т.е. представляет собой сумму функций, каждая из которых зависит лишь от одной координаты

$$f(x) = \sum_{j=1}^n h_j(x_j), \quad x \in E_n.$$

В этом случае решение задачи минимизации можно получить за один шаг поиска. Действительно, так как

$$f(x^*) = \min_{x \in E_n} f(x) = \sum_{j=1}^n \min_{x_j \in E} h_j(x_j) = \sum_{j=1}^n h_j(x_j^*),$$

то достаточно на первом шаге поиска последовательно решить  $n$  задач одномерной минимизации функций  $h_j(x_j), j=1, \dots, n$ , что позволит найти все  $n$  координат  $x_j^*$  искомой точки  $x^*$ . Для функций общего вида эффективность поиска минимума методом циклического покоординатного спуска может быть значительно ниже.

**Пример 6.2.** Решить задачу  $f(x) = x_1^2 + x_2^2 \rightarrow \min, x \in E_2$  методом циклического покоординатного спуска.

Эта функция сепарабельна, ее линии уровня — окружности с центром в начале координат (рис 6.8). Выберем произвольную начальную точку  $x$ , например,  $x = (3, 3)^T$ . Очевидно, два шага исчерпывающего спуска сначала по направлению  $e^1$ , затем —  $e^2$  приведут в точку минимума  $x^* = (0, 0)^T$ .

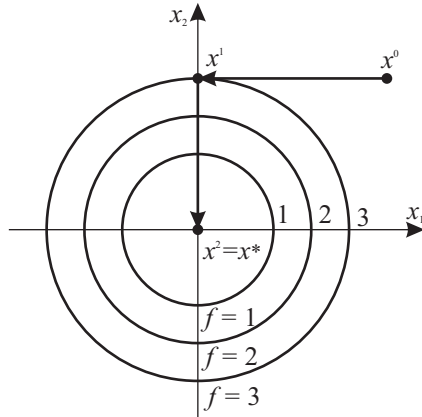


Рис. 6.8. Иллюстрация к примеру 6.2

В примере 6.2 точку минимума функции удалось найти точно за конечное число шагов. Это скорее исключение, чем правило.

**Пример 6.3.** Найти  $f(x) = 5x_1^2 + 5x_2^2 + 8x_1x_2 \rightarrow \min, x \in E_2$  методом циклического покоординатного спуска.

Заметим, что поворот системы координат на угол  $-45^\circ$  (замена переменных  $x_1 = (y_1 + y_2) / \sqrt{2}$  и  $x_2 = (-y_1 + y_2) / \sqrt{2}$ ) приводят функцию к виду  $f_1(y) = y_1^2 + 9y_2^2$ . Линии уровня этой целевой функции — эллипсы  $y_1^2 / 9 + y_2^2 = c^2$ , главные оси которых направлены вдоль осей  $(y_1, y_2)$  (рис. 6.9). Результаты решения представлены в табл. 6.2. (точность решения одномерной задачи минимизации выбрана  $10^{-2}$ ).

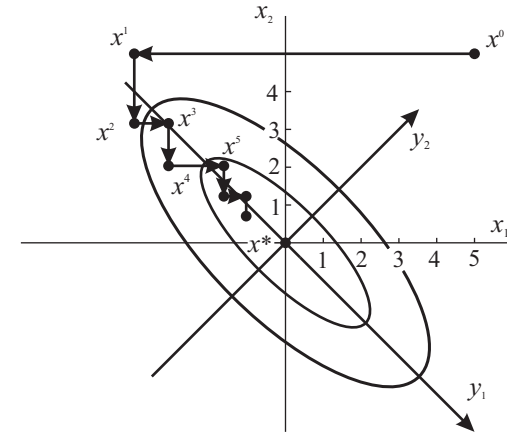


Рис. 6.9. Иллюстрация к примеру 6.3

Таблица 6.2.

Результаты расчетов в примере 6.3

| Номер итерации | $x_1$ | $x_2$ | $f(x)$ |
|----------------|-------|-------|--------|
| 0              | 5     | 5     | 450    |
| 1              | -4    | 5     | 45     |
| 2              | -4    | 3,2   | 28,8   |
| 3              | -2,56 | 3,2   | 18,43  |
| 4              | -2,56 | 2,05  | 11,8   |
| 5              | -1,64 | 2,05  | 7,55   |
| 6              | -1,64 | 1,31  | 4,83   |
| 7              | -1,05 | 1,31  | 3,09   |
| 8              | -1,05 | 0,84  | 1,98   |
| 9              | -0,67 | 0,84  | 1,27   |
| 10             | -0,67 | 0,54  | 0,81   |

Из табл. 6.2 и рис. 6.9 видно, что минимизирующая последовательность  $\{x^k\}$  сходится к точке минимума  $x^* = (0,0)^T$ . Однако, в отличие от решения задачи из примера 6.2, достижение точки минимума за конечное число шагов не гарантируется. Траектория поиска точки минимума в данной задаче имеет ярко выраженный зигзагообразный характер.

Обратим внимание, что в примере 6.3 эффективность решения задачи методом циклического покоординатного спуска можно повысить. Для этого следует дополнить его алгоритм периодически повторяющимся поиском точки минимума в направлении  $\mathbf{p}^i = \mathbf{x}^i - \mathbf{x}^{i-2}$  из точек  $\mathbf{x}^i$ . Так, например, если из точки  $\mathbf{x}^4$  провести исчерпывающий спуск в направлении  $\mathbf{p}^4 = \mathbf{x}^4 - \mathbf{x}^2$  (координаты точек приведены в табл. 6.2), то получим точку  $(-2, 2 \cdot 10^{-5}, 5, 6 \cdot 10^{-3})^T$ , расположенную значительно ближе к точке минимума  $\mathbf{x}^* = (0, 0)^T$ , чем точки  $\mathbf{x}^5, \mathbf{x}^6, \dots, \mathbf{x}^{10}$ .

Такой подход, состоящий в последовательном нахождении направлений убывания целевой функции и минимизации ее по этим направлениям, лежит в основе ряда алгоритмов. Рассмотрим один из них.

### 6.5. Метод Хука–Дживса

Эффективность прямого поиска точки минимума ограниченной снизу целевой функции можно повысить, если на каждом  $k$ -м шаге поиска последовательно выбирать направление спуска. Для этого на каждом  $k$ -м шаге выделяют предварительный этап *исследующего поиска*. Целью этого этапа является выбор направления спуска путем исследования поведения целевой функции  $f(\mathbf{x})$  в окрестности точки  $\mathbf{x}^{k-1}$ , найденной на предыдущем шаге. В результате исследующего поиска находится точка  $\bar{\mathbf{x}}^k$ , для которой  $f(\bar{\mathbf{x}}^k) < f(\mathbf{x}^{k-1})$ . Направление спуска, завершающего  $k$ -й шаг поиска, определяется вектором  $\bar{\mathbf{x}}^k - \mathbf{x}^{k-1}$ . Такая стратегия поиска, предложенная в 1961г., получила название *метода Хука–Дживса*. Это один из наиболее эффективных прямых методов.

Алгоритм метода Хука–Дживса на каждом шаге содержит две основные процедуры:

- а) исследующий поиск в окрестности данной точки  $\mathbf{x}$  для определения направления убывания функции  $f(\mathbf{x})$ . В результате получим точку  $\bar{\mathbf{x}}$ ;
- б) перемещение в направлении убывания  $(\bar{\mathbf{x}} - \mathbf{x})$ .

Поиск завершается, если после шага «а» получено, что  $\bar{\mathbf{x}} = \mathbf{x}$ .

Опишем один из возможных алгоритмов исследующего поиска — покоординатный поиск. Пусть задана точка  $\mathbf{x}$  с приращениями по каждой координате  $\Delta_j$ ,  $j = 1, \dots, n$ .

Шаг 1. Положить  $\bar{\mathbf{x}} = \mathbf{x}$ ,  $j = 1$ . Перейти к шагу 2.

Шаг 2. Сделать пробный шаг  $\mathbf{y} = \bar{\mathbf{x}} - \Delta_j \mathbf{e}^j$ , где  $\mathbf{e}^j$  —  $j$ -й базисный вектор. Если  $f(\bar{\mathbf{x}}) \leq f(\mathbf{y})$ , то перейти к шагу 3, иначе — к шагу 4.

Шаг 3. Сделать пробный шаг  $\mathbf{y} = \bar{\mathbf{x}} + \Delta_j \mathbf{e}^j$ . Если  $f(\bar{\mathbf{x}}) \leq f(\mathbf{y})$ , то перейти к шагу 5, иначе — к шагу 4.

Шаг 4. Положить  $\bar{\mathbf{x}} = \mathbf{y}$ . Перейти к шагу 5.

Шаг 5. Положить  $j = j + 1$ . Если  $j \leq n$ , то перейти к шагу 2, иначе исследующий поиск окончен, т.е. получена точка  $\bar{\mathbf{x}}$ , для которой  $f(\bar{\mathbf{x}}) < f(\mathbf{x})$ , если  $\bar{\mathbf{x}} \neq \mathbf{x}$ .

В результате исследующего поиска может оказаться, что  $\bar{\mathbf{x}} = \mathbf{x}$ . Тогда исследующий поиск считается неудачным. Если при этом норма приращения  $\Delta = (\Delta_1, \dots, \Delta_n)$  мала, т.е.  $\|\Delta\| < \varepsilon$ , где  $\varepsilon$  — заданная точность, то полагают  $\mathbf{x}^* = \mathbf{x}$ ,  $f(\mathbf{x}^*) = f(\mathbf{x})$ . Если заданная точность не достигнута, то полагают  $\Delta = \Delta / \gamma$  (где коэффициент дробления шага  $\gamma > 1$ ) и повторяют исследующий поиск.

Полный алгоритм *метода Хука – Дживса* следующий.

Шаг 1. Выбрать начальную точку  $\mathbf{x}^0$ , вектор приращений  $\Delta = (\Delta_1, \dots, \Delta_n)$ , коэффициент дробления шага  $\gamma > 1$ , параметр окончания поиска  $\varepsilon > 0$ . Перейти к шагу 2.

Шаг 2. Провести исследующий покоординатный поиск из точки  $\mathbf{x}^0$ , т.е. найти точку  $\bar{\mathbf{x}}^0$ . Если  $\bar{\mathbf{x}}^0 \neq \mathbf{x}^0$ , то перейти к шагу 4, иначе к шагу 3.

Шаг 3. Проверка окончания поиска. Если  $\|\Delta\| < \varepsilon$ , то прекратить поиск и положить  $\mathbf{x}^* = \mathbf{x}^0$ . Иначе — положить  $\Delta = \Delta / \gamma$  и перейти к шагу 2.

Шаг 4. Осуществить перемещение из точки  $\bar{\mathbf{x}}^0$  в направлении убывания  $(\bar{\mathbf{x}}^0 - \mathbf{x}^0)$  в точку  $\mathbf{x}^1$ .

$$x^1 = x^0 + a_k (\bar{x}^0 - x^0),$$

и подобрать так называемый *ускоряющий множитель*  $a_k > 0$ , чтобы  $f(x^1) < f(\bar{x}^0)$ .

Часто  $a_k \in (0, 1]$ . С увеличением  $a_k$  увеличивается длина  $a_k |\bar{x}^0 - x^0|$  шага спуска в направлении вектора  $\bar{x}^0 - x^0$ . Значение  $a_k$  можно подбирать из условия минимума функции  $f(x)$  при смещении точки  $\bar{x}^0$  в направлении этого вектора. В простейшем варианте метода Хука–Дживса значение  $a_k$  выбирают постоянным, обычно  $a_k = 2$ . В этом случае формула, по которой осуществляется спуск, имеет вид

$$x^1 = x^0 + 2(\bar{x}^0 - x^0) = 2\bar{x}^0 - x^0.$$

## 6.6. Методы случайного поиска

Основой для этих методов служит итерационный процесс

$$x^{k+1} = x^k + \alpha_k \frac{\xi}{\|\xi\|}, \quad k = 0, 1, \dots \quad (6.8)$$

где  $\alpha_k > 0$  — величина шага,  $\xi = (\xi_1, \dots, \xi_n)$  — некоторая реализация  $n$ -мерного случайного вектора  $\xi$ .

Будем считать, что координаты вектора  $\xi$  — это независимые случайные величины, равномерно распределенные на отрезке  $[-1, 1]$ . На текущей итерации при помощи генерирования случайных векторов  $\xi$  получают точки, лежащие на гиперсфере радиуса  $\alpha_k$  с центром в точке  $x^k$ . Если значение функции в полученной точке не меньше, чем в центре сферы, шаг считается неудачным. Если число неудачных шагов из данной точки достигает заданного числа  $M$ , поиск повторяется из той же точки с уменьшенным шагом до тех пор, пока шаг не станет меньше заданной точности. Если же значение функции в полученной точке меньше, чем в центре, шаг считается удачным и полученное значение выбирается за новый центр поиска.

Приведем два возможных алгоритма метода случайного поиска. Они могут использоваться как самостоятельные минимизирующие процедуры, или входить в состав других алгоритмов, например, использоваться для исследующего поиска в методе Хука–Дживса.

Алгоритм поиска с возвратом при неудачном шаге следующий.

Шаг 1. Выбрать параметр точности  $\varepsilon > 0$ , начальный шаг  $\alpha > 0$ , коэффициент уменьшения шага  $\gamma > 1$ , предельное число попыток  $M$ , точку  $x$ . Вычислить  $f(x)$ . Перейти к шагу 2.

Шаг 2. Положить счетчик числа неудачных попыток  $j = 1$ . Перейти к шагу 3.

Шаг 3. Получить реализацию случайного вектора  $\xi$ . Перейти к шагу 4.

Шаг 4. Найти пробную точку  $y = x + \alpha \frac{\xi}{\|\xi\|}$ , вычислить  $f(y)$ .

Перейти к шагу 5.

Шаг 5. Если  $f(y) < f(x)$ , то положить  $x = y$ ,  $f(x) = f(y)$  и перейти к шагу 4. Иначе — перейти к шагу 6.

Шаг 6. Положить  $j = j + 1$ . Если  $j \leq M$ , то перейти к шагу 3, иначе — к шагу 7.

Шаг 7. Если  $\alpha < \varepsilon$ , то поиск завершить, полагая  $x^* = x$ ,  $f^* = f(x)$ . Иначе — положить  $\alpha = \alpha / \gamma$  и перейти к шагу 2.

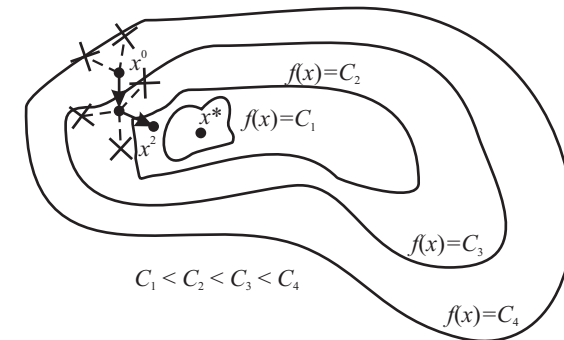


Рис. 6.10. Иллюстрация работы алгоритма с возвратом при неудачном шаге в  $E_2$

Иллюстрация построения последовательности (6.8) с помощью описанного алгоритма для функции двух переменных приведена на рис. 6.10, где пунктиром показаны неудачные попытки определения точки  $x^{k+1}$  из (6.8), не приводящие к уменьшению  $f(x)$ .

На практике предельное число неудачных попыток  $M$  обычно полагают равным  $3n$ , где  $n$  — число переменных целевой функции.

Алгоритм метода случайного поиска наилучшей пробы следующий.

Этот алгоритм отличается от предыдущего только шагами 3 и 4.

Шаг 3. Получить  $m$  реализаций случайного вектора  $\xi$ :  $\xi^1, \dots, \xi^m$ .

Шаг 4. Найти пробные точки  $y^i = x + \alpha \frac{\xi^i}{\|\xi^i\|}$ ,  $i = 1, \dots, m$ ,

вычислить  $f(y^i)$ . Найти  $y^k$  из условия  $f(y^k) = \min_i f(y^i)$  и положить  $y = y^k$ .

### Вопросы и задания для самопроверки

1. Сформулировать стратегию построения алгоритма симплексного поиска.
2. Какая нумерация вершин симплекса называется правильной?
3. Описать алгоритм отражения вершины в методе правильного симплекса.
4. Зачем необходима и в чем заключается редукция правильного симплекса?
5. Сформулировать теоретическое обоснование минимизации целевой функции методом правильного симплекса.
6. В задачах минимизации с какими целевыми функциями метод правильного симплекса не может обеспечить высокой точности?
7. Сформулировать особенности минимизации целевой функции методом Нелдера-Мида по сравнению с ее минимизацией методом правильного симплекса.

8. Назвать класс целевых функций, при минимизации которых метод Нелдера-Мида имеет преимущество перед минимизацией по регулярному симплексу.

9. Сформулировать теоретическое обоснование минимизации целевой функции методом Нелдера-Мида.

10. Назвать класс унимодальных целевых функций, для которых эффективна минимизация методом циклического покоординатного спуска.

11. Как можно дополнительно повысить эффективность поиска точки минимума целевой функции, которая ищется методом циклического покоординатного спуска?

12. В чем состоит стратегия метода Хука-Дживса?

13. Какие подходы для реализации исследующего поиска в методе Хука-Дживса Вы знаете? В чем состоит метод исследующего покоординатного поиска?

14. Перечислите способы выбора ускоряющего множителя в методе Хука-Дживса при перемещении в направлении убывания.

15. Какие алгоритмы случайного поиска Вы знаете?

16. От какого параметра в наибольшей степени зависит эффективность алгоритмов случайного поиска?

17. На основе собственного опыта дать сравнительный анализ прямых методов.

### Задание для численной реализации в среде программирования MATLAB

1. Реализовать в среде MATLAB методы правильного симплекса, Нелдера-Мида, циклического покоординатного спуска, метод Хука-Дживса и метод случайного поиска.

2. Протестировать работу реализованных методов на примере «овражной» функции

$$f(x) = x_1^2 + a x_2^2,$$

при  $a = 1, 100, 500, 1000$ . Графически отобразить линии уровня функции. Сравнить скорость работы методов при различных значениях параметра  $a$  по числу итераций и по числу вызовов функций. Параметр точности выбрать  $\epsilon = 10^{-3}$  и  $\epsilon = 10^{-5}$ .

3. Выбрать для выполнения работы тестовую функцию, номер которой соответствует номеру Вашего компьютера. Например,

для компьютера №3 это будет функция 3), для компьютера №13 – функция 4): 13-9=4; для компьютера №23 это будет функция 5): 23-9×2=5.

- 1)  $f(x) = 64x_1^2 + 126x_1x_2 + 64x_2^2 - 10x_1 + 30x_2 + 13$
- 2)  $f(x) = 129x_1^2 - 256x_1x_2 + 129x_2^2 - 51x_1 - 149x_2 - 27$
- 3)  $f(x) = 254x_1^2 + 506x_1x_2 + 254x_2^2 + 50x_1 + 130x_2 - 111$
- 4)  $f(x) = 151x_1^2 - 300x_1x_2 + 151x_2^2 + 33x_1 + 99x_2 + 48$
- 5)  $f(x) = 85x_1^2 + 168x_1x_2 + 85x_2^2 + 29x_1 - 51x_2 + 83$
- 6)  $f(x) = 211x_1^2 - 420x_1x_2 + 211x_2^2 - 192x_1 + 50x_2 - 25$
- 7)  $f(x) = 194x_1^2 + 376x_1x_2 + 194x_2^2 + 31x_1 - 229x_2 + 4$
- 8)  $f(x) = 45x_1^2 - 88x_1x_2 + 45x_2^2 + 102x_1 + 268x_2 - 21$
- 9)  $f(x) = 99x_1^2 + 196x_1x_2 + 99x_2^2 - 95x_1 - 9x_2 + 91$

4. Графически отобразить линии уровня выбранной функции. Сравнить эффективность метода покоординатного циклического спуска и метода Хука-Дживса для задачи п.2 при  $a=250$  и тестовой функции п.3 по числу итераций. Для метода Хука-Дживса величину шага исчерпывающего спуска выбирать по приближенной формуле и вычислять методом поразрядного поиска. *Объяснить полученные результаты.*

5. Используя методы прямого поиска, минимизировать функцию Розенброка

$$f(x) = 100(x_1^2 - x_2)^2 + (x_1 - 1)^2$$

с точностью  $\epsilon = 10^{-3}$  и  $\epsilon = 10^{-5}$ , контролируя точность алгоритмов исчерпывающего спуска. Начальное приближение  $x^0 = (-1, 1)^T$ . Установить, какие из примененных алгоритмов не позволяют при заданной точности поиска получить точку минимума  $x^* = (1, 1)^T$  вследствие преждевременного окончания процесса поиска.

6. На примере функции Розенброка провести сравнение прямых и градиентных методов минимизации. Определить, сколько итераций потребуется каждому методу для того, чтобы разность между численным и точным решением была меньше  $\epsilon$ .

7. На примере функции Химмельблау

$$f(x) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$$

рассмотреть особенности применения прямых методов для минимизации многомодальных функций. В качестве начального приближения взять точки  $(0, 0)$  и  $(-5, 0)$ . Как зависит работа рассматриваемых алгоритмов от выбора начального приближения?

8. Встроенная функция Matlab `fminsearch` реализует метод Нелдера-Мида. Сравнить результаты ее работы с результатами, полученными Вашим алгоритмом. В каких случаях построенный Вами алгоритм работает эффективнее функции `fminsearch`?



## Глава 7

### УСЛОВНЫЙ ЭКСТРЕМУМ ФУНКЦИИ МНОГИХ ПЕРЕМЕННЫХ. НЕОБХОДИМЫЕ И ДОСТАТОЧНЫЕ УСЛОВИЯ УСЛОВНОГО ЭКСТРЕМУМА

Пусть даны дважды непрерывно дифференцируемая целевая функция  $f(x) = f(x_1, \dots, x_n)$  и функции ограничений  $g_j(x) = g_j(x_1, \dots, x_n) = 0$ ,  $j = 1, \dots, m$ , определяющие множество допустимых решений  $U$ .

Требуется исследовать функцию  $f(x)$  на экстремум, т.е. определить точки  $x^* \in U$  ее локальных минимумов и максимумов

$$f(x^*) = \min_{x \in U} f(x), \quad f(x^*) = \max_{x \in U} f(x), \quad (7.1)$$

где  $U = \{x: g_j(x) = 0, \quad j = 1, \dots, m; \quad m < n\}$  В первой части главы будем рассматривать условные экстремумы с ограничениями только типа равенств.

#### 7.1. Условный экстремум при ограничениях типа равенств

**Определение.** Функция

$$L(x, \lambda_0, \lambda) = \lambda_0 f(x) + \sum_{j=1}^m \lambda_j g_j(x) \quad (7.2)$$

называется *обобщенной функцией Лагранжа*, числа  $\lambda_0, \lambda_1, \dots, \lambda_m$  — *множителями Лагранжа*. *Классической функцией Лагранжа* называется функция

$$L(x, \lambda) = f(x) + \sum_{j=1}^m \lambda_j g_j(x) \quad (7.3)$$

**Определение.** Градиентами обобщенной и классической функций Лагранжа по  $x$  называются вектор-столбцы, составленный из соответствующих частных производных первого порядка по  $x_i$ ,  $i = 1, \dots, n$ :

$$\nabla_x L(x, \lambda_0, \lambda) = \begin{pmatrix} \frac{\partial L(x, \lambda_0, \lambda)}{\partial x_1} \\ \dots \\ \frac{\partial L(x, \lambda_0, \lambda)}{\partial x_n} \end{pmatrix}, \quad (7.4a)$$

$$\nabla_x L(x, \lambda) = \begin{pmatrix} \frac{\partial L(x, \lambda)}{\partial x_1} \\ \dots \\ \frac{\partial L(x, \lambda)}{\partial x_n} \end{pmatrix}. \quad (7.4б)$$

**Определение.** Вторым дифференциалом обобщенной (классической) функции Лагранжа  $L(x, \lambda_0, \lambda)$  ( $L(x, \lambda)$ ) называется функция

$$d^2 L(x, \lambda_0, \lambda) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 L(x, \lambda_0, \lambda)}{\partial x_i \partial x_j} dx_i dx_j, \quad (7.5a)$$

$$d^2 L(x, \lambda) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 L(x, \lambda)}{\partial x_i \partial x_j} dx_i dx_j. \quad (7.5б)$$

**Определение.** Первым дифференциалом ограничения  $g_j(x)$  называется функция

$$dg_j(x) = \sum_{i=1}^n \frac{\partial g_j(x)}{\partial x_i} dx_i, \quad j = 1, \dots, m. \quad (7.6)$$

**Пример 7.1.** Выписать функции (7.2)–(7.6) для задачи поиска условного экстремума функции  $f(x) = x_1^2 + x_2^2$  на множестве  $U = \{x : x_2^2 - x_1 + 3 = 0\}$ , заданном ограничением  $g_1(x) = x_2^2 - x_1 + 3 = 0$ .

Обобщенная функция Лагранжа

$$L(x, \lambda_0, \lambda_1) = \lambda_0(x_1^2 + x_2^2) + \lambda_1(x_2^2 - x_1 + 3).$$

Классическая функция Лагранжа  $L(x, \lambda_1) = x_1^2 + x_2^2 + \lambda_1(x_2^2 - x_1 + 3)$ .

Градиент функций Лагранжа

$$\nabla_x L(x, \lambda_0, \lambda_1) = (2\lambda_0 x_1 - \lambda_1; 2\lambda_0 x_2 + 2\lambda_1 x_2)^T;$$

$$\nabla_x L(x, \lambda_1) = (2x_1 - \lambda_1; 2x_2 + 2\lambda_1 x_2)^T.$$

Второй дифференциал функций Лагранжа

$$d^2 L(x, \lambda_0, \lambda_1) = 2\lambda_0 dx_1^2 + (2\lambda_0 + 2\lambda_1) dx_2^2;$$

$$d^2 L(x, \lambda_1) = 2dx_1^2 + (2 + 2\lambda_1) dx_2^2.$$

Первый дифференциал ограничения  $dg_1(x) = -dx_1 + 2x_2 dx_2$ .

### Стратегия поиска локальных экстремумов

Находятся точки  $x^*$  локальных экстремумов с помощью необходимых и достаточных условий минимума и максимума первого и второго порядка при ограничениях типа равенств (порядок условий определяется порядком используемых производных). Вычисляются значения  $f(x^*)$  в найденных точках локального экстремума.

**Необходимые условия экстремума первого порядка.**

Пусть  $x^*$  — точка локального экстремума в задаче (7.1). Тогда найдутся числа  $\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*$ , не равные одновременно нулю и такие, что выполняются следующие условия:

— стационарности обобщенной функции Лагранжа по  $x$

$$\frac{\partial L(x^*, \lambda_0^*, \lambda^*)}{\partial x_i} = 0, \quad i = 1, \dots, n; \quad (7.7a)$$

— допустимости решения

$$g_j(x^*) = 0, \quad j = 1, \dots, m. \quad (7.7б)$$

Если при этом градиенты  $\nabla g_1(x^*), \dots, \nabla g_m(x^*)$  в точке  $x^*$  линейно независимы (выполняется условие регулярности), то  $\lambda_0^* \neq 0$ .

Важно отметить следующие моменты.

1. Условие (7.7a) может быть записано в векторной форме  $\nabla_x L(x^*, \lambda_0^*, \lambda^*) = 0$ .

2. Система (7.7) содержит  $(n+m)$  уравнений с  $(n+m+1)$  неизвестными  $\lambda_0^*, \lambda_1^*, \dots, \lambda_m^*, x_1^*, \dots, x_n^*$ . Точки  $x^*$ , удовлетворяющие системе при некоторых  $\lambda_0^*, \lambda^*$ , называются *условно-стационарными*.

3. При решении задач проверка регулярности затруднена, так как точка  $x^*$  заранее неизвестна. Поэтому, как правило, рассматриваются два случая:  $\lambda_0^* = 0$  и  $\lambda_0^* \neq 0$ . Если  $\lambda_0^* \neq 0$ , в системе (7.7a) полагают  $\lambda_0^* = 1$ . Это эквивалентно делению

системы уравнений (7.7a) на  $\lambda_0^*$  и замене  $\frac{\lambda_j^*}{\lambda_0^*}$  на  $\lambda_j^*$ . При том

обобщенная функция Лагранжа становится классической, а сама система (7.7) принимает вид

$$\frac{\partial L(x^*, \lambda^*)}{\partial x_i} = 0, \quad i = 1, \dots, n, \quad (7.8a)$$

$$g_j(x^*) = 0, \quad j = 1, \dots, m \quad (7.8б)$$

где число уравнений равно числу неизвестных.

**Определение.** Точка экстремума, удовлетворяющая системе (7.7) при  $\lambda_0^* \neq 0$ , называется *регулярной*, а при  $\lambda_0^* = 0$  — *нерегулярной*.

Случай  $\lambda_0^* = 0$  отражает вырожденность ограничений. При этом в обобщенной функции Лагранжа исчезает член, содержащий целевую функцию, а в необходимых условиях экстремума не используется информация, представляемая градиентом целевой функции.

**Необходимые условия экстремума второго порядка.** Пусть  $x^*$  — регулярная точка минимума (максимума) в задаче (7.1) и имеется решение  $(x^*, \lambda^*)$  системы (7.8). Тогда второй дифференциал классической функции Лагранжа, вычисленный в точке  $(x^*, \lambda^*)$ , неотрицателен (неположителен):

$$d^2 L(x^*, \lambda^*) \geq 0 \quad (d^2 L(x^*, \lambda^*) \leq 0) \quad (7.9)$$

для всех  $dx$  таких, что

$$dg_j(x^*) = \sum_{i=1}^n \frac{\partial g_j(x^*)}{\partial x_i} dx_i = 0, \quad j = 1, \dots, m. \quad (7.10)$$

**Достаточные условия экстремума.** Пусть имеется точка  $(x^*, \lambda^*)$ , удовлетворяющая системе (7.8). Если в этой точке  $d^2L(x^*, \lambda^*) > 0$ , (соответственно,  $d^2L(x^*, \lambda^*) < 0$ ) для всех ненулевых  $dx$  таких, что

$$dg_j(x^*) = \sum_{i=1}^n \frac{\partial g_j(x^*)}{\partial x_i} dx_i = 0, \quad j = 1, \dots, m,$$

то точка  $x^*$  является точкой локального минимума (максимума) задачи экстремума.

Достаточные и необходимые условия экстремума второго порядка проверяются в условно-стационарных точках, которые удовлетворяют системе (7.7) при  $\lambda_0^* \neq 0$  или системе (7.8), так как для практики представляет интерес случай, когда в функции Лагранжа присутствует целевая функция, экстремум которой ищется.

Алгоритм решения задачи следующий.

Шаг 1. Составить обобщенную функцию Лагранжа

$$L(x, \lambda_0, \lambda) = \lambda_0 f(x) + \sum_{j=1}^m \lambda_j g_j(x).$$

Шаг 2. Записать необходимые условия экстремума первого порядка

$$a) \frac{\partial L(x^*, \lambda_0^*, \lambda^*)}{\partial x_i} = 0, \quad i = 1, \dots, n; \quad б) g_j(x^*) = 0, \quad j = 1, \dots, m.$$

Шаг 3. Решить систему для двух случаев

1)  $\lambda_0^* = 0$ ;

2)  $\lambda_0^* \neq 0$  (при этом поделить условие а) на  $\lambda_0^*$  и заменить

$$\frac{\lambda_j^*}{\lambda_0^*} \text{ на } \lambda_j^*).$$

В результате найти условно-стационарные точки  $x^*$ , выдлив из них полученные при  $\lambda_0^* \neq 0$  (они могут быть регулярными точками экстремума).

Шаг 4. Для выделенных на шаге 3 точек проверить достаточные условия экстремума:

а) записать выражение для второго дифференциала классической функции Лагранжа в точке  $(x^*, \lambda^*)$

$$d^2L(x^*, \lambda^*) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 L(x^*, \lambda^*)}{\partial x_i \partial x_j} dx_i dx_j;$$

б) записать систему дополнительных условий в точке  $x^*$

$$dg_j(x^*) = \sum_{i=1}^n \frac{\partial g_j(x^*)}{\partial x_i} dx_i = 0, \quad j = 1, \dots, m;$$

в) из последних условий выразить любые из  $m$  дифференциалов  $dx_i$  через остальные  $(n-m)$  и подставить в  $d^2L(x^*, \lambda^*)$ ;

г) если  $d^2L(x^*, \lambda^*) > 0$  при ненулевых  $dx$ , то в точке  $x^*$  условный локальный минимум. Если  $d^2L(x^*, \lambda^*) < 0$  при ненулевых  $dx$ , то в точке  $x^*$  — условный локальный максимум. Если достаточные условия экстремума не выполняются, следует проверить выполнение необходимых условий второго порядка, следуя аналогичной процедуре. Если они выполняются, то требуется дополнительное исследование, в противном случае в точке  $x^*$  нет условного экстремума.

Шаг 5. Вычислить значения функции в точках условного экстремума.

Приведем в виде таблицы результаты исследования на наличие в точке  $x^*$  локальных экстремумов при соблюдении необходимых условий первого порядка

$$\frac{\partial L(x^*, \lambda_0^*, \lambda^*)}{\partial x_i} = 0, \quad i = 1, \dots, n,$$

$$g_j(x^*) = 0, \quad j = 1, \dots, m,$$

причем  $\lambda_0^* \neq 0$  (табл. 7.1).

Таблица 7.1.

Результаты исследования локальных экстремумов в точке  $x^*$  при соблюдении необходимых условий первого порядка

| № | $d^2L(x^*, \lambda^*)$ | $dg_j(x^*) \ j=1, \dots, m$ | Тип условно-стационарной точки $x^*$  |
|---|------------------------|-----------------------------|---|
| 1 | $> 0$                  | $0, dx \neq 0$              | Условный локальный минимум  |
| 2 | $< 0$                  | $0, dx \neq 0$              | Условный локальный максимум   |
| 3 | $\geq 0$               | $0$                         | Может быть условный локальный минимум, требуется дополнительное исследование  |
| 4 | $\leq 0$               | $0$                         | Может быть условный локальный максимум, требуется дополнительное исследование |
| 5 | $= 0$                  | $0$                         | Требуется дополнительное исследование   |
| 6 | $> 0, < 0$             | $0$                         | Нет экстремума  |

**Пример 7.2.** Найти  $f(x) = x_1^2 + x_2^2 \rightarrow \text{extr}$  при условии  $g_1(x) = x_1 + x_2 - 2 = 0$ .

Проверим условие регулярности. Так как  $\nabla g_1(x) = (1, 1)^T \neq 0$ , то условие выполняется. Поэтому будем пользоваться классической функцией Лагранжа.

1. Составим функцию Лагранжа

$$L(x, \lambda_1) = x_1^2 + x_2^2 + \lambda_1(x_1 + x_2 - 2).$$

2. Выпишем необходимые условия экстремума первого порядка

$$\text{а) } \frac{\partial L(x, \lambda_1)}{\partial x_1} = 2x_1 + \lambda_1 = 0, \quad \frac{\partial L(x, \lambda_1)}{\partial x_2} = 2x_2 + \lambda_1 = 0, \text{ откуда}$$

$$x_1 = -\frac{\lambda_1}{2}, \quad x_2 = -\frac{\lambda_1}{2};$$

$$\text{б) } g_1(x) = x_1 + x_2 - 2 = 0.$$

3. Решение системы:  $x_1^* = x_2^* = 1, \lambda_1^* = -2$  — условно-стационарная точка.

4. Проверим достаточные условия экстремума

$$\text{а) } d^2L(x^*, \lambda_1^*) = 2dx_1^2 + 2dx_2^2, \text{ так как } \frac{\partial^2 L(x, \lambda_1)}{\partial x_1^2} = \frac{\partial^2 L(x, \lambda_1)}{\partial x_2^2} = 2,$$

$$\frac{\partial^2 L(x, \lambda_1)}{\partial x_1 \partial x_2} = 0, \quad \frac{\partial^2 L(x, \lambda_1)}{\partial x_2 \partial x_1} = 0;$$

$$\text{б) } dg_1(x^*) = dx_1 + dx_2 = 0, \text{ так как } \frac{\partial g_1(x)}{\partial x_1} = \frac{\partial g_1(x)}{\partial x_2} = 1;$$

в) выразим дифференциал  $dx_1$  через  $dx_2$ :  $dx_1 = -dx_2$  и подставим в  $d^2L$

г) так как  $d^2L(x^*, \lambda_1^*) = 4dx_2^2 > 0$  при  $dx_2 \neq 0$ , то в точке  $x^* = (1, 1)^T$  — регулярный локальный условный минимум.

5. Вычислим значение функции в точке условного экстремума:  $f(x^*) = 2$ .

Графическое решение задачи приведено на рис.7.1.

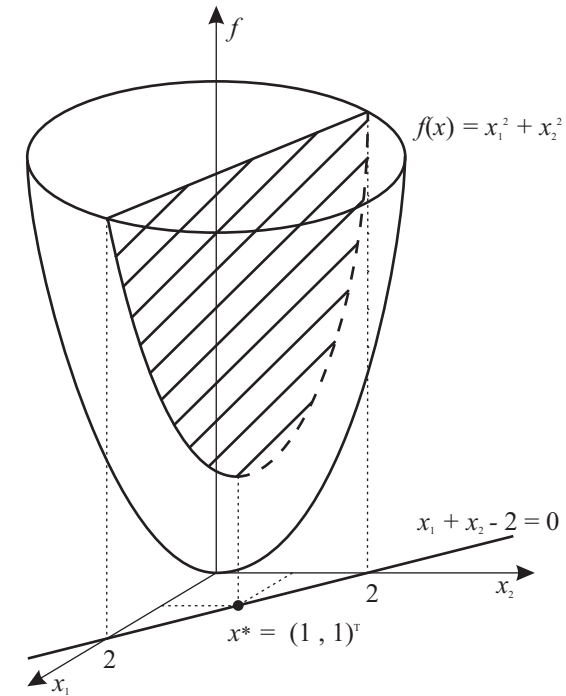


Рис.7.1. Иллюстрация к решению задачи из примера 7.2

**Пример 7.3.** Найти условный экстремум

$$f(x) = x_1 + x_2 \rightarrow \text{extr},$$

$$g_1(x) = x_1^2 + x_2^2 - 2 = 0.$$

Проверим условие регулярности. Так как  $\nabla g_1(x) = (2x_1, 2x_2)^T \neq 0$  для всех  $x \in U$ , то условие регулярности выполняется. Поэтому будем пользоваться классической функцией Лагранжа.

1. Составим функцию Лагранжа

$$L(x, \lambda_1) = x_1 + x_2 + \lambda_1(x_1^2 + x_2^2 - 2).$$

2. Выпишем необходимые условия экстремума первого порядка

$$a) \frac{\partial L(x, \lambda_1)}{\partial x_1} = 1 + 2\lambda_1 x_1 = 0, \quad \frac{\partial L(x, \lambda_1)}{\partial x_2} = 1 + 2\lambda_1 x_2 = 0, \quad \text{откуда}$$

$$x_1 = -\frac{1}{2\lambda_1}, \quad x_2 = -\frac{1}{2\lambda_1};$$

$$б) g_1(x) = x_1^2 + x_2^2 - 2 = 0.$$

3. Решением системы являются две условно-стационарные точки

$$A: x_1^* = x_2^* = 1, \quad \lambda_1^* = -\frac{1}{2} \quad B: x_1^* = x_2^* = -1, \quad \lambda_1^* = \frac{1}{2}.$$

4. Проверим достаточные условия экстремума

$$a) d^2 L(x^*, \lambda_1^*) = 2\lambda_1^* dx_1^2 + 2\lambda_1^* dx_2^2, \quad \text{так как} \quad \frac{\partial^2 L(x, \lambda_1)}{\partial x_1^2} = \frac{\partial^2 L(x, \lambda_1)}{\partial x_2^2} = 2\lambda_1^*,$$

$$\frac{\partial^2 L(x, \lambda_1)}{\partial x_1 \partial x_2} = \frac{\partial^2 L(x, \lambda_1)}{\partial x_2 \partial x_1} = 0;$$

$$б) dg_1(x^*) = 2x_1^* dx_1 + 2x_2^* dx_2 = 0, \quad \text{так как} \quad \frac{\partial g_1(x)}{\partial x_1} = 2x_1, \quad \frac{\partial g_1(x)}{\partial x_2} = 2x_2;$$

в) Исследуем точку  $A$ . Получаем  $dg_1(A) = 2dx_1 + 2dx_2 = 0$ , откуда  $dx_1 = -dx_2$ .

С учетом полученного соотношения  $d^2 L(A) = -dx_1^2 - dx_2^2 = -2dx_2^2 < 0$  при  $dx_2 \neq 0$ . Поэтому в точке  $x^* = (1, 1)^T$  имеем регулярный условный локальный максимум.

Исследуем точку  $B$ . Получаем  $dg_1(B) = -2dx_1 - 2dx_2 = 0$ , откуда  $dx_1 = -dx_2$ .

С учетом полученного соотношения  $d^2 L(B) = dx_1^2 + dx_2^2 = 2dx_2^2 > 0$  при  $dx_2 \neq 0$ .

Поэтому в точке  $x^* = (-1, -1)^T$  имеем регулярный условный локальный минимум.

5. Подсчитаем значения функции в точках экстремума  $f(A) = 2$ ,  $f(B) = -2$ .

Графическое решение задачи представлено на рис. 7.2.

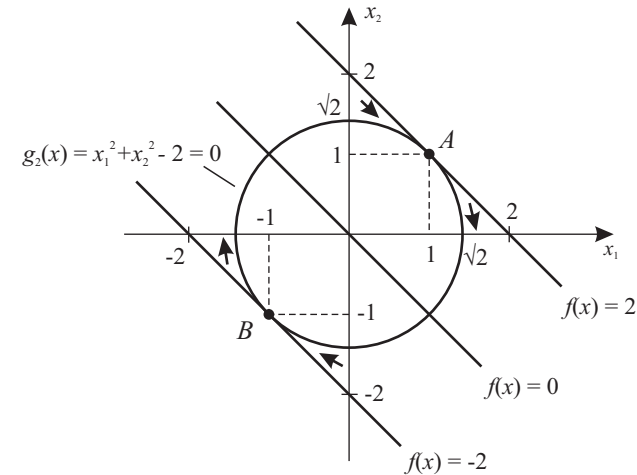


Рис. 7.2. Иллюстрация к решению задачи из примера 7.3

**Пример 7.4.** Найти условный экстремум в задаче

$$f(x) = x_1 \rightarrow \text{extr},$$

$$g_1(x) = x_2^2 - x_1^3 = 0.$$

Проверим условие регулярности. Так как  $\nabla g_1(x) = (-3x_1^2, 2x_2)^T = 0$  в точке  $x^* = (0, 0)^T$ , то условие регулярности не выполняется. Будем пользоваться алгоритмом с использованием обобщенной функции Лагранжа.

1. Составим обобщенную функцию Лагранжа

$$L(x, \lambda_0, \lambda_1) = \lambda_0 x_1 + \lambda_1(x_2^2 - x_1^3).$$

2. Выпишем необходимые условия экстремума первого порядка

$$a) \frac{\partial L(x, \lambda_0, \lambda_1)}{\partial x_1} = \lambda_0 - 3\lambda_1 x_1^2 = 0, \quad \frac{\partial L(x, \lambda_0, \lambda_1)}{\partial x_2} = 2\lambda_1 x_2 = 0;$$

$$б) g_1(x) = x_2^2 - x_1^3 = 0.$$

3. Решим систему для двух случаев.

В первом случае  $\lambda_0 = 0$ . Тогда  $\lambda_1 \neq 0$ , так как все множители Лагранжа не могут быть одновременно равными нулю. Отсюда  $x_1^* = x_2^* = 0$ ,  $\lambda_0^* = 0$ .

Во втором случае  $\lambda_0 \neq 0$ . Поделим уравнения системы, полученной на втором шаге, на  $\lambda_0$  с заменой  $\frac{\lambda_1}{\lambda_0}$  на  $\lambda_1$

$$1 - 3\lambda_1 x_1^2 = 0, \quad 2\lambda_1 x_2 = 0;$$

$$x_2^2 - x_1^3 = 0.$$

Рассмотрим второе уравнение. Если  $\lambda_1 = 0$ , то система несовместна. Если  $x_2 = 0$ , то  $x_1 = 0$  и система тоже несовместна. Таким образом, применение классической функции Лагранжа не дает результата.

4. Так как  $\lambda_0^* = 0$ , достаточные условия экстремума не проверяются. Точка  $x^*$  со значением целевой функции  $f(x^*) = 0$  является точкой нерегулярного локального и глобального минимума (рис.7.3).

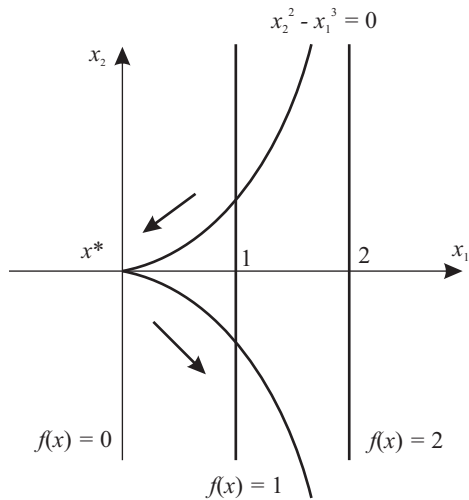


Рис.7.3. Иллюстрация к решению задачи из примера 7.4

**Пример 7.5.** Найти условный экстремум

$$f(x) = x_1^2 + x_2^2 \rightarrow \text{extr},$$

$$g_1(x) = (x_1 - 1)^2 + x_2^2 - 4 = 0.$$

Будем следовать алгоритму, не проверяя условия регулярности.

1. Составим обобщенную функцию Лагранжа

$$L(x, \lambda_0, \lambda_1) = \lambda_0(x_1^2 + x_2^2) + \lambda_1[(x_1 - 1)^2 + x_2^2 - 4].$$

2. Выпишем необходимые условия экстремума первого порядка

$$а) \frac{\partial L(x, \lambda_0, \lambda_1)}{\partial x_1} = 2\lambda_0 x_1 + 2\lambda_1(x_1 - 1) = 0, \quad \frac{\partial L(x, \lambda_0, \lambda_1)}{\partial x_2} = 2\lambda_0 x_2 + 2\lambda_1 x_2 = 0;$$

$$б) g_1(x) = (x_1 - 1)^2 + x_2^2 - 4 = 0.$$

3. Решим систему для двух случаев.

В первом случае  $\lambda_0 = 0$ . Тогда из п. 2 следует

$$2\lambda_1(x_1 - 1) = 0, \quad 2\lambda_1 x_2 = 0,$$

$$(x_1 - 1)^2 + x_2^2 - 4 = 0.$$

Так как  $\lambda_0 = 0$ , то  $\lambda_1 \neq 0$ ; поэтому из первых двух уравнений  $x_1 = 1, x_2 = 0$ . Однако при этом ограничение не выполняется:  $g_1(x) = -4 \neq 0$ , следовательно система несовместна.

Во втором случае  $\lambda_0 \neq 0$ . Поделим уравнения на  $\lambda_0$  и заменим  $\frac{\lambda_1}{\lambda_0}$  на  $\lambda_1$

$$а) \frac{\partial L(x, \lambda_1)}{\partial x_1} = 2x_1 + 2\lambda_1(x_1 - 1) = 0, \quad \frac{\partial L(x, \lambda_1)}{\partial x_2} = 2x_2 + 2\lambda_1 x_2 = 0;$$

$$б) g_1(x) = (x_1 - 1)^2 + x_2^2 - 4 = 0.$$

Рассмотрим второе уравнение. Если  $x_2 = 0$ , то из третьего уравнения следует  $x_1 = 3, x_1 = -1$ , а из первого  $\lambda_1 = -\frac{3}{2}, \lambda_1 = -\frac{1}{2}$  соответственно.

Если же  $\lambda_1 = -1$ , то первое уравнение имеет вид  $2 = 0$ , т.е. система несовместна. Таким образом, найдены две условно-стационарные точки

$$A: x_1^* = 3, x_2^* = 0, \lambda_1^* = -\frac{3}{2}; \quad B: x_1^* = -1, x_2^* = 0, \lambda_1^* = -\frac{1}{2}.$$

4. Проверим достаточные условия экстремума

$$а) d^2 L(x^*, \lambda_1^*) = 2(1 + \lambda_1^*)dx_1^2 + 2(1 + \lambda_1^*)dx_2^2;$$

$$б) dg_1(x^*) = 2(x_1^* - 1)dx_1 + 2x_2^*dx_2 = 0;$$

в, г) исследуем точку  $A$ .  $dg_1(A) = 4dx_1 + 0 = 0$ , откуда  $dx_1 = 0$  и  $d^2 L(A) = -dx_1^2 - dx_2^2 = -dx_2^2 < 0$ , если  $dx_2 \neq 0$ . Поэтому в точке  $A$  — регулярный условный максимум.

Исследуем точку  $B$ .  $dg_1(B) = 4dx_1 + 0 = 0$ , откуда  $dx_1 = 0$  и  $d^2 L(B) = dx_1^2 + dx_2^2 > 0$  при  $dx_2 \neq 0$ . Поэтому в точке  $B$  — регулярный условный минимум.

5. Подсчитаем значения функции в точках экстремума  $f(A) = 9, f(B) = 1$ .

Графическое решение задачи представлено на рис.7.4.



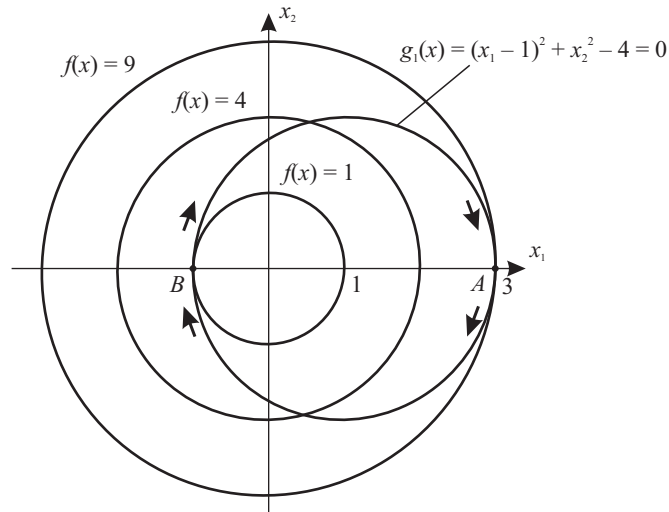


Рис. 7.4. Графическое решение задачи из примера 7.5

**Пример 7.6.** Найти условный экстремум в задаче

$$f(x) = x_1^2 - x_2^2 \rightarrow \text{extr},$$

$$g_1(x) = x_1^2 + x_2^2 - 1 = 0.$$

1. Составим обобщенную функцию Лагранжа

$$L(x, \lambda_0, \lambda_1) = \lambda_0(x_1^2 - x_2^2) + \lambda_1(x_1^2 + x_2^2 - 1).$$

2. Запишем необходимые условия экстремума первого порядка

$$\text{а) } \frac{\partial L(x, \lambda_0, \lambda_1)}{\partial x_1} = 2\lambda_0 x_1 + 2\lambda_1 x_1 = 0, \quad \frac{\partial L(x, \lambda_0, \lambda_1)}{\partial x_2} = -2\lambda_0 x_2 + 2\lambda_1 x_2 = 0;$$

$$\text{б) } g_1(x) = x_1^2 + x_2^2 - 1 = 0.$$

3. Решим систему для двух случаев.

В первом случае  $\lambda_0 = 0$ . Тогда  $\lambda_1 \neq 0$ , поэтому из первых двух уравнений следует  $x_1 = x_2 = 0$ . Однако условие «б» при этом не выполняется. Следовательно, система несовместна.

Во втором случае  $\lambda_0 \neq 0$ . Поделим уравнения на  $\lambda_0$  и заменим  $\frac{\lambda_1}{\lambda_0}$  на  $\lambda_1$

$$\text{а) } \frac{\partial L(x, \lambda_1)}{\partial x_1} = 2x_1(1 + \lambda_1) = 0, \quad \frac{\partial L(x, \lambda_1)}{\partial x_2} = 2x_2(-1 + \lambda_1) = 0;$$

$$\text{б) } g_1(x) = x_1^2 + x_2^2 - 1 = 0.$$

Пусть  $\lambda_1 = -1$ , тогда  $x_2 = 0$ ,  $x_1 = \pm 1$ . Пусть  $\lambda_1 = 1$ , тогда  $x_1 = 0$ ,  $x_2 = \pm 1$ . Других решений система не имеет. Таким образом, имеем четыре условно-стационарные точки

$$A: x_1^* = 1, x_2^* = 0, \lambda_1^* = -1; \quad B: x_1^* = -1, x_2^* = 0, \lambda_1^* = -1;$$

$$C: x_1^* = 0, x_2^* = 1, \lambda_1^* = 1; \quad D: x_1^* = 0, x_2^* = -1, \lambda_1^* = 1.$$

4. Проверим достаточные условия экстремума

$$\text{а) } d^2 L(x^*, \lambda_1^*) = 2(1 + \lambda_1^*) dx_1^2 + 2(\lambda_1^* - 1) dx_2^2;$$

$$\text{б) } dg_1(x^*) = 2x_1^* dx_1 + 2x_2^* dx_2 = 0;$$

в, г) исследуем точку  $A$ .  $dg_1(A) = 2dx_1 = 0$ , откуда получаем  $dx_1 = 0$  и  $d^2 L(A) = -4dx_2^2 < 0$ , если  $dx_2 \neq 0$ . Поэтому в точке  $A$  — регулярный локальный условный максимум.

Исследуем точку  $B$ .  $dg_1(B) = -2dx_1 = 0$ , откуда  $dx_1 = 0$  и  $d^2 L(B) = -4dx_2^2 < 0$  при  $dx_2 \neq 0$ . Поэтому в точке  $B$  — регулярный локальный условный максимум.

Исследуем точку  $C$ .  $dg_1(C) = 2dx_2 = 0$ , откуда  $dx_2 = 0$  и  $d^2 L(C) = 4dx_1^2 > 0$  при  $dx_1 \neq 0$ . Поэтому в точке  $C$  — регулярный локальный условный минимум.

Исследуем точку  $D$ .  $dg_1(D) = -2dx_2 = 0$ , откуда  $dx_2 = 0$  и  $d^2 L(D) = 4dx_1^2 > 0$  при  $dx_1 \neq 0$ . Поэтому в точке  $D$  — регулярный локальный условный минимум.

5. Вычислим значения функции в точках условного экстремума  $f(A) = f(B) = 1$ ;  $f(C) = f(D) = -1$ .

Графическое решение задачи представлено на рис. 7.5.

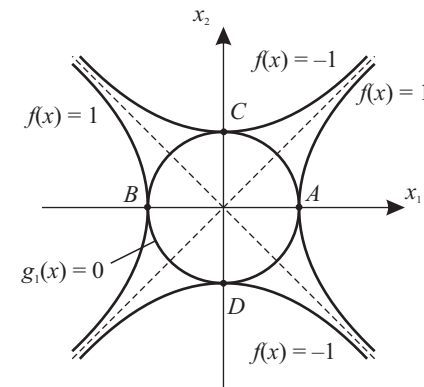


Рис. 7.5. Графическое решение задачи из примера 7.6

**Пример 7.7.** Найти условный экстремум в задаче

$$f(x) = \frac{1}{2}[(x_1 - 1)^2 + x_2^2] \rightarrow \text{extr},$$

$$g_1(x) = -x_1 + \alpha x_2^2 = 0$$

при различных  $\alpha > 0$ .

Проверим условие регулярности. Так как  $\nabla g_1(x) = (-1, 2\alpha x_2)^T \neq 0$  ни в одной точке множества  $U$ , то условие выполняется. Поэтому воспользуемся классической функцией Лагранжа.

1. Составим функцию Лагранжа

$$L(x, \lambda_1) = \frac{1}{2}[(x_1 - 1)^2 + x_2^2] + \lambda_1(-x_1 + \alpha x_2^2).$$

2. Запишем необходимые условия экстремума первого порядка

$$\text{а) } \frac{\partial L(x, \lambda_1)}{\partial x_1} = x_1 - 1 - \lambda_1 = 0, \quad \frac{\partial L(x, \lambda_1)}{\partial x_2} = x_2 + 2\alpha\lambda_1 x_2 = 0;$$

$$\text{б) } g_1(x) = -x_1 + \alpha x_2^2 = 0.$$

3. Из второго уравнения следует  $x_2(1 + 2\alpha\lambda_1) = 0$ . Если  $x_2 = 0$ , то  $x_1 = 0$ , а  $\lambda_1 = -1$ . Если  $\lambda_1 = -\frac{1}{2\alpha}$ , то  $x_1 = 1 + \lambda_1 = 1 - \frac{1}{2\alpha} = \frac{2\alpha - 1}{2\alpha}$  и при этом  $x_2 = \pm \sqrt{\frac{2\alpha - 1}{2\alpha^2}}$ .

При  $0 < \alpha < \frac{1}{2}$  решения не существует. Таким образом, имеются три условно-стационарные точки (рис.7.6)

$$A: x_1^* = 0, x_2^* = 0, \lambda_1^* = -1;$$

$$B: x_1^* = \frac{2\alpha - 1}{2\alpha}, x_2^* = \sqrt{\frac{2\alpha - 1}{2\alpha^2}}, \lambda_1^* = -\frac{1}{2\alpha};$$

$$C: x_1^* = \frac{2\alpha - 1}{2\alpha}, x_2^* = -\sqrt{\frac{2\alpha - 1}{2\alpha^2}}, \lambda_1^* = -\frac{1}{2\alpha}.$$

4. Проверим достаточные условия экстремума

$$\text{а) } d^2L(x^*, \lambda_1^*) = dx_1^2 + (1 + 2\alpha\lambda_1^*)dx_2^2;$$

$$\text{б) } dg_1(x^*) = -dx_1 + 2\alpha x_2^* dx_2 = 0;$$

$$\text{в) } \text{выразим из последнего пункта } dx_1: dx_1 = 2\alpha x_2^* dx_2;$$

$$\text{г) } d^2L(x^*, \lambda_1^*) = ((4\alpha x_2^*)^2 + (1 + 2\alpha\lambda_1^*))dx_2^2.$$

Исследуем точку  $A$ .  $d^2L(x^*, \lambda_1^*) = (1 - 2\alpha)dx_2^2$ .

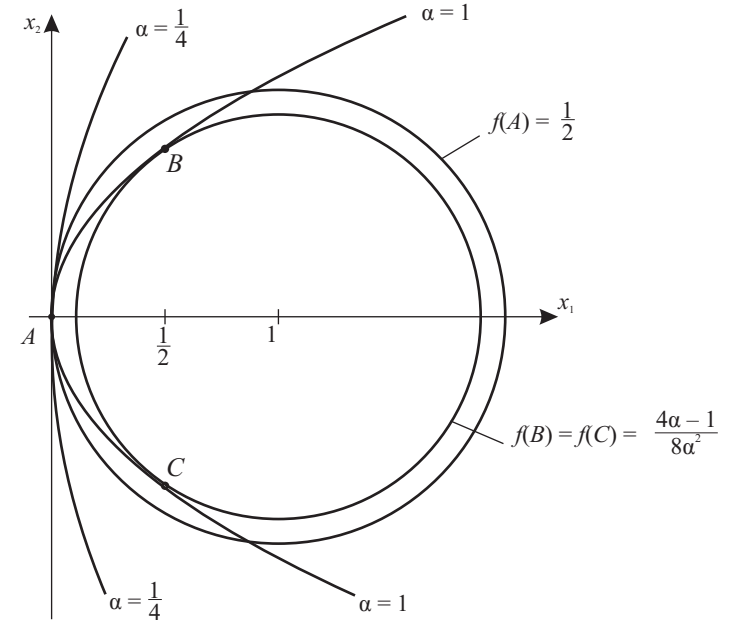


Рис.7.6. Иллюстрация к решению задачи из примера 7.7

При  $0 < \alpha < \frac{1}{2}$  получаем, что  $d^2L(x^*, \lambda_1^*) > 0$  при  $dx_2 \neq 0$  и в точке  $A$  — условный локальный минимум.

При  $\alpha > \frac{1}{2}$  находим, что  $d^2L(x^*, \lambda_1^*) < 0$  при  $dx_2 \neq 0$  и в точке  $A$  — условный локальный максимум.

При  $\alpha = \frac{1}{2}$  получаем, что  $d^2L(x^*, \lambda_1^*) = 0$  при любых  $dx_2$  и требуется дополнительное исследование.

Для этого воспользуемся *методом исключения переменных*. Выразим из уравнения ограничения одну из координат  $x_1 = \frac{1}{2}x_2^2$ .

Подставим полученное выражение в целевую функцию

$$f(x) = \frac{1}{2} \left[ \left( \frac{1}{2}x_2^2 - 1 \right)^2 + x_2^2 \right] = \frac{1}{8}x_2^4 + \frac{1}{2}x_2^2 + \frac{1}{2}.$$

Исследуем полученную функцию на безусловный экстремум.  $\frac{df}{dx_2} = \frac{1}{2}x_2^3 + x_2 = 0$ , откуда  $x_2^* = 0$  — стационарная точка. В этой точке

первая отличная от нуля производная  $f^{(4)}(x^*) = 3 > 0$  является четной и положительной. Следовательно, в точке  $x_2^*$  — локальный минимум.

С учетом связи  $x_1^* = \frac{1}{2}(x_2^*)^2 = 0$  получаем, что в точке  $x^* = (0, 0)^T$  — условный локальный минимум.

Рассмотрим точки  $B$  и  $C$ .  $d^2L(x^*, \lambda_1^*) = 16\alpha^2 \frac{2\alpha-1}{2\alpha^2} dx_2^2 = 8(2\alpha-1)dx_2^2$ .

При  $\alpha = \frac{1}{2}$  имеем  $x_1^* = x_2^* = 0$  и  $d^2L(x^*, \lambda_1^*) = 0$  при всех  $dx_2$ . Значит, требуется дополнительное исследование, которое уже было выполнено.

При  $\alpha > \frac{1}{2}$  имеем  $d^2L(x^*, \lambda_1^*) > 0$  при  $dx_2 \neq 0$  и в точках  $B$  и  $C$  условный локальный минимум.

5. Значение функции в точках экстремума

$$f(A) = \frac{1}{2}, \quad f(B) = f(C) = \frac{4\alpha-1}{8\alpha^2}.$$

Иллюстрация к решению задачи представлена на рис. 7.6.

## 7.2. Условный экстремум при ограничениях типа неравенств

Пусть требуется исследовать функцию  $f(x)$  на экстремум, т.е. определить точки  $x^* \in U$  ее локальных минимумов и максимумов

$$f(x^*) = \min_{x \in U} f(x), \quad f(x^*) = \max_{x \in U} f(x), \quad (7.11)$$

где  $U = \{x: g_j(x) \leq 0, \quad j=1, \dots, m\}$ . Рассмотрим условный экстремум с ограничениями типа неравенств.

### Стратегия решения задачи

Находятся точки  $x^*$  локального экстремума с помощью необходимых и достаточных условий минимума и максимума первого и второго порядка при ограничениях типа неравенств. Вычисляются значения  $f(x^*)$  в найденных точках локального экстремума.

**Необходимые условия минимума (максимума) первого порядка.** Пусть  $x^*$  — точка локального минимума (максимума) в задаче (7.11). Тогда найдется такое число  $\lambda_0^* \geq 0$ , и вектор

$\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)^T$ , не равные одновременно нулю и такие, что выполняются следующие условия:

— стационарности обобщенной функции Лагранжа по  $x$

$$\frac{\partial L(x^*, \lambda_0^*, \lambda^*)}{\partial x_i} = 0, \quad i=1, \dots, n; \quad (7.12a)$$

— допустимости решения

$$g_j(x^*) \leq 0, \quad j=1, \dots, m; \quad (7.12б)$$

— неотрицательности для условного минимума

$$\lambda_j^* \geq 0, \quad j=1, \dots, m \quad (7.12в)$$

(условие неположительности для условного максимума  $\lambda_j^* \leq 0, \quad j=1, \dots, m$ );

— дополняющей нежесткости

$$\lambda_j^* g_j(x^*) = 0, \quad j=1, \dots, m. \quad (7.12г)$$

Если при этом градиенты активных (т.е.  $g_j(x^*) = 0$ ) в точке  $x^*$  ограничений линейно независимы (выполняется условие регулярности), то  $\lambda_0^* \neq 0$ . Точки  $x^*$ , удовлетворяющие системе (7.12), называются *условно-стационарными*.

Необходимо обратить внимание на следующее.

1. В отличие от ограничений типа равенств необходимые условия экстремума первого порядка формулируются отдельно для максимума и минимума. Эти условия впервые были доказаны Куном и Таккером.

2. Если в задаче ограничения записаны в форме  $g_j(x^*) \geq 0$ , то их необходимо переписать как  $-g_j(x^*) \leq 0$  (7.12).

Обозначим через  $J_a$  множество индексов ограничений, активных в точке  $x^*$ .

3. При  $\lambda_0^* \neq 0$  необходимые условия минимума (максимума) первого порядка для выпуклых функций  $f(x), g_j(x), \quad j=1, \dots, m$  ( $-f(x), g_j(x), \quad j=1, \dots, m$ ) будут одновременно и достаточными условиями глобального минимума (глобального максимума).

4. Из условия дополняющей нежесткости следует, что если ограничение в точке  $x^*$  пассивное, т.е.  $g_j(x^*) < 0$ , то  $\lambda_j^* = 0$ , а если активное, т.е.  $g_j(x^*) = 0$ , то  $\lambda_j^* \geq 0$  (для минимума), и  $\lambda_j^* \leq 0$  (для максимума).

**Достаточные условия минимума (максимума) первого порядка.** Пусть имеется точка  $(x^*, \lambda^*)$ , удовлетворяющая системе (7.12) при  $\lambda_0^* \neq 0$ , число активных ограничений в точке  $x^*$  совпадает с числом  $n$  переменных (при этом условие регулярности выполняется). Если  $\lambda_j^* > 0$  для всех  $j \in J_a$ , то  $x^*$  — точка условного локального минимума. Если  $\lambda_j^* < 0$  для всех  $j \in J_a$ , то  $x^*$  — точка условного локального максимума в задаче (7.11).

**Необходимые условия минимума (максимума) второго порядка.** Пусть  $x^*$  — точка локального минимума (максимума) в задаче (7.11) и имеется решение  $(x^*, \lambda^*)$  системы (7.12). Тогда второй дифференциал классической функции Лагранжа, вычисленный в точке  $(x^*, \lambda^*)$ , неотрицателен (неположителен)

$$d^2L(x^*, \lambda^*) \geq 0 \quad (d^2L(x^*, \lambda^*) \leq 0) \quad (7.13)$$

для всех  $dx$  таких, что

$$dg_j(x^*) = \sum_{i=1}^n \frac{\partial g_j(x^*)}{\partial x_i} dx_i = 0, \quad j \in J_a, \quad \lambda_j^* > 0 \quad (\lambda_j^* < 0);$$

$$dg_j(x^*) \leq 0, \quad j \in J_a, \quad \lambda_j^* = 0.$$

**Достаточные условия минимума (максимума) второго порядка.** Пусть имеется точка  $(x^*, \lambda^*)$ , удовлетворяющая системе (7.12) при  $\lambda_0^* \neq 0$ . Если в этой точке  $d^2L(x^*, \lambda^*) > 0$  ( $d^2L(x^*, \lambda^*) < 0$ ) для всех ненулевых  $dx$  таких, что

$$dg_j(x^*) = \sum_{i=1}^n \frac{\partial g_j(x^*)}{\partial x_i} dx_i = 0, \quad j \in J_a, \quad \lambda_j^* > 0 \quad (\lambda_j^* < 0);$$

$$dg_j(x^*) \leq 0, \quad j \in J_a, \quad \lambda_j^* = 0,$$

то  $x^*$  — точка локального минимума (максимума) в задаче (7.11).

#### Алгоритм решения задачи.

Шаг 1. Составить обобщенную функцию Лагранжа

$$L(x, \lambda_0, \lambda) = \lambda_0 f(x) + \sum_{j=1}^m \lambda_j g_j(x).$$

Шаг 2. Записать необходимые условия минимума (максимума) первого порядка

$$а) \frac{\partial L(x^*, \lambda_0^*, \lambda^*)}{\partial x_i} = 0, \quad i=1, \dots, n; \quad б) g_j(x^*) \leq 0, \quad j=1, \dots, m;$$

в)  $\lambda_j^* \geq 0, \quad j=1, \dots, m$  (для минимума),  $\lambda_j^* \leq 0, \quad j=1, \dots, m$  (для максимума);

$$г) \lambda_j g_j(x^*) = 0, \quad j=1, \dots, m.$$

Шаг 3. Решить систему для двух случаев

$$1) \lambda_0^* = 0;$$

2)  $\lambda_0^* \neq 0$  (при этом поделить условия, записанные на шаге 2,

на  $\lambda_0^*$  и заменить  $\frac{\lambda_j^*}{\lambda_0^*}$  на  $\lambda_j^*$ ).

В результате найти условно-стационарные точки  $x^*$ , выдлив из них полученные при  $\lambda_0^* \neq 0$  (они могут быть регулярными точками экстремума).

В каждом из двух случаев следует начинать с рассмотрения  $2^m$  вариантов удовлетворения условия «г» дополняющей нежесткости.

Шаг 4. Для выделенных на шаге 3 точек проверить достаточные условия экстремума первого или второго порядка.

Для проверки достаточных условий первого порядка следует:

а) определить число  $l$  активных в точке  $x^*$  ограничений;

б) если  $l = n$  и  $\lambda_j^* > 0$  для всех  $j \in J_a$ , то в точке  $x^*$  — локальный минимум; если  $l = n$  и  $\lambda_j^* < 0$  для всех  $j \in J_a$ , то в точке  $x^*$  — локальный максимум. Если  $l < n$  или соответствующие множители Лагранжа не удовлетворяют достаточным условиям первого порядка, то проверить достаточные условия второго порядка.

Для проверки достаточных условий второго порядка следует:

а) записать выражение для второго дифференциала классической функции Лагранжа в точке  $(x^*, \lambda^*)$

$$d^2L(x^*, \lambda^*) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 L(x^*, \lambda^*)}{\partial x_i \partial x_j} dx_i dx_j;$$

б) записать условия, накладываемые на первые дифференциалы активных ограничений

$$dg_j(x^*) = \sum_{i=1}^n \frac{\partial g_j(x^*)}{\partial x_i} dx_i = 0, \quad j \in J_a; \quad \lambda_j^* > 0 \quad (\lambda_j^* < 0); \quad (7.14)$$

$$dg_j(x^*) = \sum_{i=1}^n \frac{\partial g_j(x^*)}{\partial x_i} dx_i \leq 0, \quad j \in J_a; \quad \lambda_j^* = 0;$$

в) исследовать знак второго дифференциала функции Лагранжа для ненулевых  $dx$ , удовлетворяющих системе (7.14). Если  $d^2L(x^*, \lambda^*) > 0$ , то в точке  $x^*$  — условный локальный минимум. Если  $d^2L(x^*, \lambda^*) < 0$ , то в точке  $x^*$  — условный локальный максимум.

Если достаточные условия первого и второго порядка не выполняются, следует проверить выполнение необходимых условий второго порядка, следуя аналогичной процедуре. Если они выполняются, то требуется дополнительное исследование, а если нет, то в точке  $x^*$  нет условного экстремума.

Шаг 5. Вычислить значения функции в точках условного экстремума.

Условия экстремума в исходной задаче (7.11) приведены в табл. 7.2. и 7.3.

Таблица 7.2.

**Необходимые и достаточные условия первого порядка в задаче поиска условного экстремума при ограничениях типа неравенств**

| Необходимые условия первого порядка |  |                           |  | Достаточные условия первого порядка ( $\lambda_0^* \neq 0$ ) |                               |                                      |
|-------------------------------------|--|---------------------------|--|--|-------------------------------|--------------------------------------|
| №                                   | $\nabla_x L(x^*, \lambda_0^*, \lambda^*);$<br>$\lambda_j g_j(x^*),$<br>$j = 1, \dots, m$ | $g_j(x^*),$<br>$j = 1, m$ | $\lambda_0^* \geq 0,$<br>$\lambda_j^*, j = 1, m$ | Число $l$  | $\lambda_j^*,$<br>$j \in J_a$ | Тип условно-стационарной точки $x^*$ |
| 1                                   | 0  | $\leq 0$                  | $\geq 0$   | $n$  | $> 0$                         | условный локальный минимум           |
| 2                                   | 0  | $\leq 0$                  | $\leq 0$   | $n$  | $< 0$                         | условный локальный максимум          |

Таблица 7.3.

**Необходимые и достаточные условия второго порядка в задаче поиска условного экстремума при ограничениях типа неравенств**

| № | $d^2L(x^*, \lambda^*)$ | $dg_j(x^*),$<br>$j \in J_a,$<br>$\lambda_j^* > 0$ | $dg_j(x^*),$<br>$j \in J_a,$<br>$\lambda_j^* < 0$ | $dg_j(x^*),$<br>$j \in J_a,$<br>$\lambda_j^* = 0$ | Тип условно-стационарной точки $x^*$                                 |
|---|------------------------|---|---|---|--|
| 1 | $> 0$                  | $0, dx \neq 0$                                    | —   | $\leq 0$  | условный локальный минимум   |
| 2 | $< 0$                  | —   | $0, dx \neq 0$                                    | $\leq 0$  | условный локальный максимум  |
| 3 | $\geq 0$               | 0   | —   | $\leq 0$  | может быть локальный минимум, требуется дополнительное исследование  |
| 4 | $\leq 0$               | —   | 0   | $\leq 0$  | может быть локальный максимум, требуется дополнительное исследование |
| 5 | $= 0$                  | 0   | —   | $\leq 0$  | требуется дополнительное исследование                                |
| 6 | $= 0$                  | —   | 0   | $\leq 0$  | требуется дополнительное исследование                                |
| 7 | $> 0, < 0$             | 0   | —   | $\leq 0$  | нет экстремума   |
| 8 | $> 0, < 0$             | —   | 0   | $\leq 0$  | нет экстремума   |

**Пример 7.8.** Найти условный экстремум

$$f(x) = x_1^2 + x_2^2 \rightarrow \text{extr},$$

$$g_i(x) = x_1 + x_2 - 2 \leq 0.$$

1. Составим обобщенную функцию Лагранжа

$$L(x, \lambda_0, \lambda_1) = \lambda_0(x_1^2 + x_2^2) + \lambda_1(x_1 + x_2 - 2).$$

2. Выпишем необходимые условия экстремума первого порядка

$$a) \frac{\partial L(x, \lambda_0, \lambda_1)}{\partial x_1} = 2\lambda_0 x_1 + \lambda_1 = 0, \quad \frac{\partial L(x, \lambda_0, \lambda_1)}{\partial x_2} = 2\lambda_0 x_2 + \lambda_1 = 0;$$

- б)  $x_1 + x_2 - 2 \leq 0$  ;  
 в)  $\lambda_1 \geq 0$  (для минимума),  $\lambda_1 \leq 0$  (для максимума);  
 г)  $\lambda_1(x_1 + x_2 - 2) = 0$  .

3. Решим систему для двух случаев.

В первом случае  $\lambda_0 = 0$  . Тогда из условия «а» следует, что  $\lambda_1 = 0$  . Это противоречит требованию существования ненулевого вектора  $(\lambda_0, \lambda_1)^T$  .

Во втором случае  $\lambda_0 \neq 0$  . Поделим систему на  $\lambda_0$  и заменим  $\frac{\lambda_1}{\lambda_0}$  на  $\lambda_1$  .

Обобщенная функция Лагранжа при этом заменяется классической

$$\text{а) } \frac{\partial L(x, \lambda_1)}{\partial x_1} = 2x_1 + \lambda_1 = 0, \quad \frac{\partial L(x, \lambda_1)}{\partial x_2} = 2x_2 + \lambda_1 = 0;$$

- б)  $x_1 + x_2 - 2 \leq 0$  ;  
 в)  $\lambda_1 \geq 0$  (для минимума),  $\lambda_1 \leq 0$  (для максимума);  
 г)  $\lambda_1(x_1 + x_2 - 2) = 0$  .

Из условия «г» дополняющей нежесткости следует

1)  $\lambda_1 = 0$  (фактически решается задача поиска безусловного экстремума).

Тогда  $x_1^* = x_2^* = 0$ ,  $\lambda_1^* = 0$  и условие «б» выполняется. Выполняются необходимые условия и для минимума, и для максимума.

2)  $\lambda_1 \neq 0$  . Тогда из системы  $x_1 + x_2 - 2 = 0$ ,  $2x_1 + \lambda_1 = 0$ ,  $2x_2 + \lambda_1 = 0$

получаем  $x_1^* = x_2^* = 1$ ,  $\lambda_1^* = -2$  . Так как  $\lambda_1^* < 0$  , то необходимое условие минимума не выполняется (в точке  $(1, 1)^T$  нет минимума), но выполняется необходимое условие максимума. Таким образом, имеем две условно-стационарные точки.

4. Проверим выполнение достаточных условий экстремума.

В точке  $x^* = (0, 0)^T$  ограничение не является активным, так как  $g_1(x) = -2 < 0$  , поэтому достаточные условия первого порядка не удовлетворяются. Проверим условия второго порядка. Так как  $d^2L(x^*, \lambda^*) = 2dx_1^2 + 2dx_2^2 > 0$  при  $dx \neq 0$  , то в точке  $x^* = (0, 0)^T$  регулярный локальный условный минимум, совпадающий в данной задаче с безусловным.

В точке  $x^* = (1, 1)^T$  ограничение является активным, но  $l = 1 < n = 2$  , поэтому достаточное условие первого порядка не выполняется. Проверим достаточное условие второго порядка.

$$d^2L(x^*, \lambda^*) = 2dx_1^2 + 2dx_2^2,$$

$$dg_1(x^*) = dx_1 + dx_2 = 0 \text{ откуда } dx_1 = -dx_2.$$

Следовательно,  $d^2L(x^*, \lambda^*) = 4dx_2^2 > 0$  при  $dx_2 \neq 0$  . Так как в этой точке  $\lambda_1^* = -2 < 0$  , то достаточное условие максимума не выполняется. Проверим достаточное условие максимума второго порядка. Так как  $d^2L(x^*, \lambda^*) = 4dx_2^2 \geq 0$  при любых  $dx_2$  , то необходимое условие максимума не выполняется, поэтому в точке  $x^* = (1, 1)^T$  максимума нет.

5. Вычислим значение функции в точке условного минимума  $f(x^*) = 0$  (рис.7.7).

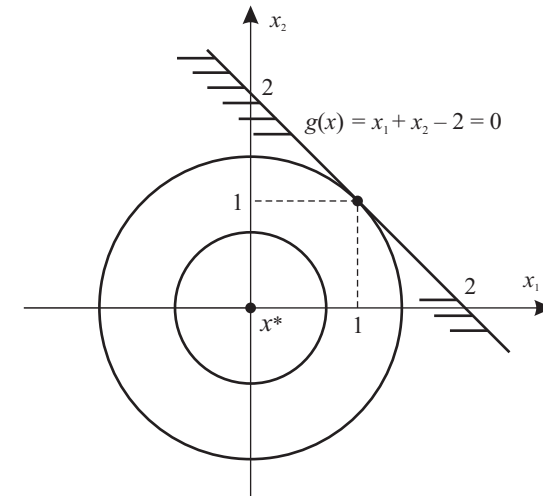


Рис.7.7. Иллюстрация к решению задачи из примера 7.8

**Пример 7.9.** Найти условный экстремум в задаче

$$f(x) = x_1 + x_2 \rightarrow \text{extr} ,$$

$$g_1(x) = x_1^2 + x_2^2 - 1 \leq 0 .$$

1. Составим обобщенную функцию Лагранжа

$$L(x, \lambda_0, \lambda_1) = \lambda_0(x_1 + x_2) + \lambda_1(x_1^2 + x_2^2 - 1) .$$



2. Выпишем необходимые условия первого порядка

$$а) \frac{\partial L(x, \lambda_0, \lambda_1)}{\partial x_1} = \lambda_0 + 2\lambda_1 x_1 = 0, \quad \frac{\partial L(x, \lambda_0, \lambda_1)}{\partial x_2} = \lambda_0 + 2\lambda_1 x_2 = 0;$$

$$б) x_1^2 + x_2^2 - 1 \leq 0;$$

$$в) \lambda_1 \geq 0 \text{ (для минимума)}, \quad \lambda_1 \leq 0 \text{ (для максимума)};$$

$$г) \lambda_1(x_1^2 + x_2^2 - 1) = 0.$$

3. Решим задачу для двух случаев.

В первом случае  $\lambda_0 = 0$ . Тогда, согласно необходимым условиям экстремума первого порядка требуется, чтобы  $\lambda_1 \neq 0$ . При этом  $x_1 = x_2 = 0$  и не удовлетворяется условие «г» дополняющей нежесткости.

Во втором случае  $\lambda_0 \neq 0$ . Поделим уравнения на  $\lambda_0$  и заменим  $\frac{\lambda_1}{\lambda_0}$  на  $\lambda_1$

$$а) \frac{\partial L(x, \lambda_1)}{\partial x_1} = 1 + 2\lambda_1 x_1 = 0 \quad \frac{\partial L(x, \lambda_1)}{\partial x_2} = 1 + 2\lambda_1 x_2 = 0;$$

$$б) x_1^2 + x_2^2 - 1 \leq 0;$$

$$в) \lambda_1 \geq 0 \text{ (для минимума)}, \quad \lambda_1 \leq 0 \text{ (для максимума)};$$

$$г) \lambda_1(x_1^2 + x_2^2 - 1) = 0.$$

Из условия «г» дополняющей нежесткости следует два варианта

1)  $\lambda_1 = 0$ . Тогда условие «а» не выполняется.

2)  $\lambda_1 \neq 0$ . Тогда  $x_1^2 + x_2^2 - 1 = 0$  и система имеет два решения (рис. 7.8)

$$\text{точка } A: x_1^* = x_2^* = -\frac{\sqrt{2}}{2}, \quad \lambda_1^* = \frac{\sqrt{2}}{2} \text{ (в точке } A \text{ может быть минимум)};$$

$$\text{точка } B: x_1^* = x_2^* = \frac{\sqrt{2}}{2}, \quad \lambda_1^* = -\frac{\sqrt{2}}{2} \text{ (в точке } B \text{ может быть максимум)}.$$

4. Проверим достаточные условия экстремума. В точках  $A$  и  $B$  ограничения являются активными, но  $l = 1 < n = 2$ . Поэтому условия первого порядка не выполняются. Проверим условия второго порядка:

$$d^2 L(x^*, \lambda_1^*) = 2\lambda_1^* dx_1^2 + 2\lambda_1^* dx_2^2,$$

$$dg_1(x^*) = 2x_1^* dx_1 + 2x_2^* dx_2 = 0.$$

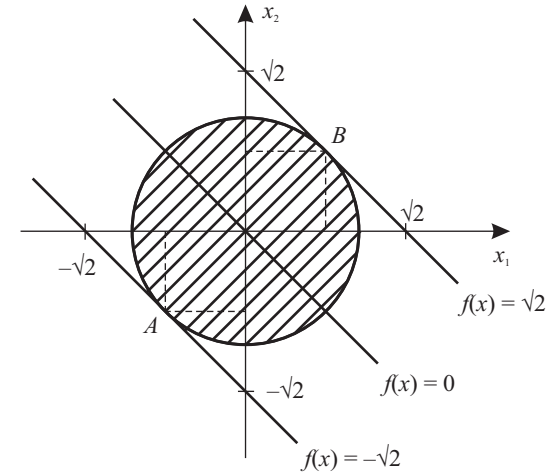


Рис. 7.8. Иллюстрация к решению задачи из примера 7.9

В точках  $A$  и  $B$  выполняется  $dx_1 = -dx_2$ . Так как  $d^2 L(A) = 4\lambda_1^* dx_2^2 = 2\sqrt{2} dx_2^2 > 0$  при  $dx_2 \neq 0$ , то в точке  $A$  — регулярный локальный условный минимум. Так как  $d^2 L(B) = 4\lambda_1^* dx_2^2 = -2\sqrt{2} dx_2^2 < 0$  при  $dx_2 \neq 0$ , то в точке  $B$  — регулярный локальный условный максимум. С другой стороны, функции  $f(x)$  и  $-f(x) = -x_1 - x_2$  — выпуклые и ограничение выпуклое, поэтому в точках  $A$  и  $B$  достигается глобальный экстремум. Достаточные условия первого и второго порядка проверялись для демонстрации методики.

5. Вычислим значение целевой функции в точках условного экстремума:  $f(A) = -\sqrt{2}$ ,  $f(B) = \sqrt{2}$ .

**Пример 7.10.** Найти условный экстремум

$$f(x) = (x_1 - 2)^2 + (x_2 - 3)^2 \rightarrow \text{extr},$$

$$g_1(x) = x_1^2 + x_2^2 - 52 \leq 0.$$

1. Составим обобщенную функцию Лагранжа

$$L(x, \lambda_0, \lambda_1) = \lambda_0((x_1 - 2)^2 + (x_2 - 3)^2) + \lambda_1(x_1^2 + x_2^2 - 52).$$

2. Выпишем необходимые условия первого порядка

$$\text{а) } \frac{\partial L(x, \lambda_0, \lambda_1)}{\partial x_1} = 2\lambda_0(x_1 - 2) + 2\lambda_1 x_1 = 0, \\ \frac{\partial L(x, \lambda_0, \lambda_1)}{\partial x_2} = 2\lambda_0(x_2 - 3) + 2\lambda_1 x_2 = 0;$$

$$\text{б) } x_1^2 + x_2^2 - 52 \leq 0;$$

$$\text{в) } \lambda_1 \geq 0 \text{ (для минимума),} \quad \lambda_1 \leq 0 \text{ (для максимума);}$$

$$\text{г) } \lambda_1(x_1^2 + x_2^2 - 52) = 0.$$

3. Решим задачу для двух случаев.

В первом случае  $\lambda_0 = 0$ . Тогда, согласно необходимым условиям экстремума первого порядка требуется, чтобы  $\lambda_1 \neq 0$ . При этом  $x_1 = x_2 = 0$  и не выполняется условие «г» дополняющей нежесткости.

Во втором случае  $\lambda_0 \neq 0$ . Поделим уравнения на  $\lambda_0$  и заменим  $\frac{\lambda_1}{\lambda_0}$  на  $\lambda_1$

$$\text{а) } \frac{\partial L(x, \lambda_1)}{\partial x_1} = 2(x_1 - 2) + 2\lambda_1 x_1 = 0 \quad \frac{\partial L(x, \lambda_1)}{\partial x_2} = 2(x_2 - 3) + 2\lambda_1 x_2 = 0;$$

$$\text{б) } x_1^2 + x_2^2 - 52 \leq 0;$$

$$\text{в) } \lambda_1 \geq 0 \text{ (для минимума),} \quad \lambda_1 \leq 0 \text{ (для максимума);}$$

$$\text{г) } \lambda_1(x_1^2 + x_2^2 - 52) = 0.$$

Из условия «г» следуют два варианта

1)  $\lambda_1 = 0$ . Тогда  $x_1 = 2, x_2 = 3$  и выполняются необходимые условия и для минимума, и для максимума. Имеем условно-стационарную точку  $A: x_1^* = 2, x_2^* = 3, \lambda_1^* = 0$ ;

2)  $\lambda_1 \neq 0$ . Тогда  $x_1^2 + x_2^2 - 52 = 0$  и система имеет решение:

$$\text{точка } B: x_1^* = 4, x_2^* = 6, \lambda_1^* = -\frac{1}{2}; \text{ точка } C: x_1^* = -4, x_2^* = -6, \lambda_1^* = -\frac{3}{2}.$$

Так как  $\lambda_1^* < 0$  в обеих точках, то в них минимума нет, но может быть максимум.

4. Проверим достаточные условия экстремума. В обеих условно-стационарных точках ограничение превращается в равенство, т.е. активно. Так как число активных ограничений  $l = 1 < 2 = n$ , то условия первого порядка не выполняются. Так как функция  $-f(x) = -(x_1 - 2)^2 - (x_2 - 3)^2$  не является выпуклой, то необходимые условия не являются достаточными. Поэтому проверим условия второго порядка

$$d^2L(x^*, \lambda_1^*) = (2 + 2\lambda_1^*)dx_1^2 + (2 + 2\lambda_1^*)dx_2^2.$$

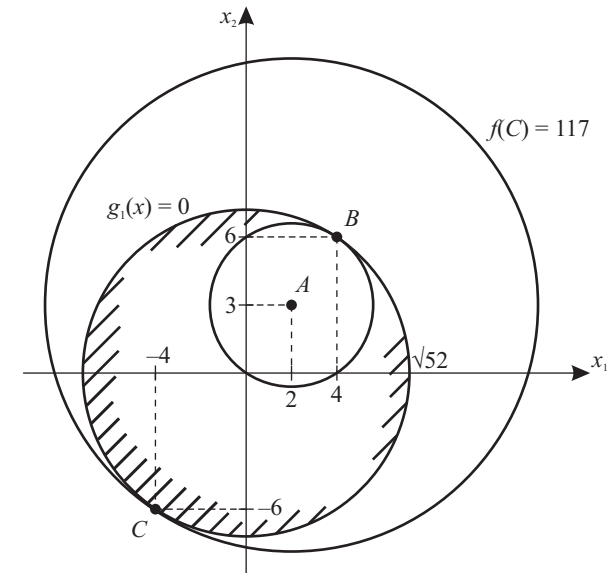


Рис.7.9. Графическое решение задачи из примера 7.10

В точке  $A$  ограничение не является активным. Так как  $\lambda_1^* = 0$ , то  $d^2L(x^*, \lambda_1^*) = 2dx_1^2 + 2dx_2^2 > 0$  при  $dx \neq 0$ . Поэтому в точке  $A$  — условный локальный минимум. Так как целевая функция выпуклая и множество допустимых значений выпукло, то можно заключить, что в данном случае необходимое условие минимума является достаточным. В точке  $A$  — глобальный минимум.

В точках  $B$  и  $C$  ограничение активно. Поэтому

$$dg_1(x^*) = 2x_1^*dx_1 + 2x_2^*dx_2 = 0. \text{ В точках } B \text{ и } C \text{ выполняется } dx_1 = -\frac{3}{2}dx_2.$$

Так как  $d^2L(B) = \left[(-\frac{3}{2})dx_2\right]^2 + dx_2^2 > 0$  при  $dx_2 \neq 0$ , а  $\lambda_1^* = -\frac{1}{2} < 0$ , то

достаточные условия максимума не выполняются. Так как  $d^2L(B) \geq 0$  при всех  $dx_2$ , то и необходимое условие максимума второго порядка в точке  $B$  не выполняется. Поэтому в ней нет экстремума. Так как

$$d^2L(C) = -\left[(-\frac{3}{2})dx_2\right]^2 - dx_2^2 < 0 \text{ при } dx_2 \neq 0, \text{ то достаточные условия}$$

максимума выполняются. В точке  $C$  условный локальный максимум.

5. Вычислим значения функций в точках экстремума  $f(A)=0$ ,  $f(C)=117$ .

Графическое решение задачи приведено на рис.7.9.

### Вопросы и задания для самопроверки

1. Что называется обобщенной (классической) функцией Лагранжа задачи условного экстремума?
2. Что называется градиентом обобщенной (классической) функции Лагранжа?
3. Что называется вторым дифференциалом обобщенной (классической) функции Лагранжа?
4. Что называется первым дифференциалом ограничения  $g_j(x)$ ?
5. Когда ограничение  $g_j(x) \leq 0$  называется активным в точке  $x^*$ ?
6. Когда ограничение  $g_j(x) \leq 0$  называется пассивным в точке  $x^*$ ?
7. При каких условиях градиенты ограничений  $g_1(x), \dots, g_m(x)$  называются линейно зависимыми в точке  $x^*$ , линейно независимыми в точке  $x^*$ ?
8. Сформулировать необходимые условия экстремума первого порядка с ограничениями типа равенств.
9. Какие точки  $x^*$  называются условно-стационарными в задаче экстремума с ограничениями типа равенств?
10. Какие точки экстремума  $x^*$  называются регулярными, а какие — нерегулярными при ограничениях типа равенств?
11. В каком случае антиградиент целевой функции является линейной комбинацией градиентов ограничений?
12. Сформулировать необходимые условия экстремума второго порядка; достаточные условия экстремума в случае ограничений типа равенств.
13. В каких точках  $x^*$  проверяются необходимые условия второго порядка и достаточные условия в случае ограничений типа равенств?
14. Сформулировать алгоритм решения задачи нахождения экстремума в случае ограничений типа равенств.
15. Сформулировать необходимые условия минимума (максимума) первого порядка с ограничениями типа неравенств.
16. В каком случае необходимые условия минимума (максимума) первого порядка с ограничениями типа неравенств яв-

ляются одновременно и достаточными условиями минимума (максимума)?

17. Сформулировать достаточные условия минимума (максимума) первого порядка в случае ограничений типа неравенств.
18. Сформулировать необходимое условие минимума (максимума) второго порядка в случае ограничений типа неравенств.
19. Сформулировать достаточные условия экстремума второго порядка в случае ограничений типа неравенств.
20. Сформулировать алгоритм решения задачи нахождения экстремума в случае ограничений типа неравенств.

## Глава 8

## ЛИНЕЙНОЕ ПРОГРАММИРОВАНИЕ

## 8.1. Определения. Примеры задач линейного программирования

*Линейное программирование* — математическая дисциплина, посвященная теории и методам решения задач об экстремумах линейных функций на множествах, задаваемых системами линейных неравенств и равенств. Линейное программирование стало развиваться в первую очередь в связи с решением задач экономики, с поиском способов оптимального распределения и использования ресурсов. Оно послужило основой широкого использования математических методов в этой сфере.

Следует подчеркнуть, что в реальных экономических задачах число независимых переменных обычно бывает очень большим (тысячи, десятки тысяч). Поэтому практическая реализация алгоритмов их решения принципиально невозможна без современной вычислительной техники.

**Пример 8.1.** Для изготовления трех видов сплавов  $A$ ,  $B$  и  $C$  используется медь, олово и цинк. Данные о сплавах приведены в табл. 8.1. В ней же указано общее количество металла каждого типа, а так же стоимость реализации одного кг сплава каждого типа.

Таблица 8.1.

Процентное содержание и общая масса компонентов

| Компоненты сплава | Содержание компонентов в % |     |     | Общая масса данной компоненты сплава, т. |
|-------------------|----------------------------|-----|-----|--|
|                   | $A$                        | $B$ | $C$ |  |
| Медь              | 20                         | 10  | 30  | 120                                      |
| Олово             | 10                         | 80  | 60  | 280                                      |
| Цинк              | 70                         | 10  | 10  | 240                                      |
| Стоимость 1кг.    | 10                         | 14  | 12  | —  |

Требуется определить, какое количество сплавов каждого вида следует изготовить предприятию, чтобы стоимость продукции была максимальной.

Предположим, что будет изготовлено  $x_1$  кг сплава  $A$ ,  $x_2$  кг сплава  $B$  и  $x_3$  кг сплава  $C$ . Тогда для производства такого количества сплавов потребуется затратить  $20x_1 + 10x_2 + 30x_3$  кг меди.

Так как общая масса меди не может превышать 120 т, то должно выполняться неравенство  $20x_1 + 10x_2 + 30x_3 \leq 120000$ .

Аналогичные рассуждения относительно возможности использования олова и цинка приведут к неравенствам

$$\begin{aligned} 10x_1 + 80x_2 + 60x_3 &\leq 280000, \\ 70x_1 + 10x_2 + 10x_3 &\leq 240000. \end{aligned}$$

При этом, так как масса сплавов не может быть отрицательной, то

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \quad (8.1)$$

Если будет изготовлено  $x_1$  кг сплава  $A$ ,  $x_2$  кг сплава  $B$  и  $x_3$  кг сплава  $C$ , то стоимость продукции составит  $F = 10x_1 + 14x_2 + 12x_3$ .

Таким образом, приходим к следующей математической задаче.

Дана система

$$\begin{cases} 20x_1 + 40x_2 + 50x_3 \leq 120000 \\ 10x_1 + 80x_2 + 60x_3 \leq 280000 \\ 70x_1 + 10x_2 + 10x_3 \leq 240000 \end{cases} \quad (8.2)$$

из четырех линейных неравенств (8.1) — (8.2) с тремя неизвестными и линейная функция относительно этих же переменных

$$F = 10x_1 + 14x_2 + 12x_3. \quad (8.3)$$

Требуется среди всех неотрицательных решений системы неравенств (8.2) найти такое, при котором функция (8.3) принимает максимальное значение.

Так как функция (8.3) линейная, а система (8.2) содержит только линейные неравенства, то задача (8.1)—(8.3) является задачей линейного программирования.

**Пример 8.2.** Продукцией молочного завода являются молоко, кефир и сметана. На производство 1 т молока, кефира и сметаны требуется соответственно 1010, 1010 и 9450 кг молока. При этом затраты рабочего

времени при разливе 1 т молока и кефира составляют 0,18 и 0,19 машино-часов. На расфасовке 1 т сметаны заняты специальные автоматы, работающие в течение 3,25 ч. Всего для производства цельномолочной продукции завод может использовать 136 т молока. Основное оборудование может быть занято в течение 21,4 машино-часа, а автоматы по расфасовке сметаны — в течение 16,25 ч. Прибыль от реализации 1 т молока, кефира и сметаны соответственно равна 30000, 26000 и 150000 руб. Завод должен ежедневно производить не менее 100 т молока, расфасованного в пакеты. Требуется определить, какую продукцию и в каком количестве следует ежедневно изготавливать заводу, чтобы прибыль от ее реализации была максимальной.

Предположим, что молочный завод будет ежедневно производить  $x_1$  тонн молока,  $x_2$  тонн кефира и  $x_3$  тонн сметаны. Тогда для производства этой продукции необходимо  $1010x_1 + 1010x_2 + 9450x_3$  кг молока.

Так как завод может использовать ежедневно не более 136000 кг молока, то должно выполняться неравенство

$$1010x_1 + 1010x_2 + 9450x_3 \leq 136000.$$

Аналогичные рассуждения, проведенные относительно возможного использования линий разлива цельномолочной продукции и автоматов по расфасовке сметаны, позволяют записать неравенства

$$\begin{aligned} 0,18x_1 + 0,19x_2 &\leq 21,4 \\ 3,25x_3 &\leq 16,25. \end{aligned}$$

Так как ежедневно должно вырабатываться не менее 100 т молока, то  $x_1 \geq 100$ .

По своему экономическому смыслу переменные  $x_2$  и  $x_3$  могут принимать лишь неотрицательные значения  $x_2 \geq 0$ ,  $x_3 \geq 0$ . Общая прибыль от реализации  $x_1$  тонн молока,  $x_2$  тонн кефира и  $x_3$  тонн сметаны равна  $30000x_1 + 22000x_2 + 150000x_3$  руб. Таким образом, приходим к следующей задаче.

Дана система

$$\begin{cases} 1010x_1 + 1010x_2 + 9450x_3 \leq 136000 \\ 0,18x_1 + 0,19x_2 \leq 21,4 \\ 3,25x_3 \leq 16,25 \\ x_1 \geq 100, \quad x_2 \geq 0, \quad x_3 \geq 0. \end{cases} \quad (8.4)$$

из четырех линейных неравенств и линейная функция относительно переменных,  $x_1$ ,  $x_2$  и  $x_3$

$$F = 30000x_1 + 22000x_2 + 150000x_3. \quad (8.5)$$

Так как система (8.4) представляет собой систему линейных неравенств и функция (8.5) линейная, то эта задача является задачей линейного программирования.

## 8.2. Общая и каноническая задачи линейного программирования

Выше были рассмотрены примеры постановки задач линейного программирования. Во всех из них требовалось найти максимум линейной функции при условии, что ее переменные принимали неотрицательные значения и удовлетворяли некоторой системе линейных уравнений или неравенств либо системе, содержащей как линейные уравнения, так и неравенства. Каждая из этих задач является частным случаем общей задачи линейного программирования.

**Определение.** *Общей (основной) задачей* линейного программирования называется задача, которая состоит в определении максимального (минимального) значения функции

$$F = \sum_{j=1}^n c_j x_j \quad (8.6)$$

при выполнении условий

$$\sum_{j=1}^n a_{ij} x_j \leq b_i \quad (i = 1, \dots, k) \quad (8.7)$$

$$\sum_{j=1}^n a_{ij} x_j = b_i \quad (i = k+1, \dots, m) \quad (8.8)$$

$$x_j \geq 0 \quad (j = 1, \dots, l; \quad l \leq n) \quad (8.9)$$

где  $a_{ij}$ ,  $b_i$  и  $c_j$  — заданные постоянные величины и  $k \leq m$ .

**Определение.** Функция (8.6) называется *целевой функцией* задачи (8.6) — (8.9), а условия (8.7) — (8.9) — *ограничениями* данной задачи.

**Определение.** Канонической задачей линейного программирования называется задача, которая состоит в определении максимального значения функции (8.6) при выполнении условий (8.8) и (8.9), где  $k=0$  и  $l=n$ .

**Определение.** Вектор  $X = (x_1, x_2, \dots, x_n)^T$ , удовлетворяющий ограничениям задачи (8.7) – (8.9), называется *допустимым решением*, или *планом*.

**Определение.** План  $X^* = (x_1^*, x_2^*, \dots, x_n^*)^T$ , при котором целевая функция принимает свое максимальное (минимальное) значение, называется *оптимальным*.

Чтобы перейти от одной формы записи задачи линейного программирования к другой, нужно в общем случае уметь, во-первых, сводить задачу минимизации функции к задаче максимизации; во-вторых, переходить от ограничений-неравенств к ограничениям-равенствам; в-третьих, заменять переменные, которые не подчинены условиям неотрицательности.

В том случае, когда требуется найти минимум функции  $F = c_1x_1 + c_2x_2 + \dots + c_nx_n$ , можно перейти к нахождению максимума функции  $F_1 = -F = -c_1x_1 - c_2x_2 - \dots - c_nx_n$  поскольку  $\min F = \max(-F)$ .

Нестрогое ограничение-неравенство общей задачи линейного программирования можно преобразовать в ограничение-равенство добавлением к его левой части дополнительной неотрицательной переменной. Таким образом, ограничение-неравенство

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \leq b_i$$

преобразуется в ограничение-равенство

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n + x_{n+1} = b_i, \quad x_{n+1} \geq 0,$$

а ограничение-неравенство

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \geq b_i$$

преобразуется в ограничение-равенство

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n - x_{n+1} = b_i, \quad x_{n+1} \geq 0.$$

Вводимые дополнительные переменные имеют вполне определенный смысл в рамках решения экономических задач. Так, если в ограничениях исходной задачи линейного программирования отражается расход и наличие производ-

ственных ресурсов, то числовое значение дополнительной переменной в плане задачи, как это будет показано в дальнейших примерах, равно объему неиспользуемого соответствующего ресурса.

Наконец, если переменная  $x_k$  по условию задачи отрицательна, то ее следует заменить двумя неотрицательными переменными  $u_k$  и  $v_k$ , приняв  $x_k = u_k - v_k$ .

**Пример 8.3.** Записать в форме канонической задачи линейного программирования задачу нахождения максимума функции  $F = 3x_1 - 2x_2 - 5x_4 + x_5$  при условиях

$$\begin{cases} 2x_1 + x_3 - x_4 + x_5 \leq 2, \\ x_1 - x_3 + 2x_4 + x_5 \leq 3, \\ 2x_1 + x_3 - x_4 + 2x_5 \leq 6, \\ x_1 + x_4 + 5x_5 \geq 8, \\ x_1, x_2, x_3, x_4, x_5 \geq 0 \end{cases}$$

Для записи в форме канонической задачи нужно перейти от ограничений-неравенств к ограничениям-равенствам. Этот переход может быть осуществлен введением четырех дополнительных неотрицательных переменных. При этом к левым частям каждого неравенства вида « $\leq$ » соответствующая переменная прибавляется, а из левых частей каждого из неравенств вида « $\geq$ » вычитается. В результате, получается следующая каноническая задача: максимизировать функцию  $F = 3x_1 - 2x_2 + 5x_3 + x_5$  при условиях

$$\begin{cases} 2x_1 + x_3 - x_4 + x_5 + x_6 = 2, \\ x_1 - x_3 + 2x_4 + x_5 + x_7 = 3, \\ 2x_1 + x_3 - x_4 + 2x_5 + x_8 = 6, \\ x_1 + x_4 + 5x_5 - x_9 = 8, \\ x_1, \dots, x_9 \geq 0. \end{cases}$$

### 8.3. Геометрическое истолкование задач линейного программирования

**Определение.** Непустое множество планов общей (основной) задачи линейного программирования называется многогранником решений.



**Свойство.** Если основная задача линейного программирования имеет оптимальный план, то максимальное значение целевая функция принимает в одной из вершин многогранника решений.

Вершину многогранника решений, в котором целевая функция принимает максимальное значение, найти не сложно, если задача содержит не более двух переменных (в общем случае задача, записанная в основной форме, содержит не более двух свободных переменных, то есть  $n - r \leq 2$ , где  $n$  — число переменных,  $r$  — ранг матрицы, составленной из коэффициентов в системе ограничений задачи).

Найдем решение задачи, состоящей в определении максимального значения функции

$$F = c_1x_1 + c_2x_2 \quad (8.10)$$

при условиях

$$a_{i1}x_1 + a_{i2}x_2 \leq b_i, \quad i = 1, \dots, k \quad (8.11)$$

$$x_1, x_2 \geq 0. \quad (8.12)$$

Каждое из неравенств (8.11), (8.12) геометрически определяет полуплоскость соответственно с граничными прямыми  $a_{i1}x_1 + a_{i2}x_2 = b_i, i = 1, \dots, k, x_1 = 0, x_2 = 0$ . Областью решений системы является множество точек, принадлежащих всем полуплоскостям, называемое в двумерном случае многоугольником решений.

Исходная задача линейного программирования состоит в нахождении вершины многоугольника решений, в которой целевая функция (8.10) принимает максимальное значение. Такая вершина существует тогда, когда многоугольник решений не пуст и на нем целевая функция ограничена сверху. Для определения этой вершины построим линию уровня  $c_1x_1 + c_2x_2 = h$ , проходящую через многоугольник решений, и будем перемещать ее в направлении вектора  $\vec{C} = (c_1, c_2)$ , ортогонального ей, до тех пор, пока она не пройдет через последнюю общую точку с многоугольником решений. Координаты этой точки и определяют оптимальный план задачи.

Отметим, что нахождение минимального значения линейной функции при данной системе ограничений отличается от нахождения ее максимального значения при тех же ограничениях лишь тем, что линия уровня  $c_1x_1 + c_2x_2 = h$  перемещается не в направлении вектора  $\vec{C} = (c_1, c_2)$ , а в противоположном направлении.

**Пример 8.4.** Для изготовления двух видов изделий  $A$  и  $B$  предприятие использует три вида сырья. Нормы расхода каждого вида сырья на изготовление единицы продукции данного вида приведены в табл. 8.2.

Таблица 8.2.

Нормы затрат и общее количество сырья в задаче 1

| Вид сырья                     | Нормы затрат сырья на одно изделие |     | Общее количество сырья |
|-------------------------------|------------------------------------|-----|------------------------|
|                               | $A$                                | $B$ |                        |
| I                             | 12                                 | 4   | 300                    |
| II                            | 4                                  | 4   | 120                    |
| III                           | 3                                  | 12  | 252                    |
| Прибыль от реализации изделия | 30                                 | 40  | —                      |

Требуется определить такой план выпуска, при котором прибыль предприятия от реализации совокупности изделий будет максимальной.

Предположим, что предприятие изготовит  $x_1$  единиц изделий вида  $A$ ,  $x_2$  единиц изделий вида  $B$ . Так как производство продукции ограничено сырьем каждого вида и количество изделий не может быть отрицательным, должны выполняться неравенства

$$\begin{cases} 12x_1 + 4x_2 \leq 300, \\ 4x_1 + 4x_2 \leq 120, \\ 3x_1 + 12x_2 \leq 252, \\ x_1, x_2 \geq 0. \end{cases}$$

Общая прибыль от реализации всей продукции составит  $F = 30x_1 + 40x_2$ .

Найдем решение данной задачи, используя ее геометрическую интерпретацию. Сначала определим многоугольник решений. Для этого в неравенствах системы ограничений и условиях неотрицательности переменных следует заменить знаки неравенств знаками точных равенств и найти соответствующие прямые.

Эти прямые изображены на рис.8.1. Каждая из прямых делит плоскость  $(x_1, x_2)$  на две полуплоскости. Координаты точек одной полуплоскости удовлетворяют исходному неравенству, а координаты другой — нет. Пересечение полученных полуплоскостей определяет многоугольник решений данной задачи.

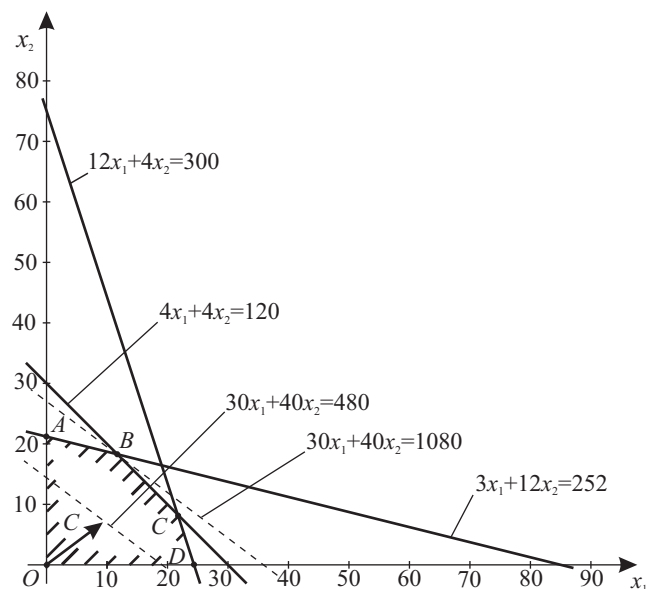


Рис. 8.1. Графическое решение примера 8.4

Из рис. 8.1 следует, что многоугольником решений является пятиугольник  $OABCD$ . Задача будет решена, если найдется точка, принадлежащая пятиугольнику  $OABCD$ , в которой функция  $F = 30x_1 + 40x_2$  принимает максимальное значение. Для определения этой точки построим вектор  $\vec{C} = (30; 40)$  и прямую  $30x_1 + 40x_2 = h$ , где  $h$  — некоторое число, при котором прямая имеет общие точки с многоугольником решений. Положим например,  $h = 480$  и построим прямую  $30x_1 + 40x_2 = 480$  (см. рис.8.1).

Координаты любой точки, принадлежащей построенной прямой и многоугольнику решений определяют допустимое решение (план производства изделий), при котором значение целевой функции равно 480. При увеличении значения  $h$  будут получаться параллельные прямые, соответствующие прибыли от реализации более 480 единиц.

Последней общей точкой прямой с многоугольником решений является точка  $B$ . Ее координаты определяют план выпуска изделий, при котором прибыль от его реализации является максимальной, а координаты точки  $B$  удовлетворяют уравнениям прямых

$$\begin{cases} 4x_1 + 4x_2 = 120, \\ 3x_1 + 12x_2 = 252. \end{cases}$$

Решив эту систему, получим  $x_1 = 12, x_2 = 18$ . Следовательно, если предприятие изготовит 12 изделий вида  $A$  и 18 вида  $B$ , то оно получит максимальную прибыль  $F = 30 \cdot 12 + 40 \cdot 18 = 1080$ .

**Пример 8.5.** Найти минимум и максимум функции  $F = x_1 + x_2$  при условиях

$$\begin{cases} 2x_1 + 4x_2 \leq 16, \\ -4x_1 + 2x_2 \leq 8, \\ x_1 + 3x_2 \geq 9, \\ x_1, x_2 \geq 0. \end{cases}$$

Построим многоугольник решений. Для этого в неравенствах системы ограничений и в условиях неотрицательности переменных следует заменить знаки неравенств знаками точных равенств. Построив полученные прямые, найдем соответствующие полуплоскости и их пересечение (рис. 8.2).

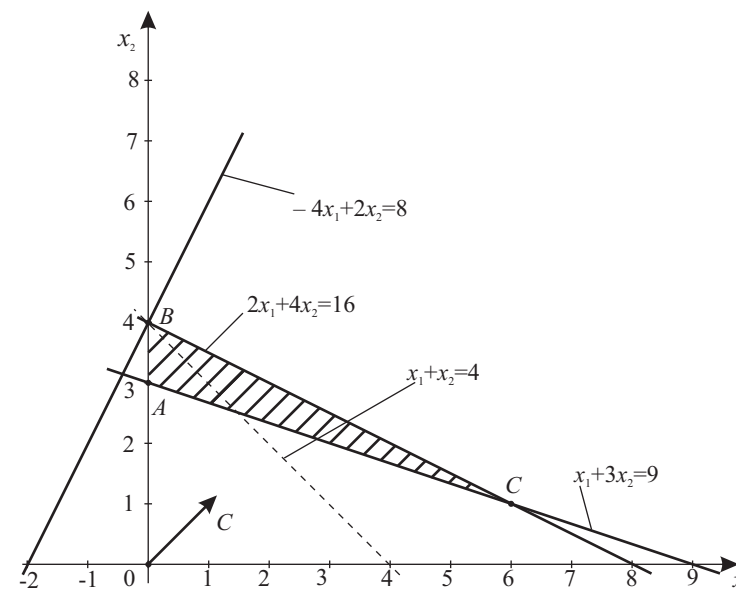


Рис. 8.2 Графическое решение примера 8.5

Из рис. 8.2 следует, что многоугольником решений является треугольник  $ABC$ . Задача будет решена, если среди точек треугольника  $ABC$  найти такие, в которых функция  $F = x_1 + x_2$  примет свои максимальное и минимальное значения. Для определения этих точек построим вектор  $\vec{C} = (1, 1)$  и, например, прямую  $x_1 + x_2 = 4$ .

Перемещая прямую  $x_1 + x_2 = 4$  параллельно самой себе в направлении вектора  $\vec{C}$ , найдем, что ее последней общей точкой с треугольником  $ABC$  является точка  $C$ , в которой функция  $F = x_1 + x_2$  принимает свое максимальное значение. Координаты точки  $C$  находим из системы

$$\begin{cases} 2x_1 + 4x_2 = 16, \\ x_1 + 3x_2 = 9, \end{cases}$$

отсюда  $x_1 = 6$ ,  $x_2 = 1$ , максимальное значение функции равно  $F = 7$ .

Для определения минимального значения целевой функции задачи перемещаем прямую  $x_1 + x_2 = 4$  в противоположном направлении. Последней общей точкой прямой и треугольника  $ABC$  является точка  $A$ . В ней целевая функция принимает минимальное значение. Координаты точки  $A$  определяются из системы

$$\begin{cases} x_1 + 3x_2 = 9, \\ x_1 = 0, \end{cases}$$

откуда  $x_1 = 0$ ,  $x_2 = 3$  и значение целевой функции равно  $F = 3$ .

#### 8.4. Аналитическое решение задач линейного программирования

Рассмотрим каноническую задачу линейного программирования: найти максимальное значение функции  $F = \sum_{j=1}^n c_j x_j$  при условиях  $\sum_{j=1}^n a_{ij} x_j = b_i$ ,  $i = 1, \dots, m$ ,  $x_j \geq 0$   $j = 1, \dots, n$ . Перепишем условия в форме

$$x_1 P_1 + x_2 P_2 + \dots + x_n P_n = P_0 \quad (8.13)$$

$$x_j \geq 0 \quad (j = 1, \dots, n) \quad (8.14)$$

где  $P_0, \dots, P_n$  —  $m$ -мерные векторы — столбцы, составленные из коэффициентов  $a_{ij}$  при неизвестных  $x_j \geq 0$ ,  $j = 1, \dots, n$  и свободных членах  $b_i$ , ( $i = 1, \dots, m$ ) системы уравнений задачи:

$$P_0 = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}; P_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix}; P_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{pmatrix}; \dots; P_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{pmatrix}.$$

**Определение.** Если система векторов  $\{P_j\}$ , входящих в левую часть уравнения (8.13), (8.14) с положительными коэффициентами  $\{x_j > 0\}$  линейно независима, то план  $X = (x_1, x_2, \dots, x_n)^T$  называется *опорным планом* канонической задачи линейного программирования.

Числа  $x_j$ ,  $j = 1, \dots, n$ , входящие в план  $X = (x_1, x_2, \dots, x_n)^T$ , называются его компонентами. Так как векторы  $P_j$  являются  $m$ -мерными, то из определения опорного плана следует, что число его ненулевых компонент не может быть больше чем  $m$ .

Отсюда следует, что вектор  $P_0$  может быть представлен в виде линейной комбинации из  $m$  векторов  $P_j$ , соответствующих положительным компонентам опорного плана. Таким образом, существует, по крайней мере, одно решение уравнения (8.13), содержащее ровно  $m$  компонент.

**Определение.** Опорный план  $X = (x_1, x_2, \dots, x_n)^T$  называется *невыврожденным*, если он содержит ровно  $m$  ненулевых компонент  $\{x_j > 0\}$ , а если их меньше  $m$ , то *вырожденным*.

Необходимо обратить внимание на следующее.

1. Если задача линейного программирования разрешима, то максимум целевой функции достигается хотя бы при одном опорном плане.

2. Опорный план является угловой точкой множества планов, и в соответствии с геометрической интерпретацией п. 8.3, целевая функция принимает максимальное значение именно на множестве опорных планов задачи (в одной из вершин многогранника решений).

Число допустимых опорных планов, как и вершин многогранника решений, не может превышать величины  $C_n^m$ . Поэтому задачу линейного программирования можно решить методом полного перебора всех возможных вариантов, сравнивая значения целевой функции в этих точках. Однако, при большой размерности задачи этот подход затруднителен. Идея рассматриваемого далее *симплекс-метода* состоит в направленном переборе опорных планов с последовательным увеличением значения целевой функции.

## Симплекс-метод

Симплекс-метод основан на направленном переходе от одного опорного плана к другому, при котором значение целевой функции возрастает. Такой переход возможен, если уже известен какой-нибудь исходный опорный план.

Нахождение решения симплекс-методом гарантирует только в том случае, когда на каждом шаге рассматриваемый опорный план является невырожденным. В противном случае задача линейного программирования называется *вырожденной*, возможны *холостые* шаги симплекс-метода, т.е. шаги, в результате которых значение целевой функции не изменяется. При этом теоретически возможно и *заикливание*, т.е. бесконечное повторение холостых шагов.

Симплекс-метод решения задачи линейного программирования не следует путать с симплекс-методом решения задачи безусловной минимизации.

Рассмотрим применение симплекс-метода на примере задачи, для которой исходный опорный план легко определяется из формы записи ограничений.

**Постановка задачи и определение исходного опорного плана.**

Пусть требуется найти максимальное значение функции

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

при условиях

$$\begin{cases} x_1 + a_{1m+1}x_{m+1} + \dots + a_{1n}x_n = b_1, \\ x_2 + a_{2m+1}x_{m+1} + \dots + a_{2n}x_n = b_2, \\ \dots \\ x_m + a_{mm+1}x_{m+1} + \dots + a_{mn}x_n = b_m, \\ x_j \geq 0 \\ i = 1, \dots, m; \quad j = 1, \dots, n; \quad m < n, \end{cases}$$

где  $a_{ij}$ ,  $b_i$  и  $c_j$  — заданные постоянные числа, причем все  $b_i > 0$ .

В векторной форме задача формируется следующим образом: найти максимум функции

$$F = \sum_{j=1}^n c_j x_j \quad (8.15)$$

при условиях

$$x_1P_1 + x_2P_2 + \dots + x_nP_n = P_0, \quad (8.16)$$

$$x_j \geq 0 \quad (j = 1, \dots, n), \quad (8.17)$$

где

$$P_0 = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}, P_1 = \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}, P_2 = \begin{pmatrix} 0 \\ 1 \\ \dots \\ 0 \end{pmatrix}, \dots, P_m = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 1 \end{pmatrix}, P_{m+1} = \begin{pmatrix} a_{1m+1} \\ a_{2m+1} \\ \dots \\ a_{mm+1} \end{pmatrix}, P_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{pmatrix}.$$

Исходным опорным планом для данной задачи является совокупность чисел  $X = (b_1, b_2, \dots, b_m, 0, \dots, 0)^T$ , которая удовлетворяет определению опорного плана, поскольку  $b_1P_1 + b_2P_2 + \dots + b_mP_m + 0 \cdot P_{m+1} + \dots + 0 \cdot P_n = P_0$ . Этот план основан на системе единичных векторов  $P_1, P_2, \dots, P_m$ , которые образуют базис  $m$ -мерного пространства.

**Проверка оптимальности опорного плана.** Выясним, является ли выбранный опорный план оптимальным для данной задачи, или необходим переход к новому опорному плану.

Представим функцию  $F = \sum_{j=1}^n c_j x_j$  в виде

$$\begin{aligned} F &= \sum_{j=1}^n c_j x_j = \sum_{i=1}^m c_i x_i + \sum_{j=m+1}^n c_j x_j = \sum_{i=1}^m c_i (b_i - \sum_{j=m+1}^n a_{ij} x_j) + \sum_{j=m+1}^n c_j x_j = \\ &= \sum_{i=1}^m c_i b_i + \sum_{j=m+1}^n (c_j - \sum_{i=1}^m c_i a_{ij}) x_j = \sum_{i=1}^m c_i b_i - \sum_{j=m+1}^n (z_j - c_j) x_j = \\ &= \sum_{i=1}^m c_i b_i - \sum_{j=m+1}^n \Delta_j x_j = F_0 - \sum_{j=m+1}^n \Delta_j x_j, \end{aligned}$$

где  $z_j = \sum_{i=1}^m c_i a_{ij}$ ,  $\Delta_j = z_j - c_j = \sum_{i=1}^m c_i a_{ij} - c_j$ . Здесь  $F_0 = \sum_{i=1}^m c_i b_i$  — значение целевой функции при выбранном опорном плане

$X = (b_1, b_2, \dots, b_m, 0, \dots, 0)^T$ . Величина  $\Delta_j = \sum_{i=1}^m c_i a_{ij} - c_j$  показы-

вает, на сколько изменится  $F_0$  при увеличении компоненты  $x_j$  ( $j = m+1, \dots, n$ ). Следует говорить только об увеличении  $x_j$ , так как одним из условий задачи является  $x_j \geq 0$  ( $j = 1, \dots, n$ ), а при данном опорном плане все  $x_j = 0$  ( $j = m+1, \dots, n$ ).

Если  $\Delta_j \geq 0$  для любого  $j$  ( $j = m+1, \dots, n$ ), то, очевидно, что увеличение компоненты  $x_j$  ( $j = m+1, \dots, n$ ) приведет к

уменьшению значения функции  $F = F_0 - \sum_{j=m+1}^n \Delta_j x_j < F_0$ . Сле-

довательно, если  $\Delta_j \geq 0$  для любого  $j$  ( $j = 1, \dots, n$ ), то данный опорный план задачи (8.15) – (8.17) является оптимальным, а

максимальное значение функции найдено и равно  $F = \sum_{i=1}^m c_i b_i$ .

**Переход к новому опорному плану.** Если среди  $\Delta_j$  ( $j = m+1, \dots, n$ ) найдутся  $\Delta_p < 0$ , то будет иметь смысл ввести в опорный план ту переменную  $x_k$ , которой соответствует значение  $\Delta_k = -\max_p |\Delta_p|$ , что приведет к максимальному увеличению целевой функции на данном шаге. При этом можно не рассчитывать значения  $\Delta_j$  при  $j = 1, \dots, m$ , так как  $P_j$  – базисные вектора, следовательно, для таких  $j$

$$\Delta_j = \sum_{i=1}^m c_j a_{ij} - c_j = c_1 \times 0 + \dots + c_j \times 1 + \dots + c_m \times 0 - c_j = 0.$$

Рассмотрим, как изменится значение целевой функции при включении в опорный план компоненты  $x_k$

$$\begin{aligned} F &= F_0 - \Delta_k x_k \equiv F_0 - \left( \sum_{i=1}^m c_i a_{ik} - c_k \right) x_k = \sum_{i=1}^m c_i b_i - \sum_{i=1}^m c_i a_{ik} x_k + c_k x_k = \\ &= \sum_{i=1}^m c_i (b_i - a_{ik} x_k) + c_k x_k. \end{aligned}$$

Если при этом все  $a_{ik} < 0, i = 1, \dots, m$ , то

$$F = \sum_{i=1}^m c_i (b_i - a_{ik} x_k) + c_k x_k \xrightarrow{x_k \rightarrow \infty} \infty,$$

т. е. функция не ограничена на множестве планов.

Если среди величин  $a_{ik}, i = 1, \dots, m$  есть такие, что  $a_{ik} > 0$ , то необходимо, чтобы выполнялось условие  $(b_i - a_{ik} x_k) \geq 0$  для всех  $a_{ik} > 0$ . Это условие будет выполняться при

$x_k = \min_{i: a_{ik} > 0} \left( \frac{b_i}{a_{ik}} \right) = \frac{b_l}{a_{lk}}$ . В результате получим, что  $(b_l - a_{lk} x_k) = 0$  и

компоненту  $x_l$  следует исключить из опорного плана, а вместо включить компоненту  $x_k$ . Базисный вектор  $P_l$  исключается, вместо него в базис включается вектор  $P_k$ . Значение функции

при этом увеличится  $F = F_0 - \Delta_k x_k = F_0 - \Delta_k \frac{b_l}{a_{lk}} > F_0$ .

Возможна ситуация, когда  $x_k = \min_{i: a_{ik} > 0} \left( \frac{b_i}{a_{ik}} \right)$  достигается при

нескольких  $i$ . В таком случае получим, что несколько компонент нового опорного плана будут равны нулю, и невырожденный опорный план превратится в вырожденный.

**Переход к новому базису.** Рассмотрим как изменятся условия (8.16) исходной задачи при переходе к новому базису, в который вместо вектора  $P_l$  включается вектор  $P_k$

$$\begin{aligned} \begin{pmatrix} 1 \\ \dots \\ 0 \\ \dots \\ 0 \end{pmatrix} x_1 + \dots + \underbrace{\begin{pmatrix} 0 \\ \dots \\ 1 \\ \dots \\ 0 \end{pmatrix}}_{P_l} x_l + \dots + \begin{pmatrix} 0 \\ \dots \\ 0 \\ \dots \\ 1 \end{pmatrix} x_m + \begin{pmatrix} a_{1m+1} \\ \dots \\ a_{lm+1} \\ \dots \\ a_{mm+1} \end{pmatrix} x_{m+1} + \dots + \\ + \underbrace{\begin{pmatrix} a_{1k} \\ \dots \\ a_{lk} \\ \dots \\ a_{mk} \end{pmatrix}}_{P_k} x_k + \dots + \begin{pmatrix} a_{1n} \\ \dots \\ a_{ln} \\ \dots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ \dots \\ b_l \\ \dots \\ b_m \end{pmatrix}. \end{aligned}$$

Поменяем местами вектор  $P_k$ , который соответствует  $\Delta_k = -\max_p |\Delta_p|$  для  $\Delta_p < 0$  и вектор базиса  $P_l$ , который соответствует компоненте, исключаемой из опорного плана.

$$\begin{pmatrix} 1 \\ \dots \\ 0 \\ \dots \\ 0 \end{pmatrix} x_1 + \dots + \begin{pmatrix} a_{1k} \\ \dots \\ a_{lk} \\ \dots \\ a_{mk} \end{pmatrix} x_k + \dots + \begin{pmatrix} 0 \\ \dots \\ 0 \\ \dots \\ 1 \end{pmatrix} x_m + \begin{pmatrix} a_{1m+1} \\ \dots \\ a_{lm+1} \\ \dots \\ a_{mm+1} \end{pmatrix} x_{m+1} + \dots + \\
+ \begin{pmatrix} 0 \\ \dots \\ 1 \\ \dots \\ 0 \end{pmatrix} x_l + \dots + \begin{pmatrix} a_{1n} \\ \dots \\ a_{ln} \\ \dots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_l \\ \dots \\ b_l \\ \dots \\ b_m \end{pmatrix}.$$

Далее необходимо привести вектор-столбец  $P_k$  к виду вектор-столбца  $P_l$ , т. е. сделать так, что бы на  $l$ -й позиции у него стояла единица, а все остальные элементы были заменены нулями. Для этого поделим  $l$ -ю строку на  $a_{lk}$  и вычтем данную строку из остальных, умножая ее на  $a_{ik}$  для  $i$ -й строки

$$\begin{pmatrix} 1 \\ \dots \\ 0 \\ \dots \\ 0 \end{pmatrix} x_1 + \dots + \begin{pmatrix} 0 \\ \dots \\ 1 \\ \dots \\ 0 \end{pmatrix} x_k + \dots + \begin{pmatrix} 0 \\ \dots \\ 0 \\ \dots \\ 1 \end{pmatrix} x_m + \begin{pmatrix} a_{1m+1} - a_{1k} \frac{a_{lm+1}}{a_{lk}} \\ \dots \\ \frac{a_{lm+1}}{a_{lk}} \\ \dots \\ a_{mm+1} - a_{mk} \frac{a_{lm+1}}{a_{lk}} \end{pmatrix} x_{m+1} + \dots + \\
+ \begin{pmatrix} 0 - a_{1k} \frac{a_{lm+1}}{a_{lk}} \\ \dots \\ \frac{1}{a_{lk}} \\ \dots \\ 0 - a_{mk} \frac{a_{lm+1}}{a_{lk}} \end{pmatrix} x_l + \dots + \begin{pmatrix} a_{1n} - a_{1k} \frac{a_{lm+1}}{a_{lk}} \\ \dots \\ \frac{a_{ln}}{a_{lk}} \\ \dots \\ a_{mn} - a_{mk} \frac{a_{lm+1}}{a_{lk}} \end{pmatrix} x_n = \begin{pmatrix} b_l - a_{1k} \frac{b_l}{a_{lk}} \\ \dots \\ \frac{b_l}{a_{lk}} \\ \dots \\ b_m - a_{mk} \frac{b_l}{a_{lk}} \end{pmatrix}.$$

Новый опорный план  $X'$  можно найти, приравняв нулю все компоненты  $x_j$  ( $j = m+1, \dots, n$ ). В результате получим

$$X' = (x_1, \dots, x_k, \dots, x_m, 0, \dots, 0)^T = \left( b_l - a_{1k} \frac{b_l}{a_{lk}}, \dots, \frac{b_l}{a_{lk}}, \dots, b_m - a_{mk} \frac{b_l}{a_{lk}}, 0, \dots, 0 \right)^T.$$

Таким образом, задаче в форме (8.15) – (8.17), соответствует новый опорный план  $X'$ , причем  $F(X') > F(X)$ . Далее следует проверить новый опорный план  $X'$  на оптимальность и т.д.

\* \* \*

Проведенный анализ задачи (8.15) – (8.17) позволяет сформулировать следующие выводы три следующих три утверждения.

**Утверждение 1.** (Признак оптимальности опорного плана). Опорный план  $X = (b_1, b_2, \dots, b_m, 0, \dots, 0)^T$  задачи (8.15) – (8.17) является оптимальным, если  $\Delta_j \geq 0$  для любого  $j = 1, \dots, n$ .

**Утверждение 2.** Если для некоторого  $\Delta_k < 0$  среди чисел  $a_{ik}$  ( $i = 1, \dots, m$ ) нет положительных, то целевая функция (8.15) не ограничена на множестве ее планов.

**Утверждение 3.** Если опорный план  $X$  задачи (8.15) – (8.17) не вырожден и  $\Delta_k < 0$  и есть  $a_{ik} > 0$  есть положительные, то существует опорный план  $X'$  такой, что  $F(X') > F(X)$ .

Утверждения 1–3 позволяют проверить найденный опорный план на оптимальность и оценить целесообразность перехода к новому опорному плану.

Для решения конкретных задач в дальнейшем будет применяться симплекс-таблица. Прием, использующий симплекс-таблицу, включает в себя запись начальных данных в виде таблицы и набор правил перехода к новому базису, которые соответствуют преобразованиям, проведенным выше.

Правила преобразования симплекс-таблицы можно применять формально, не обращаясь к проведенным выше выкладкам.

**Описание симплекс-таблицы.** Исследование опорного плана на оптимальность, а так же переход к новому опорному плану удобнее вести, если условия задачи (8.15) – (8.17) записать в виде симплекс-таблицы (табл. 8.3).



Таблица 8.3.

Общий вид симплекс-таблицы

| $i$ | Базис | $C_0$ | $P_0$ | $c_1$ | ... | $c_m$ | $c_{m+1}$      | ... | $c_k$      | ... | $c_n$      |
|-----|-------|-------|-------|-------|-----|-------|----------------|-----|------------|-----|------------|
|     |       |       |       | $P_1$ | ... | $P_m$ | $P_{m+1}$      | ... | $P_k$      | ... | $P_n$      |
| 1   | $P_1$ | $c_1$ | $b_1$ | 1     | ... | 0     | $a_{1m+1}$     | ... | $a_{1k}$   | ... | $a_{1n}$   |
| 2   | $P_2$ | $c_2$ | $b_2$ | 0     | ... | 0     | $a_{2m+1}$     | ... | $a_{2k}$   | ... | $a_{2n}$   |
| ... | ...   | ...   | ...   | ...   | ... | ...   | ...            | ... | ...        | ... | ...        |
| $l$ | $P_l$ | $c_l$ | $b_l$ | 0     | ... | 0     | $a_{lm+1}$     | ... | $a_{lk}$   | ... | $a_{ln}$   |
| ... | ...   | ...   | ...   | ...   | ... | ...   | ...            | ... | ...        | ... | ...        |
| $m$ | $P_m$ | $c_m$ | $b_m$ | 0     | ... | 1     | $a_{mm+1}$     | ... | $a_{mk}$   | ... | $a_{mn}$   |
|     |       |       | $F_0$ | 0     | ... | 0     | $\Delta_{m+1}$ | ... | $\Delta_k$ | ... | $\Delta_n$ |

В столбце  $C_0$  симплекс-таблицы записывают коэффициенты при неизвестных целевой функции, имеющие те же индексы, что и векторы данного базиса.

В столбце  $P_0$  записывают положительные компоненты исходного опорного плана, в нем же в результате вычислений получают компоненты оптимального плана. Столбцы векторов  $P_j$  представляют собой коэффициенты разложения этих векторов по векторам данного базиса.

В табл. 8.3 первые  $m$  строк определяются исходными данными, а показатели  $(m+1)$ -й строки вычисляются. На пересечении этой строки со столбцом  $P_0$  записывают значение целевой функции, которое она принимает на данном опорном плане, а в столбце вектора  $P_j$  — значение величины  $\Delta_j = z_j - c_j$ .

Величина  $z_j$  находится как скалярное произведение вектора  $P_j$  ( $j=1, \dots, n$ ) на вектор  $P_0$ . Значение целевой функции  $F_0$  равно скалярному произведению вектора  $P_0$  на вектор  $C_0$ .

**Проверка оптимальности опорного плана по симплекс-таблице.** После заполнения табл. 8.3 исходный опорный план проверяют на оптимальность. Для этого просматривают элементы  $(m+1)$ -й строки симплекс-таблицы. В результате может иметь место один из трех случаев:

- 1)  $\Delta_j \geq 0$  для любого  $j$  ( $j=1, \dots, n$ );
- 2)  $\Delta_j < 0$  для некоторого  $j$  и при этом величины  $a_{ij} \leq 0$  ( $i=1, \dots, m$ );

3)  $\Delta_j < 0$  для некоторых индексов  $j$  и для каждого такого  $j$  по крайней мере одно из чисел  $a_{ij}$  положительно.

В первом случае на основании признака оптимальности исходный опорный план является оптимальным, во втором — целевая функция не ограничена сверху на множестве планов, а в третьем — можно перейти от исходного плана к новому опорному плану, при котором значение целевой функции увеличивается. Этот переход осуществляется исключением из исходного базиса какого-нибудь из векторов и введением в базис нового вектора. В качестве вектора, вводимого в базис, можно взять любой из векторов  $P_j$ , для которого  $\Delta_j < 0$ . Пусть, например,  $\Delta_k < 0$  и решено ввести в базис вектор  $P_k$ .

Для определения вектора, подлежащего исключению из базиса, находят  $\min(b_i / a_{ik})$  для всех  $a_{ik} > 0$ . Пусть этот минимум достигается при  $i = l$ . Тогда из базиса исключают вектор  $P_l$ , а число  $a_{lk}$  называют *разрешающим элементом*. Столбец и строку, на пересечении которых находится разрешающий элемент, называют *направляющими*.

**Поиск нового опорного плана в методе симплекс-таблицы.** После выделения направляющих строки и столбца находят новый опорный план и коэффициенты разложения вектора  $P_j$  по векторам нового базиса. Это можно реализовать, если воспользоваться методом Жордана-Гаусса. При этом можно показать, что положительные компоненты нового опорного плана вычисляются по формулам

$$b'_i = \begin{cases} b_i - (b_l / a_{lk}) a_{ik} & \text{при } i \neq l, \\ b_l / a_{lk} & \text{при } i = l, \end{cases} \quad (8.18)$$

а коэффициенты разложения векторов  $P_j$  — по векторам нового базиса по формулам

$$a'_{ij} = \begin{cases} a_{ij} - (a_{lj} / a_{lk}) a_{ik} & \text{при } i \neq l, \\ a_{lj} / a_{lk} & \text{при } i = l. \end{cases} \quad (8.19)$$

Элементы  $(m+1)$ -й строки новой таблицы могут быть вычислены либо по формулам

$$F'_0 = F_0 - (b_l / a_{lk}) \Delta_k, \quad (8.20)$$

$$\Delta'_j = \Delta_j - (a_{lj} / a_{lk}) \Delta_k, \quad (8.21)$$

либо на основании их определения.

Из формулы (8.20) следует, что при переходе от одного опорного плана к другому наиболее целесообразно ввести в базис вектор  $P_j$ , при котором максимальным по абсолютной величине является число  $(b_l / a_{lk})\Delta_k$ . Однако с целью упрощения вычислительного процесса вектор, вносимый в базис, будем определять исходя из максимальной абсолютной величины чисел  $\Delta_k$ . Если таких чисел будет несколько, то в базис будем вводить вектор, имеющий тот же индекс, что и максимальное из чисел  $c_j$ .

**Правила перехода к новой симплекс-таблице.** Итак, переход от одного опорного плана к другому сводится к переходу от одной симплекс-таблицы к другой. Элементы новой симплекс-таблицы можно вычислить с помощью рекуррентных формул (8.18) — (8.21). Так же этот переход можно осуществлять по правилам, непосредственно вытекающим из формул. Эти правила состоят в следующем.

В столбцах векторов, входящих в базис, на пересечении строк и столбцов одноименных векторов проставляются единицы, а все остальные элементы полагают равными нулю.

Элементы векторов  $P_0$  и  $P_j$  в строке новой симплекс-таблицы, в которой записан вектор, вводимый в базис, получают из элементов этой же строки исходной таблицы делением их на величину разрешающего элемента. В столбце  $C_0$  в строке вводимого вектора проставляют величину  $c_k$ , где  $k$  — индекс вводимого вектора.

Остальные элементы столбцов вектора  $P_0$  и  $P_j$  новой симплекс-таблицы вычисляют по правилу треугольника. Для вычисления какого-нибудь из этих элементов находят три числа:

1) число, расположенное в исходной таблице на месте искомого элемента новой симплекс-таблицы;

2) число, расположенное в исходной симплекс-таблице на пересечении строки, в которой находится искомый элемент новой симплекс-таблицы и столбца, соответствующего вектору, вводимому в базис;

3) число, расположенное в новой симплекс-таблице на пересечении столбца, в котором стоит искомый элемент, и строки вновь вводимого в базис вектора. Как отмечено выше, эта строка получается из строки исходной таблицы делением ее элементов на разрешающий элемент.

Для определения искомого элемента новой симплекс-таблицы из первого вычитают произведение второго и третьего.

После заполнения новой симплекс-таблицы просматривают элементы  $(m+1)$ -й строки. Если все  $\Delta'_j = z'_j - c_j \geq 0$ , то новый опорный план является оптимальным. Если же среди указанных чисел имеются отрицательные, то находят новый опорный план. Этот процесс продолжают до тех пор, пока либо не получат оптимальный план задачи, либо не установливают ее неразрешимость.

Напомним, что при решении задач линейного программирования предполагается, что они имеют опорные планы, и каждый такой план является невырожденным. В противном случае возможна ситуация, когда функция сохраняет свое значение в течение нескольких итераций, и так же возможен возврат к первоначальному базису. В последнем случае говорят, что произошло заикливание.

**Пример 8.6.** Для изготовления различных изделий  $A$ ,  $B$  и  $C$  предприятие использует три различных видов сырья. Нормы расхода сырья на производство одного изделия каждого вида, цена одного изделия, а так же общее количество сырья каждого вида приведены в табл. 8.4.

Таблица 8.4.

Нормы затрат и общее количество сырья

| Вид сырья           | Нормы затрат сырья на одно изделие |     |     | Общее количество сырья (ч) |
|---------------------|------------------------------------|-----|-----|----------------------------|
|                     | $A$                                | $B$ | $C$ |                            |
| I                   | 18                                 | 15  | 12  | 360                        |
| II                  | 6                                  | 4   | 8   | 192                        |
| III                 | 5                                  | 3   | 3   | 180                        |
| Цена одного изделия | 9                                  | 10  | 16  | —                          |

Составить план производства изделий, при котором общая стоимость произведенной продукции будет максимальной.

Искомый выпуск изделий  $A$  обозначим через  $x_1$ , изделий  $B$  — через  $x_2$ , изделий  $C$  — через  $x_3$ . Поскольку имеются ограничения на выделенный предприятию фонд сырья каждого вида, переменные  $x_1$ ,  $x_2$  и  $x_3$  должны удовлетворять системе неравенств

$$\begin{cases} 18x_1 + 15x_2 + 12x_3 \leq 360, \\ 6x_1 + 4x_2 + 8x_3 \leq 192, \\ 5x_1 + 3x_2 + 3x_3 \leq 180. \end{cases} \quad (8.22)$$

Общая стоимость произведенной предприятием продукции составляет

$$F = 9x_1 + 10x_2 + 16x_3. \quad (8.23)$$

По своему экономическому смыслу переменные  $x_1$ ,  $x_2$  и  $x_3$  могут принимать лишь неотрицательные значения

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \quad (8.24)$$

Запишем задачу в форме канонической. Для этого перейдем от ограничений — неравенств к ограничениям — равенствам. Введем три дополнительные переменные, в результате ограничения запишутся в виде системы уравнений

$$\begin{cases} 18x_1 + 15x_2 + 12x_3 + x_4 = 360, \\ 6x_1 + 4x_2 + 8x_3 + x_5 = 192, \\ 5x_1 + 3x_2 + 3x_3 + x_6 = 180. \end{cases} \quad (n = 6, m = 3).$$

Эти дополнительные переменные по экономическому смыслу означают не используемое при данном плане производства количество сырья того или иного типа, например,  $x_4$  — это неиспользуемое сырье I вида.

Преобразованную систему уравнений запишем в векторной форме

$$x_1 P_1 + x_2 P_2 + x_3 P_3 + x_4 P_4 + x_5 P_5 + x_6 P_6 = P_0,$$

где

$$P_1 = \begin{pmatrix} 18 \\ 6 \\ 5 \end{pmatrix}, P_2 = \begin{pmatrix} 15 \\ 4 \\ 3 \end{pmatrix}, P_3 = \begin{pmatrix} 12 \\ 8 \\ 3 \end{pmatrix}, P_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, P_5 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, P_6 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, P_0 = \begin{pmatrix} 360 \\ 192 \\ 180 \end{pmatrix}.$$

Поскольку среди векторов  $P_1, \dots, P_6$  имеются три единичных вектора, то можно непосредственно записать опорный план  $X = (0, 0, 0, 360, 192, 180)^T$ , определяемый системой трехмерных единичных векторов  $P_4, P_5, P_6$ , которые образуют базис векторного пространства.

Составим симплекс-таблицу для итерации 1 (табл. 8.5) и проверим исходный план на оптимальность:

$$F_0 = (C_0, P_0) = 0; z_1 = (C_0, P_1) = 0; z_2 = (C_0, P_2) = 0; z_3 = (C_0, P_3) = 0; \\ z_1 - c_1 = 0 - 9 = -9; z_2 - c_2 = 0 - 10 = -10; z_3 - c_3 = 0 - 16 = -16.$$

Для векторов базиса  $\Delta_j = z_j - c_j = 0$ .

Таблица 8.5.

Симплекс-таблица итерации 1

| $i$ | Базис | $C_0$ | $P_0$ | 9     | 10    | 16    | 0     | 0     | 0     |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     |       |       |       | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
| 1   | $P_4$ | 0     | 360   | 18    | 15    | 12    | 1     | 0     | 0     |
| 2   | $P_5$ | 0     | 192   | 6     | 4     | 8     | 0     | 1     | 0     |
| 3   | $P_6$ | 0     | 180   | 5     | 3     | 3     | 0     | 0     | 1     |
| 4   |       |       | 0     | -9    | -10   | -16   | 0     | 0     | 0     |

Опорный план  $X$  не является оптимальным, поскольку в четвертой строке симплекс-таблицы имеются три отрицательных числа. Это свидетельствует о возможности увеличения общей стоимости производимой продукции и показывает, на сколько она увеличится при введении в план единицы того или иного вида.

Так, число (-9) означает, что при включении в план производства одного изделия  $A$  обеспечивается увеличение выпуска продукции на 9 единиц. Поэтому с экономической точки зрения наиболее целесообразным является включение в план производства изделий  $C$  (соответствующие вектору  $P_3$ ). Это же необходимо сделать и на основании формального признака симплекс-метода, поскольку максимальное по абсолютной величине отрицательное число  $\Delta_j$  стоит в четвертой строке таблицы и соответствует вектору  $P_3$ , который следует ввести в базис. Для определения вектора, подлежащего исключению, находим  $\min(b_i / a_{i3})$  при  $a_{i3} > 0$ , то есть  $\min(360 / 12; 192 / 8; 180 / 3) = 192 / 8 = 24$ . Таким образом, из базиса следует исключить вектор  $P_5$ . Ограничивающим фактором для производства изделий  $C$  является имеющийся объем сырья II вида. С учетом этого предприятие может выпустить 24 изделия  $C$ .

Итак, столбец  $P_3$  и строка  $P_5$  являются в симплекс-таблице направляющими. Составим таблицу для итерации 2 (табл. 8.6).

Таблица 8.6.

Симплекс-таблица итерации 2

| $i$ | Базис | $C_0$ | $P_0$ | 9     | 10    | 16    | 0     | 0     | 0     |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     |       |       |       | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
| 1   | $P_4$ | 0     | 72    | 9     | 9     | 0     | 1     | -3/2  | 0     |
| 2   | $P_3$ | 16    | 24    | 3/4   | 1/2   | 1     | 0     | 1/8   | 0     |
| 3   | $P_6$ | 0     | 108   | 11/4  | 3/2   | 0     | 0     | -3/8  | 1     |
| 4   |       |       | 384   | 3     | -2    | 0     | 0     | 2     | 0     |

Сначала заполним строку для вновь введенного в базис вектора  $P_3$  (вторую строку в табл. 8.6). Элементы этой строки получаются из элементов второй строки табл. 8.5 делением их на разрешающий элемент (на 8). При этом в столбце  $C_6$  записываем коэффициент  $c_3 = 16$ , соответствующий вводимому в базис вектору  $P_3$ . Для определения остальных элементов табл. 8.6 применяем правило треугольника. Эти элементы могут быть вычислены непосредственно по формулам (8.18) – (8.21).

Вычислим, например, число, являющееся первым элементом вектора  $P_0$  в табл. 8.6. Для его вычисления находим три числа расположенные:

- 1) в табл. 8.5 на пересечении столбца вектора  $P_0$  и первой строки (360);
- 2) в табл. 8.5 на пересечении столбца вектора  $P_3$  и первой строки (12);
- 3) в табл. 8.6 на пересечении столбца вектора  $P_0$  и второй строки (24).

Вычитая из первого числа произведение двух других, находим искомый элемент:  $360 - 12 \cdot 24 = 72$ . Аналогично находим элементы столбцов  $P_1, P_2, P_5$ .

По окончании расчета всех элементов табл. 8.6 получаем новый опорный план  $X = (0, 24, 0, 72, 0, 108)^T$  и значения  $\Delta'_j$  и  $F'_0$ . При данном плане производства изготавливается 24 изделия  $C$  и остается неиспользованным 72 сырья вида I и 108 сырья вида III. Стоимость всей производимой при этом плане продукции равна 384.

Найденный на итерации 2 план также не является оптимальным, поскольку в четвертой строке табл. 8.6, на пересечении со столбцом  $P_2$  стоит отрицательное число  $-2$ . Далее в базис следует ввести вектор  $P_2$ . Найдем вектор, подлежащий исключению из базиса. Для этого находим  $\min(72/9, 48/1, 360/3) = 72/9 = 8$ . Следовательно, исключению из базиса подлежит вектор  $P_4$ . Число 9 является разрешающим элементом, а столбец  $P_2$  и строка  $P_4$  являются направляющими. Составим таблицу 8.7 для итерации 3.

Таблица 8.7.

Симплекс-таблица итерации 3

| i | Базис | $C_6$ | $P_0$ | 9     | 10    | 16    | 0     | 0     | 0     |
|---|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|   |       |       |       | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
| 1 | $P_2$ | 10    | 8     | 1     | 1     | 0     | 1/9   | -1/6  | 0     |
| 2 | $P_3$ | 16    | 20    | 1/4   | 0     | 1     | -1/18 | 5/24  | 0     |
| 3 | $P_6$ | 0     | 96    | 5/4   | 0     | 0     | -1/6  | -1/8  | 1     |
| 4 |       |       | 400   | 5     | 0     | 0     | 2/9   | 5/3   | 0     |

В табл. 8.7 сначала заполняем элементы первой строки (вектор  $P_2$ ). Элементы этой строки получаем из элементов первой строки табл. 8.6 делением на разрешающий элемент (на 9). При этом в столбце  $C_6$  записываем коэффициент  $c_2 = 10$ , соответствующий вводимому в базис вектору  $P_2$ .

Затем по правилу треугольника вычисляем элементы остальных столбцов. В результате получаем новый опорный план  $X = (0, 8, 20, 0, 0, 96)^T$  и коэффициенты разложения векторов  $P_i$  ( $i = 1, \dots, 6$ ) через базисные векторы  $P_2, P_3, P_6$ , а также значения  $\Delta''_j$  и  $F''_0$ .

Проверяем, является ли данный опорный план оптимальным. Для этого рассмотрим четвертую строку табл. 8.7. В этой строке среди чисел  $\Delta''_j$  нет отрицательных. Следовательно, найденный план является оптимальным и  $F_{\max} = 400$ .

С экономической точки зрения, план выпуска продукции, включающий изготовление 8 изделий  $B$  и 20 изделий  $C$ , является оптимальным. При данном плане выпуска полностью используется сырье видов I и II, остается неиспользованными 96 единиц сырья вида III, а стоимость произведенной продукции равна 400. Оптимальным планом не предусматривается изготовление изделий  $A$ .

**Пример 8.7.** Найти максимум функции  $F = 2x_1 - 6x_2 + 5x_3$  при условиях

$$\begin{cases} -2x_1 + x_2 + x_3 + x_5 = 20, \\ -x_1 - 2x_2 + x_4 + 3x_5 = 24, \\ 3x_1 - x_2 - 12x_5 + x_6 = 18, \\ x_j \geq 0 \quad (j = 1, \dots, 6). \end{cases}$$

Систему уравнений задачи перепишем в векторной форме

$$x_1 P_1 + x_2 P_2 + x_3 P_3 + x_4 P_4 + x_5 P_5 + x_6 P_6 = P_0,$$

где

$$P_1 = \begin{pmatrix} -2 \\ -1 \\ 3 \end{pmatrix}, P_2 = \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}, P_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, P_4 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, P_5 = \begin{pmatrix} 1 \\ 3 \\ -12 \end{pmatrix}, P_6 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, P_0 = \begin{pmatrix} 20 \\ 24 \\ 18 \end{pmatrix}.$$

Так как среди этих векторов имеются три единичных вектора, для данной задачи можно непосредственно записать опорный план  $X = (0, 0, 20, 24, 0, 18)^T$ . Составляем таблицу 8.8 и проверяем, является ли данный план оптимальным.

Таблица 8.8.

Симплекс-таблица итерации 1

| $i$ | Базис | $C_6$ | $P_0$ | 2     | -6    | 0     | 0     | 5     | 0     |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     |       |       |       | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
| 1   | $P_3$ | 0     | 20    | -2    | 1     | 1     | 0     | 1     | 0     |
| 2   | $P_4$ | 0     | 24    | -1    | -2    | 0     | 1     | 3     | 0     |
| 3   | $P_6$ | 0     | 18    | 3     | -1    | 0     | 0     | -12   | 1     |
| 4   |       |       | 0     | -2    | 6     | 0     | 0     | -5    | 0     |

Как видно из табл. 8.8, исходный опорный план не является оптимальным. Заметим, что в столбцах  $P_1$  и  $P_5$  четвертая строка содержит отрицательные числа. Кроме того, в этих столбцах имеются и положительные элементы. Поэтому мы можем перейти к новому опорному плану. Исключению из базиса подлежит вектор  $P_4$ . Для перехода к новому опорному плану введем в базис вектор  $P_5$ . Составим таблицу 8.9 для итерации 2.

Таблица 8.9.

Симплекс-таблица итерации 2

| $i$ | Базис | $C_6$ | $P_0$ | 2     | -6    | 0     | 0     | 5     | 0     |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     |       |       |       | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
| 1   | $P_3$ | 0     | 12    | -5/3  | 5/3   | 1     | -1/3  | 0     | 0     |
| 2   | $P_5$ | 5     | 8     | -1/3  | -2/3  | 0     | 1/3   | 1     | 0     |
| 3   | $P_6$ | 0     | 114   | -1    | -9    | 0     | 4     | 0     | 1     |
| 4   |       |       | 40    | -11/3 | 8/3   | 0     | 5/3   | 0     | 0     |

Из табл. 8.9 следует, что новый опорный план задачи не является оптимальным, так как в четвертой строке таблицы стоит отрицательное число: -11/3 (вектор  $P_1$ ). Поскольку среди элементов этого вектора нет положительных чисел, данная задача не имеет оптимального плана, а целевая функция бесконечно возрастает при увеличении  $x_1$ .

**Пример 8.8.** Найти максимум функции  $F = 2x_1 + x_2 - x_3 + x_4 - x_5$  при условиях

$$\begin{cases} x_1 + x_2 + x_3 = 5, \\ 2x_1 + x_2 + x_4 = 9, \\ x_1 + 2x_2 + x_5 = 7, \\ x_j \geq 0 \quad (j=1, \dots, 5). \end{cases}$$

Систему уравнений задачи запишем в векторной форме

$$x_1 P_1 + x_2 P_2 + x_3 P_3 + x_4 P_4 + x_5 P_5 = P_0,$$

$$\text{где } P_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, P_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, P_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, P_4 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, P_5 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, P_0 = \begin{pmatrix} 5 \\ 9 \\ 7 \end{pmatrix}.$$

Так как среди этих векторов имеются три единичных вектора  $P_3, P_4$  и  $P_5$ , то для данной задачи можно непосредственно написать опорный план и найти оптимальное решение симплекс-методом.

Заметим, что решение задачи симплекс-методом можно проводить, используя лишь одну сводную таблицу. В ней (табл. 8.10) последовательно записываются итерации вычисления оптимального плана.

Таблица 8.10.

Сводная симплекс-таблица

| $i$ | Базис | $C_6$ | $P_0$ | 2     | 1     | -1    | 1     | -1    |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|
|     |       |       |       | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ |
| 1   | $P_3$ | -1    | 5     | 1     | 1     | 1     | 0     | 0     |
| 2   | $P_4$ | 1     | 9     | 2     | 1     | 0     | 1     | 0     |
| 3   | $P_5$ | -1    | 7     | 1     | 2     | 0     | 0     | 1     |
| 4   |       |       | -3    | -2    | -3    | 0     | 0     | 0     |
| 1   | $P_3$ | -1    | 3/2   | 1/2   | 0     | 1     | 0     | -1/2  |
| 2   | $P_4$ | 1     | 11/2  | 3/2   | 0     | 0     | 1     | -1/2  |
| 3   | $P_2$ | 1     | 7/2   | 1/2   | 1     | 0     | 0     | 1/2   |
| 4   |       |       | 15/2  | -1/2  | 0     | 0     | 0     | 3/2   |
| 1   | $P_1$ | 2     | 3     | 1     | 0     | 2     | 0     | -1    |
| 2   | $P_4$ | 1     | 1     | 0     | 0     | -3    | 1     | 1     |
| 3   | $P_2$ | 1     | 2     | 0     | 1     | -1    | 0     | 1     |
| 4   |       |       | 9     | 0     | 0     | 1     | 0     | 1     |

Из решения с помощью табл. 8.10 следует, что  $X = (3, 2, 0, 1, 0)^T$  является оптимальным планом исходной задачи. При этом плане значение целевой функции равно  $F_{\max} = 9$ .

**Пример 8.9.** Решить задачу линейного программирования

$$F = x_1 + 9x_2 + 5x_3 + 3x_4 + 4x_5 + 14x_6 \rightarrow \min,$$

$$\begin{cases} x_1 + x_4 = 20, \\ x_2 + x_5 = 50, \\ x_3 + x_6 = 30, \\ x_4 + x_5 + x_6 = 60, \\ x_j \geq 0 \ (j=1, \dots, 6). \end{cases}$$

Равносильными преобразованиями приведем задачу к виду

$$F = -x_1 - 9x_2 - 5x_3 - 3x_4 - 4x_5 - 14x_6 \rightarrow \max,$$

$$\begin{cases} x_1 + x_2 + x_3 = 40, \\ x_1 + x_4 = 20, \\ x_3 + x_6 = 30, \\ x_5 - x_1 - x_3 = 10, \\ x_j \geq 0 \ (j=1, \dots, 6). \end{cases}$$

Полученную систему уравнений можно записать в векторной форме

$$x_1 P_1 + x_2 P_2 + x_3 P_3 + x_4 P_4 + x_5 P_5 + x_6 P_6 = P_0,$$

$$\text{где } P_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix}, P_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, P_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}, P_4 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, P_5 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, P_6 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, P_0 = \begin{pmatrix} 40 \\ 20 \\ 10 \\ 30 \end{pmatrix}.$$

Так как среди этих векторов имеются четыре единичных вектора  $P_2, P_4, P_5$  и  $P_6$ , то можно записать опорный план и найти решение симплекс-методом. Составим симплекс-таблицу (табл. 8.11).

Таблица 8.11.

Сводная симплекс-таблица

| $i$ | Базис | $C_6$ | $P_0$ | -1    | -9    | -5    | -3    | -4    | -14   |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     |       |       |       | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
| 1   | $P_2$ | -9    | 40    | 1     | 1     | 1     | 0     | 0     | 0     |
| 2   | $P_4$ | -3    | 20    | 1     | 0     | 0     | 1     | 0     | 0     |
| 3   | $P_5$ | -4    | 10    | -1    | 0     | -1    | 0     | 1     | 0     |
| 4   | $P_6$ | -14   | 30    | 0     | 0     | 1     | 0     | 0     | 1     |
| 5   |       |       | -880  | -7    | 0     | -14   | 0     | 0     | 0     |

| $i$ | Базис | $C_6$ | $P_0$ | -1    | -9    | -5    | -3    | -4    | -14   |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     |       |       |       | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
| 1   | $P_2$ | -9    | 10    | 1     | 1     | 0     | 0     | 0     | -1    |
| 2   | $P_4$ | -3    | 20    | 1     | 0     | 0     | 1     | 0     | 0     |
| 3   | $P_5$ | -4    | 40    | -1    | 0     | 0     | 0     | 1     | 1     |
| 4   | $P_3$ | -5    | 30    | 0     | 0     | 1     | 0     | 0     | 1     |
| 5   |       |       | -460  | -7    | 0     | 0     | 0     | 0     | 14    |
| 1   | $P_1$ | -1    | 10    | 1     | 1     | 0     | 0     | 0     | -1    |
| 2   | $P_4$ | -3    | 10    | 0     | -1    | 0     | 1     | 0     | 1     |
| 3   | $P_5$ | -4    | 50    | 0     | 1     | 0     | 0     | 1     | 0     |
| 4   | $P_3$ | -5    | 30    | 0     | 0     | 1     | 0     | 0     | 1     |
| 5   |       |       | -390  | 0     | 7     | 0     | 0     | 0     | 7     |

Так как в четвертой строке нет отрицательных элементов, то полученный план  $X = (10, 0, 30, 10, 50, 0)^T$  является оптимальным и значение целевой функции  $F = 390$ .

### Метод искусственного базиса

При описании симплекс-метода было показано, что для задачи линейного программирования, записанной в форме канонической, можно непосредственно указать ее опорный план только в том случае, если среди векторов  $P_j$  имеется  $m$  единичных. Когда это не выполняется, то следует применять метод искусственного базиса.

Пусть требуется найти максимум функции

$$F = c_1 x_1 + c_2 x_2 + \dots + c_n x_n \quad (8.25)$$

при условиях

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + \dots + a_{2n}x_n = b_2, \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m, \end{cases} \quad (8.26)$$

$$x_j \geq 0 \ (j=1, \dots, n) \quad (8.27)$$



$$\text{и среди векторов } P_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix}, \quad P_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \dots \\ a_{m2} \end{pmatrix}, \quad \dots \quad P_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{pmatrix}$$

нет  $m$  единичных.

**Определение.** По отношению к задаче (8.25) — (8.27) нахождение максимума функции

$$F = -Mx_1 - \dots - Mx_m + c_{m+1}x_{m+1} + c_{m+2}x_{m+2} + \dots + c_{m+n}x_{m+n} \quad (8.28)$$

при условиях

$$\begin{cases} x_1 + a_{1m+1}x_{m+1} + \dots + a_{1m+n}x_{m+n} = b_1, \\ x_2 + a_{2m+1}x_{m+1} + \dots + a_{2m+n}x_{m+n} = b_2, \\ \dots \\ x_m + a_{mm+1}x_{m+1} + \dots + a_{mm+n}x_{m+n} = b_m, \end{cases} \quad (8.29)$$

$$x_j \geq 0 \quad (j = 1, \dots, n+m) \quad (8.30)$$

где  $M$  — некоторое достаточно большое положительное число (конкретное значение которого обычно не задается) называется *расширенной задачей*.

Расширенная задача (8.28) — (8.30) может быть решена симплекс-методом, поскольку записана в форме канонической задачи и имеет  $m$  единичных векторов  $P_1, \dots, P_m$ , которые определяют исходный опорный план  $X = (b_1, \dots, b_m, \underbrace{0, 0, \dots, 0}_n)^T$ .

**Определение.** Вектора  $P_1, \dots, P_m$ , переменные  $x_1, \dots, x_m$  и базис, образованный этими векторами называют *искусственными*.

Справедливо следующее утверждение.

Если в оптимальном плане  $X^* = (x_1^*, \dots, x_m^*, x_{m+1}^*, \dots, x_{m+n}^*)^T$  расширенной задачи (8.28) — (8.30) значения искусственных переменных  $(x_1^*, \dots, x_m^*)$  равны нулю, то план  $X = (x_{m+1}^*, \dots, x_{m+n}^*)^T$  является оптимальным для задачи (8.25) — (8.27).

Рассмотрим подробно решение расширенной задачи.

Построим итерационный процесс, на каждом шаге которого будем по очереди исключать из базиса искусственные

вектора, обнуляя при этом соответствующие искусственные переменные.

Исходное значение целевой функции  $F = -M \sum_{i=1}^m x_i + \sum_{i=m+1}^{n+m} c_i x_i$ .

На итерации 1 вместо исключенной *искусственной* переменной  $x_1$  в опорный план записывается одна из *неискусственных* переменных  $x_j$ , где  $(j = m+1, \dots, m+n)$ . Искусственная переменная  $x_1$  обнуляется и поэтому в записи функции опускается.

В результате получим  $F^1 = c_1 x_1 - M \sum_{i=2}^m x_i + \sum_{j=m+1}^{n+m-1} c_j x_j$ ,

где общее число переменных стало на единицу меньше. Продолжая этот процесс далее, на итерации  $k$  целевую функцию можно представить в виде

$$\begin{aligned} F^k &= \sum_{i=1}^k c_i x_i - M \sum_{i=k+1}^m x_i + \sum_{j=m+1}^{n+m-k} c_j x_j = \\ &= \sum_{i=1}^k c_i (b_i - \sum_{j=m+1}^{n+m-k} a_{ij} x_j) - M \sum_{i=k+1}^m (b_i - \sum_{j=m+1}^{n+m-k} a_{ij} x_j) + \sum_{j=m+1}^{n+m-k} c_j x_j = \\ &= \sum_{i=1}^k c_i b_i - \sum_{i=1}^k c_i \sum_{j=m+1}^{n+m-k} a_{ij} x_j - M \sum_{i=k+1}^m b_i + \sum_{i=k+1}^m M \sum_{j=m+1}^{n+m-k} a_{ij} x_j + \sum_{j=m+1}^{n+m-k} c_j x_j = \\ &= (\sum_{i=1}^k c_i b_i - M \sum_{i=k+1}^m b_i) - (\sum_{i=1}^k c_i \sum_{j=m+1}^{n+m-k} a_{ij} x_j - \sum_{i=k+1}^m M \sum_{j=m+1}^{n+m-k} a_{ij} x_j - \sum_{j=m+1}^{n+m-k} c_j x_j) = \\ &= \sum_{i=1}^k c_i b_i - M \sum_{i=k+1}^m b_i - \sum_{j=m+1}^{n+m-k} \left( \sum_{i=1}^k c_i a_{ij} - M \sum_{i=k+1}^m a_{ij} - c_j \right) x_j = F_0^k - \sum_{j=m+1}^{n+m-k} \Delta_j x_j, \end{aligned}$$

$$\text{где } F_0^k = \sum_{i=1}^k c_i b_i - M \sum_{i=k+1}^m b_i, \text{ а } \Delta_j = z_j - c_j = \sum_{i=1}^k c_i a_{ij} - M \sum_{i=k+1}^m a_{ij} - c_j.$$

Таким образом,  $F_0^k$  и  $\Delta_j$  состоят каждая из двух независимых частей, одна из которых зависит от параметра  $M$ , а другая — нет.

На каждой итерации заполняется расширенная симплекс-таблица, содержащая на одну строку больше, чем обычная. Первые  $m$  строк заполняются обычным образом. В  $(m+1)$ -ю строку записываются независимые от параметра  $M$  части  $F_0^k$  и  $\Delta_j$ . В последнюю,  $(m+2)$ -ю, строку записывается множитель при параметре  $M$  из  $F_0^k$  и  $\Delta_j$ .

При переходе от одного опорного плана к другому в базис вводят вектор, соответствующий наибольшему по абсолютной величине отрицательному числу из  $(m+2)$ -й строки. Искусственный вектор, исключенный из базиса в результате некоторой итерации, в дальнейшем нецелесообразно вводить в последующие базисы, поэтому, преобразование его излишне.

Пересчет таблиц при переходе от одного опорного плана к другому производят по общим правилам симплекс-метода.

Итерационный процесс по  $(m+2)$ -й строке ведут до тех пор, пока:

1) либо все искусственные векторы не будут исключены из

базиса, то есть значение линейной формы станет  $F_0 = \sum_{i=1}^m c_i b_i$ ,

а значения  $\Delta_j = \sum_{i=1}^m c_i a_{ij} - c_j$ . Следовательно, базис отвечает

некоторому опорному плану и определение оптимального плана продолжают по  $(m+1)$ -й строке.

2) либо не все искусственные векторы исключены,  $(m+2)$ -я строка не содержит больше отрицательных элементов в столбцах  $P_1, \dots, P_{n+m-k}$ . В этом случае, если элемент, стоящий в  $(m+2)$ -й строке  $P_0$ , отрицателен, то исходная задача не имеет решения, если он равен нулю, то найденный опорный план исходной задачи является вырожденным и базис содержит по крайней мере один из векторов искусственного базиса.

**Пример 8.10.** Найти минимум функции  $F = -2x_1 + 3x_2 - 6x_3 - x_4$  при условиях

$$\begin{cases} 2x_1 + x_2 - 2x_3 + x_4 = 24, \\ x_1 + 2x_2 + 4x_3 \leq 22, \\ x_1 - x_2 + 2x_3 \geq 10, \\ x_j \geq 0 \quad (j = 1, \dots, 4). \end{cases}$$

Запишем задачу в форме канонической: найти максимум функции  $F = 2x_1 - 3x_2 + 6x_3 + x_4$  при условиях

$$\begin{cases} 2x_1 + x_2 - 2x_3 + x_4 = 24, \\ x_1 + 2x_2 + 4x_3 + x_5 = 22, \\ x_1 - x_2 + 2x_3 - x_6 = 10, \\ x_j \geq 0 \quad (j = 1, \dots, 6). \end{cases}$$

В системе уравнений последней задачи рассмотрим векторы из коэффициентов при неизвестных

$$P_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, P_2 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, P_3 = \begin{pmatrix} -2 \\ 4 \\ 2 \end{pmatrix}, P_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, P_5 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, P_6 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$

Среди этих векторов только два единичных  $P_5, P_6$ . Поэтому в левую часть третьего уравнения системы ограничений задачи добавим дополнительную неотрицательную переменную  $x_7$  и рассмотрим расширенную задачу, состоящую в максимизации функции  $F = -2x_1 + 3x_2 - 6x_3 - x_4 - Mx_7$  при условиях

$$\begin{cases} 2x_1 + x_2 - 2x_3 + x_4 = 24, \\ x_1 + 2x_2 + 4x_3 + x_5 = 22, \\ x_1 - x_2 + 2x_3 - x_6 + x_7 = 10, \\ x_j \geq 0 \quad (j = 1, \dots, 7). \end{cases}$$

Расширенная задача имеет опорный план  $X = (0, 0, 0, 24, 22, 0, 10)^T$ , определяемый системой трех единичных векторов  $P_5, P_6, P_7$ .

Составим таблицу итерации 1 (табл. 8.12), содержащую пять строк. Для заполнения четвертой и пятой строк найдем  $F_0$  и значения разностей  $\Delta_j = z_j - c_j$  ( $j = 1, \dots, 7$ ):

$$\begin{aligned} F_0 &= 24 - 10M, \quad z_1 - c_1 = 0 - M, \quad z_2 - c_2 = 4 + M, \quad z_3 - c_3 = -8 - 2M, \quad z_4 - c_4 = 0, \\ z_5 - c_5 &= 0, \quad z_6 - c_6 = 0 + M, \quad z_7 - c_7 = 0. \end{aligned}$$

Таблица 8.12.

Симплекс-таблица итерации 1

| $i$ | Базис | $C_0$ | $P_0$ | 2     | -3    | 6     | 1     | 0     | 0     | -M    |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     |       |       |       | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ | $P_7$ |
| 1   | $P_4$ | 1     | 24    | 2     | 1     | -2    | 1     | 0     | 0     | 0     |
| 2   | $P_5$ | 0     | 22    | 1     | 2     | 4     | 0     | 1     | 0     | 0     |
| 3   | $P_7$ | -M    | 10    | 1     | -1    | 2     | 0     | 0     | -1    | 1     |
| 4   |       |       | 24    | 0     | 4     | -8    | 0     | 0     | 0     | 0     |
| 5   |       |       | -10   | -1    | 1     | -2    | 0     | 0     | 1     | 0     |

Значения  $F_0$  и  $\Delta_j = z_j - c_j$  состоят из двух слагаемых, одно из которых содержит  $M$ , а другое — нет. Для удобства итерационного процесса число, стоящее при  $M$ , записываем в пятой строке, а слагаемое, не содержащее  $M$ , — в четвертой строке.

В пятой строке таблицы имеются два отрицательных числа, наличие которых говорит о том, что данный опорный план расширенной задачи не является оптимальным. Переходим к новому опорному плану расширенной задачи. В базис вводим вектор  $P_3$ . Чтобы определить вектор, исключаемый из базиса, находим  $\min(22/4; 10/2) = 10/2$ , следовательно, таким вектором оказывается  $P_7$ . Этот вектор не имеет смысла вводить ни в один из последующих базисов, поэтому в дальнейшем столбец, соответствующий данному вектору, не заполняется.

Составим таблицу итерации 2 (табл. 8.13), содержащую только четыре строки, так как искусственный вектор из базиса исключен.

Таблица 8.13.

Симплекс-таблица итерации 2

| $i$ | Базис | $C_6$ | $P_0$ | 2     | -3    | 6     | 1     | 0     | 0     |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     |       |       |       | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
| 1   | $P_4$ | 1     | 34    | 3     | 0     | 0     | 1     | 0     | -1    |
| 2   | $P_5$ | 0     | 2     | -1    | 4     | 0     | 0     | 1     | 2     |
| 3   | $P_3$ | 6     | 5     | 1/2   | -1/2  | 1     | 0     | 0     | -1/2  |
| 4   |       |       | 64    | 4     | 0     | 0     | 0     | 0     | -4    |

Из табл. 8.13 следует, что для исходной задачи опорным является план  $X = (0, 0, 5, 34, 2, 0)^T$ . Для проверки его на оптимальность рассмотрим элементы четвертой строки. В столбце  $P_6$  имеется отрицательное число, следовательно, опорный план может быть улучшен благодаря введению в базис вектора  $P_6$ , а из базиса исключается вектор  $P_5$ . Составим таблицу итерации 3.

Таблица 8.14.

Симплекс-таблица итерации 3

| $i$ | Базис | $C_6$ | $P_0$ | 2     | -3    | 6     | 1     | 0     | 0     |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     |       |       |       | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |
| 1   | $P_4$ | 1     | 35    | 5/2   | 2     | 0     | 1     | 1/2   | 0     |
| 2   | $P_6$ | 0     | 1     | -1/2  | 2     | 0     | 0     | 1/2   | 1     |
| 3   | $P_3$ | 6     | 11/2  | 1/4   | 1/2   | 1     | 0     | 1/4   | 0     |
| 4   |       |       | 68    | 2     | 8     | 0     | 0     | 2     | 0     |

В четвертой строке табл. 8.14 среди чисел  $\Delta_j$  нет отрицательных, что означает оптимальность опорного плана исходной задачи  $X = (0, 0, 11/2, 35, 0, 1)^T$  и целевая функция  $F = 68$ .

**Пример 8.11.** Найти минимум функции  $F = 2x_1 - x_2 - x_4$  при условиях

$$\begin{cases} x_1 - 2x_2 + x_3 = 10, \\ -2x_1 - x_2 - 2x_4 \geq 18, \\ 3x_1 + 2x_2 + x_4 \geq 36, \\ x_j \geq 0 \quad (j = 1, \dots, 4). \end{cases}$$

Запишем задачу в форме канонической задачи линейного программирования: найти максимум функции  $F = -2x_1 + x_2 + x_4$  при условиях

$$\begin{cases} x_1 - 2x_2 + x_3 = 10, \\ -2x_1 - x_2 - 2x_4 - x_5 = 18, \\ 3x_1 + 2x_2 + x_4 - x_6 = 36, \\ x_j \geq 0 \quad (j = 1, \dots, 6). \end{cases}$$

Так как среди векторов

$$P_1 = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, P_2 = \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}, P_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, P_4 = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix}, P_5 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, P_6 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}$$

имеется только один единичный вектор  $P_3$ , то ищем решение расширенной задачи, состоящей в определении максимального значения функции

$$F = -2x_1 + x_2 + x_4 - Mx_7 - Mx_8$$

при условиях

$$\begin{cases} x_1 - 2x_2 + x_3 = 10, \\ -2x_1 - x_2 - 2x_4 - x_5 + x_7 = 18, \\ 3x_1 + 2x_2 + x_4 - x_6 + x_8 = 36, \\ x_j \geq 0 \quad (j = 1, \dots, 8). \end{cases}$$

Расширенная задача имеет опорный план

$$X = (0, 0, 10, 0, 0, 0, 18, 36)^T.$$

Составим таблицу для итерации 1.

Таблица 8.15.

Симплекс-таблица итерации 1

| $i$ | Базис | $C_6$ | $P_0$ | -2    | 1     | 0     | 1     | 0     | 0     | -M    | -M    |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     |       |       |       | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ | $P_7$ | $P_8$ |
| 1   | $P_3$ | 0     | 10    | 1     | -2    | 1     | 0     | 0     | 0     | 0     | 0     |
| 2   | $P_7$ | -M    | 18    | -2    | -1    | 0     | -2    | -1    | 0     | 1     | 0     |
| 3   | $P_8$ | -M    | 36    | 3     | 2     | 0     | 1     | 0     | -1    | 0     | 1     |
| 4   |       |       | 0     | 2     | -1    | 0     | -1    | 0     | 0     | 0     | 0     |
| 5   |       |       | -54   | -1    | -1    | 0     | 1     | 1     | 1     | 0     | 0     |

В пятой строке табл. 8.15 имеются отрицательные числа, поэтому следует перейти к новому опорному плану расширенной задачи. В базис введем вектор  $P_2$ , а из базиса исключим вектор  $P_8$ . Этот вектор не имеет смысла вводить ни в один из последующих базисов, поэтому столбец, соответствующий ему, не заполняется и в таблице отсутствует.

Таблица 8.16.

Составим таблицу итерации 2.

| $i$ | Базис | $C_6$ | $P_0$ | -2    | 1     | 0     | 1     | 0     | 0     | -M    |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|     |       |       |       | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ | $P_7$ |
| 1   | $P_3$ | 0     | 46    | 4     | 0     | 1     | 1     | 0     | -1    | 0     |
| 2   | $P_7$ | -M    | 36    | -1/2  | 0     | 0     | -3/2  | -1    | -1/2  | 1     |
| 3   | $P_2$ | 1     | 18    | 3/2   | 1     | 0     | 1/2   | 0     | -1/2  | 0     |
| 4   |       |       | 18    | 7/2   | 0     | 0     | -1/2  | 0     | -1/2  | 0     |
| 5   |       |       | -36   | 1/2   | 0     | 0     | 3/2   | 1     | 1/2   | 0     |

В пятой строке табл. 8.16 в столбцах  $P_1, \dots, P_7$  не содержится отрицательных элементов. На пересечении столбца  $P_0$  с этой строкой стоит отрицательное число -36, следовательно, исходная задача не имеет опорного плана.

## Вопросы и задания для самоконтроля

1. Какая задача оптимизации называется задачей линейного программирования?
2. Дать определение общей задачи линейного программирования.
3. Дать определение канонической задачи линейного программирования.
4. Описать алгоритм сведения общей задачи к задаче в канонической форме линейного программирования. Привести пример.
5. Какие задачи линейного программирования можно решить графически?
6. Описать алгоритм графического решения задачи линейного программирования.
7. Дать определение плана и оптимального плана задачи линейного программирования.
8. Какую задачу линейного программирования можно решить с помощью симплекс-метода?
9. Дать определение опорного плана задачи линейного программирования.
10. Сформулировать признак оптимальности опорного плана для задачи линейного программирования симплексным методом.
11. Показать, что если для некоторого  $\Delta_k < 0$  среди чисел  $a_{ik}$  ( $i = 1, \dots, m$ ) нет положительных, то целевая функция задачи линейного программирования в канонической форме (8.15) — (8.17) не ограничена на множестве ее планов.
12. Сформулировать правила пересчета ограничений задачи линейного программирования при переходе к новому базису. Привести пример.
13. Всегда ли решение задачи линейного программирования, записанной в канонической форме (8.15) — (8.17), может быть найдено за конечное число шагов?
14. Каким образом можно улучшить приведенный алгоритм решения задачи линейного программирования симплексным методом?
15. В каких случаях для решения задачи линейного программирования необходимо добавлять искусственные переменные? Привести пример.
16. Показать, что для метода искусственного базиса величины  $F_0^k$  и  $\Delta_j$  зависят от параметра  $M$ .

17. В каких случаях метод искусственного базиса не даст решения задачи линейного программирования?

18. Описать алгоритм расчета задачи линейного программирования для метода искусственного базиса.

19. Решить задачи линейного программирования графическим методом

Задача 1.

$$\begin{aligned} f(x_1, x_2) &= -x_1 - 4x_2 \rightarrow \min, \\ x_1 &\leq 2, \quad x_1 + 2x_2 \geq 2, \\ x_2 &\leq 2, \quad x_1 + x_2 \leq 3, \\ x_1, x_2 &\geq 0. \end{aligned}$$

Ответ:  $x^* = (1, 2)^T$ ;  $f^* = -9$ .

Задача 2.

$$\begin{aligned} f(x_1, x_2) &= -x_1 - x_2 \rightarrow \min, \\ x_1 &\leq 3, \quad x_2 \leq 2, \\ x_1 + x_2 &\leq 1, \\ x_1, x_2 &\geq 0. \end{aligned}$$

Ответ: Бесконечное множество решений:

$$x^* = (\alpha, 1 - \alpha)^T, \quad \alpha \in [0, 1]; \quad f^* = 1.$$

Задача 3.

$$\begin{aligned} f(x_1, x_2) &= -2x_1 - x_2 \rightarrow \min, \\ 2x_1 + x_2 &\geq 1, \quad 3x_1 - x_2 \geq -1, \\ x_1 - 4x_2 &\leq 2, \\ x_1, x_2 &\geq 0. \end{aligned}$$

Ответ: Нет решений.

Задача 4.

$$\begin{aligned} f(x_1, x_2) &= -x_1 - 3x_2 \rightarrow \min, \\ 2x_1 + x_2 &\leq 2, \quad x_1 - x_2 \geq 0, \\ x_1 - x_2 &\leq 1, \\ x_1, x_2 &\geq 0. \end{aligned}$$

Ответ:  $x^* = (2/3, 2/3)^T$ ;  $f^* = -8/3$ .

Задача 5.

$$\begin{aligned} f(x_1, x_2) &= x_1 - 2x_2 \rightarrow \min, \\ -x_1 + x_2 &\leq 0, \quad 2x_1 + x_2 \leq 3, \\ x_1 - x_2 &\leq 1, \\ x_1, x_2 &\geq 0. \end{aligned}$$

Ответ:  $x^* = (1, 1)^T$ ;  $f^* = -1$ .

## Литература

1. *Аттеков А. В.* Методы оптимизации / А. В. Аттеков, С. В. Галкин, В. С. Зарубин. М.: изд-во МГТУ им Н. Э. Баумана, 2003. 440с.
2. *Васильев, Ф. П.* Численные методы решения экстремальных задач / Ф. П. Васильев. М.: Наука, 1988. 552с.
3. *Измаилов, А. Ф.* Численные методы оптимизации / А. Ф. Измаилов, М. В. Солодов. М.: ФИЗМАТЛИТ, 2005. 304с.
4. *Лесин, В. В.* Математическое программирование / В. В. Лесин, Ю. П. Лисовец. М.: изд. МИЭТ, 1988. 116с.
5. *Лесин, В. В.* Основы методов оптимизации / В. В. Лесин, Ю. П. Лисовец. М.: изд. МАИ, 1995. 344с.
6. Методы оптимизации / под общ. ред. А. В. Ефимова и А. С. Поспелова // Сборник задач по математике для втузов. – В 4-х ч. М.: ФИЗМАТЛИТ, 2002. Ч. 3. – 576с.
7. *Пантелеев, А. В.* Методы оптимизации в примерах и задачах / А. В. Пантелеев, Т. А. Летова. М.: Высшая школа, 2005. 544с.
8. *Сухарев А. Г.* Курс методов оптимизации / А. Г. Сухарев, А. В. Тимохов, В. В. Федоров. М.: ФИЗМАТЛИТ, 2005. 368с.

*Учебное издание*

**В. А. Гончаров**

## Методы оптимизации

УЧЕБНОЕ ПОСОБИЕ

Руководитель направления  
Редактор  
Корректор  
Компьютерная верстка

Формат 84×108 <sup>1</sup>/<sub>32</sub>.  
Гарнитура «PetersburgC». Печать офсетная.  
Усл. печ. л. . Тираж экз. Заказ №

**Издательство «Высшее образование»**

140004, Московская обл., г. Люберцы, 1-й Панковский проезд, д. 1.  
Тел.: (495) 744-00-12. E-mail: publish@urait.ru. www.urait.ru



**стр. 206-208 реклама**

